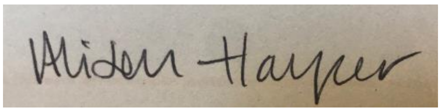
A Hybrid Modelling Framework for Real-time Decision-support for Urgent and Emergency Healthcare

Submitted by Alison Harper, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Management Studies, September 2020.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signed: 

Abstract

In healthcare, opportunities to use real-time data to support quick and effective decision-making are expanding rapidly, as data increases in volume, velocity and variety. In parallel, the need for short-term decision-support to improve system resilience is increasingly relevant, with the recent COVID-19 crisis underlining the pressure that our healthcare services are under to deliver safe, effective, quality care in the face of rapidly-shifting parameters.

A real-time hybrid model (HM) which combines real-time data, predictions, and simulation, has the potential to support short-term decision-making in healthcare. Considering decision-making as a consequence of situation awareness focuses the HM on what information is needed where, when, how, and by whom with a view toward sustained implementation. However the articulation between real-time decision-support tools and a sociotechnical approach to their development and implementation is currently lacking in the literature.

Having identified the need for a conceptual framework to support the development of real-time HMs for short-term decision-support, this research proposed and tested the Integrated Hybrid Analytics Framework (IHAF) through an examination of the stages of a Design Science methodology and insights from the literature examining decision-making in dynamic, sociotechnical systems, data analytics, and simulation. Informed by IHAF, a HM was developed using real-time Emergency Department data, time-series forecasting, and discrete-event simulation. The application started with patient questionnaires to support problem definition and to act as a formative evaluation, and was subsequently evaluated using staff interviews.

Evaluation of the application found multiple examples where the objectives of people or sub-systems are not aligned, resulting in inefficiencies and other quality problems, which are characteristic of complex adaptive sociotechnical systems. Synthesis of the literature, the formative evaluation, and the final evaluation found significant themes which can act as antecedents or evaluation criteria for future real-time HM studies in sociotechnical systems, in particular in healthcare. The generic utility of IHAF is emphasised for supporting future applications in similar domains.

Acknowledgements

The completion of this thesis has, in a large part, been thanks to those who have challenged, supported, and encouraged me over the last few years. First and foremost, I would like to express sincere gratitude to my supervisors, Prof. Nav Mustafee and Prof. Martin Pitt. It has been a privilege to work with them both.

I have been tremendously fortunate to have Prof. Nav Mustafee as my supervisor. His support, guidance and unending patience kept me going on this journey. Nav's belief in me has been unwavering, and his selfless and generous supervision will always be remembered. I would like to thank him for his friendship, kindness and sense of humour, which made difficult times easier and easier times enjoyable.

Prof. Martin Pitt has been a steady source of wisdom, support and inspiration. I want to thank him for his positivity, motivation and understanding. Martin's feedback, helpful ideas and expert advice have guided me, and I gratefully acknowledge his support, and encouragement. I would also like to thank him for believing in me, and for his warm welcome and kindness in my new role.

Special thanks are due to our NHS partners, in particular to Susan Martin and Andrew Fordyce, for their time, enthusiasm, facilitation and willing assistance, and to those who supported data collection. I would like to thank my colleagues and fellow researchers in CSAM, for their support and encouragement over the years, and to the team at PenCHORD who have always made me feel welcome. Thanks are due to Prof. John Powell, a kind mentor, to Prof. Mike Yearworth, for his guidance and encouragement, and to Prof. Mickey Howard, Prof. Andi Smart and Dr Sarah Hartley, for their welcome advice.

I wouldn't have got here without the encouragement, love and endless tolerance of my family, so thank you to my husband Richard, and my children Joseph, Olive, and Ivy for being the best. I hope I haven't put you off academia, kids! To my friends who have been behind me all the way, I am very grateful, there are too many of you to name. I would like to thank my friends in CSAM whose personal support and humour I could rely on, in particular Dr Pikakshi Manchanda and Adela Boak.

Most importantly, to my parents, who always taught me that no value can be put on education, I am eternally grateful. My love of learning and a good challenge comes from them both. My mother Bridie has modelled resilience, positivity and humour my whole life. Despite everything, her emotional support and encouragement has remained solid.

Finally, I gratefully acknowledge the financial support that has been made available to fund this research. This has primarily been via a joint studentship from the South West Doctoral Training College, and the University of Exeter Business School, to whom I extend my appreciation. In addition, the following funding is acknowledged: South West Academic Health Sciences Network, Torbay Medical Research Funds, ACM/SigSim Travel Award, IDSAI Early Career Travel Fund.

Table of Contents

List of Illustrations

List of Tables

Publications related to this thesis:

Tolk, A., **Harper, A**., and Mustafee, N. “Hybrid Models as Transdisciplinary Research Enablers” *European Journal of Operational Research. (R&R)*

Harper, A., Mustafee, N., and Yearworth, M. “Facets of Trust in Simulation Studies” *European Journal of Operational Research.* doi.org/10.1016/j.ejor.2020.06.043

Mustafee, N., **Harper, A.** and Onggo, S. (2020) “Hybrid Modelling and Simulation (M&S): Driving Innovation in the Theory and Practice of M&S”. Accepted in: *Proceedings of 2020 Winter Simulation Conference.*

Harper, A. and Mustafee, N. (2019) “A Hybrid Modelling Approach using Forecasting and Real-Time Simulation to Prevent Emergency Department Overcrowding.” In: *Proceedings of 2019 Winter Simulation Conference,* December 3-6, Washington, IEEE.

Harper, A. and Mustafee, N. (2019) “Proactive Service Recovery in Emergency Departments: A Hybrid Modelling Approach using Forecasting and Real-time Simulation” In *Proceedings of the 2019 SigSim PADS Conference, June 2-5*, Chicago, USA, IEEE.

Mustafee, N., Powell, J.H. and **Harper, A.** (2018) “RH-RT: A Data Analytics Framework for Reducing Wait Time at Emergency Departments and Centres for Urgent Care”. In: *Proceedings of 2018 Winter Simulation Conference,* December 9-12, Gothenburg, Sweden, IEEE.

Mustafee, N., Powell, J.H. and **Harper, A.** (2018) “Right hospital-right time: A business analytics framework for analyzing urgent care/A&E wait time data”. In *Proceedings of the 2018 Operational Research Society Simulation Workshop,* April 11-13, Stratford, Worcs.

Harper, A. and Mustafee, N. (2017). “A Hybrid Approach using Discrete-Event Simulation and Forecasting for Endoscopy Services” In: *Proceedings of 2017 Winter Simulation Conference,* December 3-6, Las Vegas, Nevada, IEEE.

Mustafee, N., Powell, J.H., Martin, S., Fordyce, A. and **Harper, A.** (2017). “Investigating the Use of Real-time Data in Nudging Patients’ Emergency Department Attendance Behaviour” In *Proceedings of the 2017 Spring Simulation Multi-Conference*, April 23-26, Virginia Beach, VA.

Abbreviations

ABS	Agent-Based Simulation
ACF	Autocorrelation Function
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
PACF	Partial Autocorrelation Function
StD	Standard Deviation
A&E	Accident & Emergency
ADF	Augmented Dickey Fuller Test
AIC	Akaike's Information Criterion
CDM	Critical Decision Method
CDU	Clinical Decision Unit
CINAHL	Cumulative Index To Nursing And Allied Health Literature
CS	Computer Science
DA	Data Analytics
DES	Discrete-Event Simulation
DDDAS	Dynamic Data-Driven Application Systems
DSA	Distributed Situation Awareness
DV	Dependent Variable
ED	Emergency Department
ED-RAG	ED Resilience Assessment Grid
GP	General Practice/General Practitioner
HCP	Health Care Professional
HF	Human Factors
HM	Hybrid Modelling
HS	Hybrid Simulation
IHAF	Integrated Hybrid Analytics Framework
IMPACT Network	Information, Modelling, Prediction and Evaluation to Inform ACTion) Network
IoT	Internet of Things
IS	Information Systems
IT	Information Technology
IV	Independent Variable
KPI	Key Performance Indicator
LoS	Length of Stay
LWBS	Leave without Being Seen
M&S	Modelling and Simulation
MA	Moving Average
MIU	Minor Injury Unit
ML	Machine Learning
MSE	Mean Squared Error
MCS	Montecarlo Simulation
NHS	National Health Service
OPEL	Operational Pressure Escalation Level
OR	Operational Research
PLE	Personal Learning Edition
PI	Prediction Interval
QI	Quality Improvement
RAG	Resilience Assessment Grid
RFID	Radio Frequency Identification
RMSE	Root Mean Squared Error
RQ	Research Question
SA	Situation Awareness
SAGAT	Situation Awareness and Global Assessment Tool

SARIMA	Seasonal ARIMA
SCM	Supply Chain Management
SPAM	Situation Present Assessment Method
SD	Systems Dynamics
SE	Standard Error
STS	Sociotechnical System
TAM	Technology Acceptance Model
UCC	Urgent Care Centre
UCN	Urgent Care Network
UK	United Kingdom
UTAUT	Unified Theory of Acceptance and Use of Technology
UTC	Urgent Treatment Centre
V&V	Verification & Validation
VARMA	Vector Autoregressive Moving Average
WIC	Walk-In Centre

Chapter 1: Introduction

1.1 Background

Healthcare is a basic need in any society, however the provision of healthcare increasingly faces enormous challenges. In the United Kingdom (UK), these include expanding costs, an ageing population, new disease patterns associated with wealth and human behaviour, and changes in patient expectations and actions. Managing limited healthcare resources is further challenged by variability in demand, which can lead to unbalanced utilisation of resources. Emergency care in particular suffers from high variability and the UK National Health Service (NHS) has been under sustained pressure with unpredictable demand surges, in particular during the winter months when emergency hospital admissions rise (British Medical Association, 2019). A goal of healthcare policy is to move more acute healthcare provision into community settings, and deliver the “right care at the right time in the right place” (NHS, 2019). This is often assumed to bring benefits such as reduced costs, improved access to services and improved operational performance in relation to quality of care and time (Munton et al., 2011), however evaluating this is difficult. A further assumption is that levelling demand and capacity across services will improve patient satisfaction in terms of reducing waiting times and perceived quality of services (Abo-Hamad & Arisha, 2013; Zhao et al., 2015). This is also relevant for operational performance against targets, and may support objectives such as improving the way Emergency Departments (ED) work within the wider urgent care network, and encouraging patients to access alternatives to emergency care where appropriate (Murray et al., 2018). This chapter sets the context, rationale and scope of this research, with this section outlining the status of healthcare in the UK, albeit prior to the acute worldwide COVID-19 crisis. The applicability and challenges of simulation modelling in healthcare, a complex system, are appraised.

Healthcare systems are complex social systems, with non-linear processes and unpredictable outcomes, challenging operational decision-making and evaluation of interventions. Simulation modelling is an effective decision-support tool for complex systems (Jahangirian et al., 2010; Marshall et al., 2016; Zhang et al., 2020), allowing an understanding of the interdependencies between human and system variables (Almagooshi, 2015). System Dynamics (SD), agent-based simulation (ABS), Monte-Carlo simulation (MCS) and discrete-event simulation (DES) are the most commonly used simulation methods in Operations Research (OR), although other methods are used. In healthcare, simulation modelling has been used for decades, aiming to improve outcomes, evidence changes in delivery, and reduce costs (Katsaliaki & Mustafee, 2011; Fakhimi & Probert, 2013).

Despite the large number of studies applying simulation in healthcare, considerable challenges exist and its application is far from routine in practice. Reviews of simulation modelling in healthcare highlight deficiencies in research design (Aboueljinane et al., 2013; Mohiuddin et al., 2017; Zhang et al., 2020), alongside an ongoing interest in the challenges of conducting simulation studies in healthcare (Brailsford et al., 2013; Jahangirian et al., 2015; Klein & Young, 2015; Tako & Robinson, 2015; Long, McDermott & Meadows, 2019). These centre on the need for increased stakeholder engagement and the difficulty of accessing data, alongside messy problems and rapid organisational change, and have resulted in low levels of real-world implementation of the results (Katsaliaki & Mustafee, 2011; Jahangirian, 2016; Pitt et al., 2016). The number of reviews of simulation in healthcare published in the last decade suggest that applications are rapidly rising. For example, an umbrella review by Salleh et al. (2017) synthesised 37 reviews of healthcare simulation modelling, of which 30 were published since 2010. Of these, 21 were focussed on operational performance, and five reviews were specific to emergency care. Yet Mohiuddin et al. (2017) reported that only 14% of results from simulation studies in their review were implemented, while Katsaliaki and Mustafee (2011) found that just over 5% of published papers in healthcare simulation modelling reported evidence of real-world implementation of results.

Despite the challenges, the potential value and impact of simulation methods for healthcare operational improvement remains undisputed (Pitt et al., 2016; Brailsford et al., 2018; Crema & Verbano, 2019; Zhang et al., 2020). However, potential efficiency gains must be balanced against risks to quality and safety of care. This is particularly important in the current environment, where a crisis of public health has followed a prolonged period of financial austerity. Additionally, rising demand, increasing costs, and changing standards have placed huge pressure on health systems, risking serious breakdowns in care. For example, Vincent and Amalberti (2016) took a direct approach to safety, arguing that healthcare in the UK is frequently delivered at ‘the illegal normal’ level, where care is unreliable, and quality is poor. The patient will usually escape harm, with staff using adaptive strategies to cope (Kadri et al., 2014; Vincent & Amalberti, 2016; Amalberti & Vincent, 2019). This is particularly relevant in emergency care, where variability in patient arrivals, illness severity, and health resources needed for treatment can lead to system strain (Cildoz et al., 2019). The Kings Fund (Murray et al., 2018) reported a continuing, significant decline against the ED 4-hour standard in Type 1 departments, which are attached to acute hospitals. In these, the standard has not been met since 2014, while 4-12 hour waits rose by just under 20% between December 2017 and January 2018. In comparison, the performance of Type 3 departments, that is Minor Injury Units (MIU), Walk-in Centres (WIC), and Urgent Care Centres (UCC), has remained stable at around 99% since 2010.

Meanwhile, public satisfaction with the NHS is falling: between 2014 and 2017, the level of dissatisfaction almost doubled, rising from 15% to 29%, and in 2018 public satisfaction was at its lowest level for more than a decade (Robertson et al., 2019). Considered on a continuum, both quality and safety improvements are a priority where care is operating at this level. While improving reliability of service is one approach to safety, emergency care clinical activities are necessarily adaptive hour-to-hour, due to the variation in patient numbers and case mix presenting throughout the day and across the wider system. The aim for hospitals is to deliver care that is safe, even if not ideal, when working conditions are difficult such as at times of high workload or during times of major emergencies. From a quality improvement perspective, where quality and safety are on a continuum, there should be no trade-off to be made with efficiency and cost because prioritising quality and safety can reduce costs and enhance efficiency, either directly or indirectly (Graban, 2009; MacArthur et al., 2012), for example by lowering complications, or reducing hospital readmissions (Huerta et al., 2008). For The Kings Fund, Jabbal and Lewis (2018) found some common factors across hospitals that show how cost and quality can be successfully balanced to provide value. This shift from a ‘quality’ focus to a ‘value’ focus reflects the necessary emphasis on cost, where ‘value’, as defined by Porter

(2010), involves delivering the highest quality health outcomes for patients at the lowest possible cost (Jabbal & Lewis, 2018).

The next section focuses the research on short-term decision support, and its precursor, situation awareness, and proposes research questions and objectives to meet these.

1.2 Research focus

When care is operating close to the boundaries of capacity as has been the case in the UK for some years now (British Medical Association, 2018; Department of Health, 2018; House of Commons, 2019), the risk of a critical event occurring is high. Anandaciva (2019) outlined a range of contributing factors at the start of 2019/2020 winter, including high levels of 'flu, the pension crisis, lack of funding, and preparations for a no-deal Brexit. In general, the NHS sees improvements in waiting time performance over summer, allowing it to be prepared for winter in terms of staffing and waiting lists. However by end of 2019 financial year, the NHS reported 100,000 staff vacancies (NHS Improvement, 2019), significantly impacting on system resilience.

System resilience is defined as the ability to *anticipate*, to *react* and to *mobilise resources* for rebuilding and recovering after a degraded or critical state (Hollnagel, 2009, 2011a, Hollnagel et al., 2019). In emergency healthcare, this requires adaptive behaviour from staff to maintain system functioning, and the ability to make effective short-term decisions (Kadri et al., 2015). Situation awareness (SA) is an important constituent in decision-making processes, defined as an operator's understanding of 'what is going on' while interacting with a complex, dynamic system (Endsley, 1995; 2016). This can occur at the task-level (individual or teams) and at the system-level, and a range of validated metrics have been used for measuring SA at both levels. In ED, the variability in patient arrival rate and severity can interfere with SA as demand exceeds the capacity of available resources (Levin et al., 2012). This subsequently impacts on system resilience, as the ability to anticipate, react and recover from crowding or other critical situations reduces. Appropriate information can support SA and subsequent decision-making toward action, to support the delivery of safe, quality care.

One approach to dealing with the variability in patient throughput in EDs was proposed by Espinoza et al. (2014), who explored the use of real-time simulation to improve daily operations in an ED by portraying the current state of the system and predicting ED performance. With the rise in digitalisation and Industry 4.0 applications, real-time methods are increasingly used for online decision-support, in particular in manufacturing environments (Rodič, 2017). While data security challenges to this approach limit its application in healthcare, greater volume and availability of standardised operational datasets can support the investigation of real-time decision-support tools. However uptake in healthcare has been slow. Activity in ED needs considerable hour-by-hour adaptation because of the variation in patient presentations, the urgency of patient conditions, and the vulnerabilities of the healthcare system (Vincent & Amalberti, 2016). Real-time simulation holds promise as an approach for supporting decision-making in this situation. It can allow visibility over the current system state for supporting short-term decision making, where the simulation behaviour can be close to real behaviour. Once initialised with the current system state, the simulation is run over a short time period, enabling the results to be more accurate (Bahrani et al., 2013; Oakley, Onggo & Worthington 2020). Though there is a growing interest in real-time simulation, few published studies have used real-time data in healthcare simulation applications with the majority using historical data sourced from hospital databases, observation, and expert opinion (Almagooshi, 2015; Salleh et al., 2017). With increasing availability of healthcare operational data, opportunities exist to guide real-time or near-real-time decision-making (Weiner et al., 2016). Given the growth of digitalisation and the increasing availability of data made possible by the advancement of technologies related to the Internet of Things, sensors, and the use of personal devices, this is likely to be a rich area for research and applications in the future. As research in this area is seeing rapid advances in domains such as manufacturing and transport, understanding the applicability and challenges to such an approach in healthcare is essential. For these reasons, the first research question (RQ) to be addressed by this thesis is:

> ***RQ 1:*** *How can simulation approaches support short-term operational decision making in healthcare?*

Using simulation as a short-term decision-support tool, where quality and safety remain a priority, is potentially challenging, particularly in care networks where it can be hard to measure and interpret the impact that changes to one part of the system can have on another. In accord with Gul and Guneri (2015), Palmer et al. (2018) argued for the need for more innovative approaches to collecting and analysing healthcare data. Palmer et al. (2018) reviewed simulation modelling of patient flow across settings and found that multiple services, patient mix and different health-states within these services are rarely considered, in particular where services are time-dependent, or where capacity, demand and timing of patient use varies. Long and Meadow's (2018) review of simulation studies in mental health found a rise in multi-method modelling, which recognised that different segments of health systems have different profiles, which need different modelling methods. Similarly, Mielczarek and Uziałko-Mydikowska (2012), Gul and Guneri (2015), and Salmon et al. (2018) each identified in their reviews of healthcare simulation studies a trend towards multi-method platforms and multimethod modelling. Hybrid models, where two or more methods are combined, are not new (e.g. Shanthikumar & Sargent, 1983) but have seen increasing interest over the last decade or so, particularly in healthcare. This is alongside a more general, rapidly-growing academic interest in hybrid approaches to simulation studies as a comparatively new research area (Eldabi et al., 2016; Mustafee et al., 2017; Brailsford et al., 2019). Representing a problem area adequately to inform change suggests the need for a mixed-methods approach. These may better capture the complexity of healthcare problem-situations, and are therefore of interest in applied studies, as single methods may require significant assumptions and over-simplifications, or isolating single aspects of a real-world problem.

The rationale for using hybrid methods has been described as combining the methodological strengths of individual modelling techniques (Alvanchi, Lee & AbouRizk, 2011; Mustafee et al., 2015a), or to better capture the breadth of a problem situation (Lynch et al., 2014; Mielczarek & Zabawa, 2016; Abohamad et al., 2017). Mustafee et al. (2015a) and Powell and Mustafee (2016) made a distinction between hybrid simulation (HS), where two or more simulation methods are combined, and hybrid modelling (HM), where simulation is combined with other distinct methods at specific stages of a simulation study. HM may combine qualitative approaches such as Soft Systems Methodology (e.g. Kotiadis, Tako & Vasilakis, 2013), or quantitative approaches, for example predictive or descriptive analytics (e.g. Uriarte et al., 2017). These hybrid approaches extend simulation modelling methodology, arguably adding further value to the study in practice (Greasley & Edwards, 2019). This is of particular interest in real-world applications such as healthcare, where uncertainty and variability are key features (Mustafee et al., 2017).

Trkman et al. (2010) and Chae et al. (2014) argued that data analysis lies at the heart of organisational decision-making, and descriptive and predictive techniques are used increasingly in healthcare to optimise health, operational and financial outcomes (e.g. Hersh, 2014). Analytics techniques offer support to enhance the application of simulation models, while simulation can greatly enhance the value delivered by analytics applications by informing approaches to healthcare delivery (Marshall et al., 2016). This can help to link policy, management or logistical problems at strategic, tactical or operational levels of decision making. For example, a straightforward but costly solution to healthcare crowding resulting from a rise in, or variability in demand is to increase capacity and/or resources. Simulation methods can model alternative scenarios by optimising resource allocation and utilisation more efficiently against expected changes in demand. However to best achieve this, knowledge about the expected level of demand using forecasting methods can be inputted into the simulation model for planning (Mielczarek & Zabawa, 2016; Harper, Mustafee & Feeney, 2017). Given the demonstrated advantages of using mixed-methods approaches to better capture a problem situation, and the proposed benefits of using real-time simulation the second research question is:

RQ2: *How can an integrated hybrid approach using real time simulation and predictive analytics support short-term operational decision-making?*

Positioning healthcare research within the common principles of quality improvement (QI) is arguably important for any study that aims to intervene in frontline healthcare. In practice, this means involving stakeholders including frontline staff and patients, and prioritising safety and quality alongside cost and efficiency considerations. Involving patients in healthcare OR research is not common, but aims to start with an understanding of what is important to patients, to ensure that modelling efforts focus on measures that patients view as important as well as improving an in-depth understanding of the problem situation (Pearson et al., 2013). Many new products and processes come from a technology push of a new application, regardless whether there is a demand for the application (Brem & Voigt, 2009). By involving end-users, who are part of the system under investigation, an understanding of the current requirements and perceived value for patients, as end-users, can be considered in the design and function of a decision-support tool using a 'market pull' approach (Horbach et al., 2012).

It is generally accepted that involving stakeholders in simulation projects is fundament to success (e.g. Long & Meadows, 2018), however barriers in healthcare need to be considered. These include communication gaps between the stakeholders and researcher, poor management support and high clinician workload (Jahangirian et al., 2015). QI principles provide a starting point for healthcare service research. These principles are aligned with previous OR research which have investigated the challenges of conducting simulation research in the healthcare domain (Tako & Robinson, 2012; Brailsford et al., 2013; Jahangirian et al., 2015; Klein & Young, 2015; Long et al., 2019). The position paper by Pearson et al. (2013) aligns with healthcare QI definitions and recommends that interventions focussed on better patient outcomes and better system performance require the engagement of all stakeholders, including patients, the intended users of the system. This provides a suitable context for positioning healthcare OR research, subtly shifting the more traditional emphasis of OR to one where quality and safety are priorities alongside operational benefits, and where the constructivist model of stakeholder engagement includes patients as end-users, enhancing the ability to bring about health service change for the benefit of both the health service and the patients served by it. Given this, the final research question is:

> ***RQ3:*** *What are the implications and the added value to the system of using real-time data applications for both patient and NHS decision-support?*

1.3 Research aims and objectives

Research into real-time simulation in healthcare is still novel and untested, though advancements in model development, validation and applications are being seen

(Oakley et al., 2020). However its proposed benefits remain uncertain in practice. Firstly, it is necessary to understand what those proposed benefits are, by understanding the characteristics of healthcare, in particular emergency care, which may benefit from short-term decision-support, and by what mechanism. Situation awareness is dynamic by nature, it changes constantly as tasks, environment and sociotechnical interactions occur. Viewing a system through a sociotechnical lens, with a focus on the needs of the users of the system, supports effective performance, for example by paying attention to what type of information is needed to support system goals (Jones, 2015). Analytic methods are one approach to decision-support, and how these might contribute to the use of real-time simulation in healthcare needs is of interest, as such methods have been proposed to increase the value of modelling and simulation studies (Molloy et al., 2019; Greasley & Edwards, 2019). These groups of methods can be evaluated in the context of real-time data and real-time simulation, to determine their contribution to situation awareness and short-term decision support in healthcare.

To test this approach, a conceptual framework to support its development is required. It is advantageous if this framework is generic, as it can be tested in practice and transferable learnings can support similar future work. Within a use-case, its effectiveness against success criteria can be determined through evaluation, for example of staff who will be using the application. Evaluation is important, as the technical proficiency of the application is only one part of the overall perceived ‘usefulness’ of the approach in practice, and modelling and simulation (M&S) continues to experience challenges in healthcare in terms of evidence of real-world impact (e.g. Jahangirian, 2016).

From a QI perspective, the overall value of such methods at the system level needs to be understood. By involving end-users, who are part of the system under investigation, an understanding of the current requirements and perceived value for patients, as end-users, can be considered in the design and function of a decision-support tool using a ‘technology pull’ approach. The implications, and the added value to the system, of using real-time data applications for patient decision-support should be synthesised with the views of staff to determine the potential value of this approach at the system level.

The specific research objectives which will be addressed to achieve the three research questions in this thesis are stated below (Table 1.1). The research questions have been restated as aims to focus their relationship to the research objectives. In this thesis, the term 'simulation' refers to computer simulation, the use of a computer to represent the dynamic responses of one system by the behaviour of another system modelled after it.

Table 1-1 Research Questions, Aims and Objectives

Research Questions	**Aim**	**Objectives**
1. *How can simulation approaches support short-term operational decision-making in healthcare?*	To determine the need for short-term decision-support in healthcare, and to examine how simulation, real-time simulation, and hybrid modelling approaches using analytics have been used for short-term operational decision-support in the healthcare context, in particular emergency care.	**1.** To understand the need for short-term decision-support in healthcare, in particular emergency care. **2.** To explore how analytics methods can be used for short-term decision-support. **3.** To evaluate simulation approaches used in healthcare for decision-support and to identify how simulation is used for short-term decision-support. **4.** To determine the criteria for evaluation of a hybrid simulation approach for short-term decision-support in healthcare.
2. *How can an integrated hybrid approach using real-time simulation and predictive analytics support short-term operational decision-making?*	To test and evaluate the potential of an integrated hybrid approach for short-term decision-support in healthcare combining real-time simulation with other analytics approaches.	**1.** To propose a generic framework supporting an integrated hybrid approach for short-term decision-making in healthcare. **2.** To apply the framework within the case study in a hospital ED. **3.** To evaluate the application of the framework.

3. What are the implications and the added value to the system of using real-time data applications for both patient and NHS decision-support?	To analyse the system level impact of the use of real-time data for both patient and staff decision-support.	**1.** To critically evaluate the value that real-time applications provide at the system level. **2.** To synthesise previous findings and to evaluate the framework in light of the application.

1.4 Audience and scope

This research is expected to be of interest to OR researchers, both the simulation and analytics/data science communities, in particular those with an interest in applied healthcare research and real-time approaches to decision-support. While this is an expanding area, and significant research has attempted to understand the barriers to research in this area, there remains little understanding of what is needed and useful in practice to support short-term decision-making.

The healthcare domain is an example of a complex system with identified challenges in executing and implementing simulation studies, however other such complex systems exist for which real-time simulation and hybrid modelling approaches can be applied. This research, and its transferable knowledge, will be of interest, for example to researchers working in Industry 4.0, in particular in sociotechnical systems.

Human Factors researchers with an interest in SA and decision-support, and applied healthcare researchers with an interest in QI, or QI researchers/practitioners may find this work to be applicable. It may be of particular concern to those interested in how data applications may support SA and decision-making, and how a data analytics study may fit within a QI approach or support existing QI priorities.

Additionally, healthcare staff and policy-makers with an interest in research-based system and service improvement, and patients/the public with an interest in the challenges and potential for decision-support tools for supporting health service delivery toward better quality outcomes may find the following research applicable.

1.5 Outline of the thesis

Table 1.2 outlines the flow of the thesis, and the links between chapters and research questions. The literature review aims to address Research Question 1 and its objectives, and to determine the gaps in understanding to be addressed, and the criteria for evaluation, of an approach to short-term decision-support in healthcare.

Table 1-2 Outline of the thesis

Chapter	Purpose	Research Questions	
1	Introduction		
2	Literature Review	RQ1, Obj. 1-4	
3	Methodology		
4	Framework	RQ2, Obj. 1	
5	Case Study: Problem definition	RQ2, Obj. 2	RQ3, Obj. 1
6	Case Study: Hybrid model	RQ2, Obj. 2	
7	Case Study: Evaluation	RQ2, Obj. 3	RQ3, Obj. 1
8	Evaluation of framework	RQ3, Obj. 2	
9	Conclusion		

Research Question 2 is addressed by Chapters 3 (framework development), 5 and 6 (applying the framework), and 7 (evaluating the application). Chapter 5 (patient evaluation) provides an understanding of the current requirements and perceived value for patients, as end-users, to be considered in the design and function of a decision-support tool using a ‘market pull’ approach. It additionally supports the problem definition. Subsequently, Chapters 5 and 7 will inform the discussion and evaluation of this approach to short-term decision-support in healthcare, which addresses Research Question 3, in Chapters 7 and 8. Chapter 8 revisits the framework, and Chapter 9 concludes the research, and outlines its contributions.

Figure 1.1 is a graphical representation of the thesis, and the links between chapters to address research questions. As per Table 1.2, the research questions are indicated by their supporting chapters. The purpose of each chapter is clarified in terms of the overall structure of the thesis.

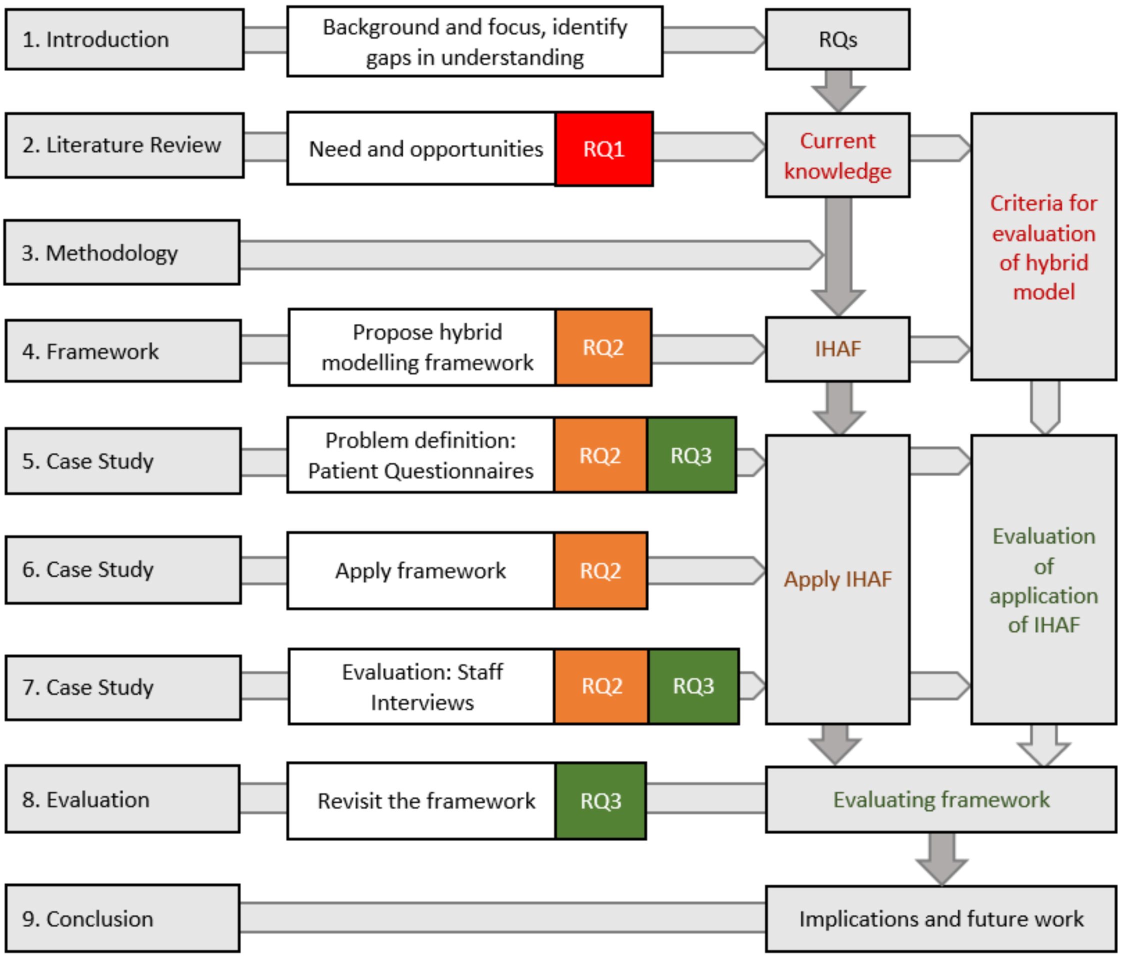


Figure 1-1 Graphical structure of the thesis including links between chapters

The chapters are summarised below:

Chapter 1: Introduction

This chapter has provided the background to the need for short-term decision-support in healthcare, in particular emergency care. It has outlined the challenges faced by emergency care, and how analytic decision-support, including real-time simulation, may contribute to supporting SA and short-term decision-making. The chapter outlines why a QI approach is important, with its focus on quality and safety as well as efficiency and productivity. Three research questions have been articulated, and the specific objectives needed to achieve the research questions are outlined. The structure of the thesis is presented.

Chapter 2: Literature Review

The literature review addresses the first research question. Research into real-time simulation in healthcare is still novel and largely untested in practice. Therefore its proposed benefits remain uncertain. Firstly, it is necessary to understand what those proposed benefits are, by understanding the characteristics of healthcare, in particular emergency care, which might benefit from short-term decision-support, and by what mechanism. Analytic methods are one approach to decision-support, and how these might contribute to the use of simulation in healthcare are explored, through hybrid approaches to simulation studies. The application and challenges to real-time simulation are outlined, as well as the criteria for evaluation of a short-term decision-support tool in healthcare. Finally, gaps in the literature are identified, and how this thesis will attempt to address these gaps are explicated.

Chapter 3: Methodology

This chapter outlines the research strategy, the research design, and the methods used to address the research questions. It discusses the research philosophy, and why it is important and relevant for OR real-world research, and outlines the association between the research strategy and each of the research questions and the methods. Research ethics are addressed.

Chapter 4: Proposed framework

This chapter develops and proposes a generic framework for the development and application of real-time hybrid modelling and simulation studies in sociotechnical systems such as healthcare. It addresses the first objective of the second research question: to propose a conceptual framework supporting an integrated hybrid approach for short-term decision-making in healthcare. The framework is developed to be generic, with transferable knowledge to support similar future work.

Chapter 5: Case study: Introduction and problem definition

This chapter introduces the use-case to test the framework, and the problem definition stage. It outlines the development, application, analysis, and results of a patient questionnaire. The questionnaire provides an understanding of the current requirements and perceived value for patients, as end-users, to be considered in the design and function of a decision-support tool. The results of the questionnaire will be subsequently integrated with the results of the evaluation and synthesised with the literature, to address Research Question 3 in Chapter 7.

Chapter 6: Case study: hybrid model

This chapter applies each of the components of the hybrid model: the descriptive, diagnostic, predictive and prescriptive components. The predictive component uses a SARIMA model to forecast the number of patients in the department, while the prescriptive component is a DES model of the emergency department for supporting recovery based on forecasted crowding.

Chapter 7: Case Study: Evaluation

Evaluation is the final component of the framework. The development, application, analysis, and results of the staff interviews are presented. Interviews are used to evaluate the hybrid model, how it might support SA and subsequent decision-making in practice, and how it might be improved. Research Question 3 is addressed by synthesising the patient questionnaires and staff interviews with the literature to evaluate the system-level value of the application.

Chapter 8: Evaluation of the framework

Subsequent to the application which tests the framework in practice, the framework itself is revisited and modified in light of its application.

Chapter 9: Conclusion

This chapter provides a summary of the contribution of the research, its limitations, and directions for future research.

1.6 Chapter Summary

This chapter has provided the background to the need for short-term decision-support in healthcare, in particular emergency care. It has outlined the challenges faced by emergency care in today's environment, and how analytic decision-support, including real-time simulation, may contribute to supporting SA and subsequent short-term decision-making in these environments. The necessity for a QI approach to frontline healthcare research is argued. Three research questions have been articulated and justified, and the specific objectives needed to achieve the research questions are outlined. The structure of the thesis has been presented graphically, and in summary. The next chapter, the literature review, addresses the first research question.

Chapter 2: Literature Review

2.1 Introduction

The aim of this chapter is to address the first research question and its three aims (Table 2.1), and to provide a basis for the subsequent empirical research. The research aims to explore the value that real-time simulation within a hybrid model (HM) can provide in a sociotechnical system for short-term decision-support. It is positioned in situation awareness (SA) theory, as adaptive behaviour from staff is required to maintain system functioning, which includes effective short-term decision-making.

SA is considered to be an important constituent of decision-making processes, and the provision of new information can support this understanding of ‘what is going on’. In ED, the hour-to-hour change in patient presentations can interfere with SA, and subsequently impact on system resilience, that is, the ability to anticipate, react and recover from crowding or other critical situations. Real-time simulation has been proposed as a solution to supporting systems that are highly stochastic in the short-term. It allows visibility over the current system state and can be closer to real system behaviour as the simulation is running over a short time period. As technologies collecting data in real-time advance, unlocking the value in real-time analytics, including simulation, is of burgeoning interest. This is of particular concern in healthcare, which in the UK, and worldwide, is under significant pressure. In the healthcare domain, simulation methods for decision-support have strived to gain a foothold in practice, yet is an approach which is considered to offer high value.

Data analytics (DA), defined as the use of data, statistical and quantitative analysis, and fact-based management to drive decisions and actions (Davenport & Harris, 2007), are used increasingly in healthcare to optimise health, operational and financial outcomes (e.g. Hersh, 2014; Chen et al., 2020). Analytics techniques can also enhance the application of simulation models, while simulation can greatly augment the value delivered by analytics applications by informing approaches to healthcare delivery. This can help to link policy, management or logistical problems at strategic, tactical or operational levels of decision-making. However an understanding of the challenges, exploring potential unexpected outcomes in practice, and how to optimise an intervention in terms of the value it might provide are important considerations. Positioning healthcare research within the common principles of quality improvement is arguably important for any study that aims to impact on frontline healthcare. In practice, this means involving stakeholders, including frontline staff and patients, and prioritising safety and quality alongside cost and efficiency. It also means considering the effects of a new application at the system-level, as non-linear outcomes can occur in complex sociotechnical systems. A conceptualisation of the relationships between constructs is illustrated in Figure 2.1. Data, analytics, and simulation all contribute to SA, and together aim to support quality improvement in terms of efficient and effective care delivery, patient experience and outcomes, staff capability, and ultimately, potential financial savings.

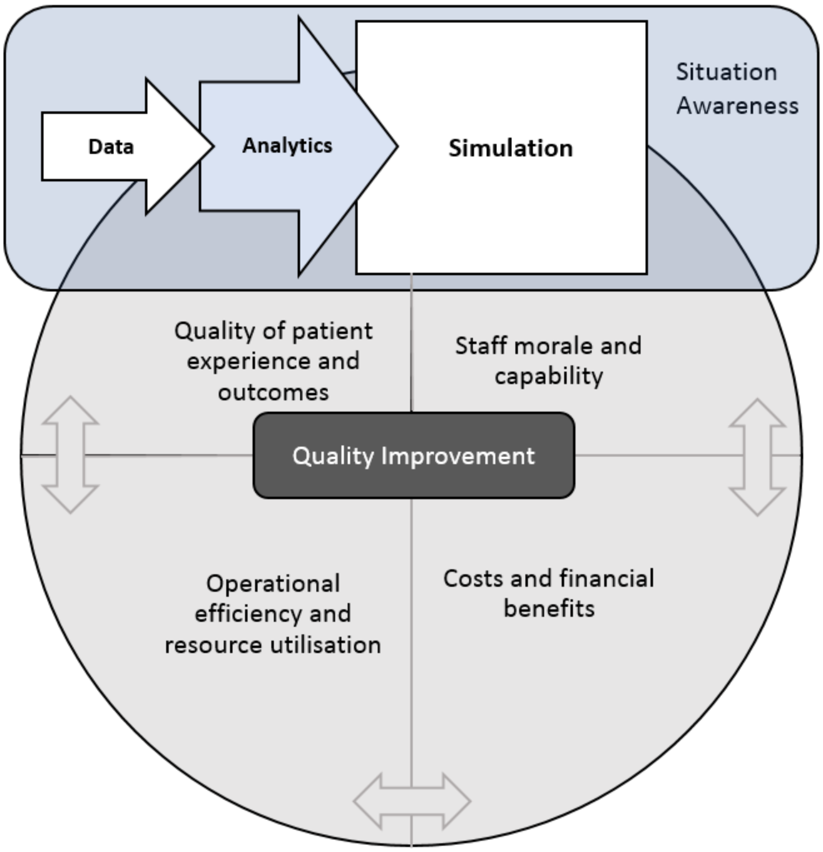


Figure 2-1 The relationship between real-time data, data analytics, SA and QI in healthcare decision support

Within the broader OR community, much effort is made to gain a coherent understanding of the future direction of the discipline, specifically how it may maintain its unique position finding solutions for real-world problems, while continuing to evolve as a discipline. These challenges have been the subject of research and debate both specific to healthcare (Royston, 2009, 2011; Monks, 2015) and more widely (Royston 2013; Holsapple et al., 2014; Ranyard, Fildes and Hu, 2015; Mortenson, Doherty & Robinson, 2015). Within healthcare, the question centres on real-world benefit, and in particular the lack of evidence supporting the use of simulation in healthcare (Pitt et al., 2016). Other important issues are the co-design of OR interventions with stakeholders, including frontline staff (Young et al., 2009; Van Lent et al., 2012; Kotiadis & Tako, 2016); raising the awareness of OR and simulation modelling with stakeholders (Royston, 2009; Monks, 2015), and methodological issues including coping with complexity, behavioural considerations and extending methodologies (Gunal, 2012; Royston, 2013; Klein & Young, 2015; Ranyard et al., 2015; Mortenson et al., 2015; Franco & Hämäläinen, 2016; Long et al., 2019).

Several authors have also raised the importance of engaging patients and the public in healthcare OR studies (Pearson et al., 2013; Monks, 2015; Batalden et al., 2016) while others have drawn attention to patient-centred care, patient safety, access and inequality in healthcare (Royston, 2009; Pitt et al., 2016). Pearson et al. (2013) argued that this enhances model credibility and relevance for patients and staff, ensuring that modelling efforts are focussed on issues that patients find important. The theory of co-produced services explains that consumers and providers of services work together to co-create value, and Balaban et al. (2016) extended this to healthcare to analyse the relationship between patients and service providers in designing, improving and evaluating healthcare services from a quality improvement (QI) perspective. Examples of patient involvement in OR studies are few and far between, but not unusual in the healthcare QI literature (e.g. Boivin et al., 2014 Coulter et al., 2014; Hardyman et al., 2015; Robert & Cornwall, 2015; Ocloo & Matthews, 2016). With a focus on implementation, improvement, and multi-stakeholder engagement, QI and OR can synergise in healthcare studies, and as the dominant framework for improvement in the NHS is a natural context for positioning OR studies. This expands the focus, rather than the scope of OR practice, as comparable tools are used in each discipline. Indeed simulation has been shown to be a useful tool in QI, adding value to the development of the organisation and helping staff face future challenges collaboratively (Rutberg et al., 2015; Hvitfeldt-Forsberge et al., 2017; Guise et al., 2017).

The following section outlines the role and structure of this review, and how it will meet the objectives of the first research question, and provide a basis for the subsequent empirical work.

2.2 Structure of the review

The integrative literature review is a form of research that reviews, critiques, and synthesizes representative literature on a topic in an integrated way such that new perspectives on the topic are generated (Torraco, 2005). This complements primary empirical research by narratively integrating the evidence to arrive at new insights (Esbach & van Knippenberg, 2020). An integrative literature review can be considered a form of research that can stand alone, as it results in novel outcomes, such as a new conceptual framework that defines the topic under review. Alvesson and Sandberg (2011; 2020) emphasise the need to problematize integrative reviews as a way to challenge and reimagine current ways of thinking and to consider broader knowledge domains. They emphasise broad but selective reading rather than systematic searches, and questioning rather than identifying gaps. Problematizing reviews aim for insight and rethinking, rather than rigour, or pseudo-rigour along dominant lines of logic.

This is the approach used to address the first research question. The value of the literature review is in synthesising aspects of two bodies of literature within a domain of application, with a view to establishing what the decision-support literature can learn from the decision-making literature to inform the development of OR tools for short-term decision-support which are useful in the real-world. Alongside the methodology chapter, it significantly informs the framework proposed in Chapter 4. This serves to open up new conversations around the purpose and use of short-term decision-support tools.

Table 2.1 re-states Research Question (RQ) 1, its aim, and its four objectives, which will be addressed in this chapter.

Table 2-1 Research Question 1

Research Question	Aim	Objectives
1. *How can simulation approaches support*	To understand the need for short-term decision-support in healthcare, and to	**1.** To understand the need for short-term decision-support in

short-term operational decision-making in healthcare?	examine how simulation, real-time simulation, and hybrid modelling approaches using analytics have been used for short-term operational decision-support in the healthcare context, in particular emergency care.	healthcare, in particular emergency care. **2.** To explore how analytics methods can be used for short-term decision-support. **3.** To evaluate simulation approaches used in healthcare for decision-support and to identify how simulation is used for short-term decision-support. **4.** To determine the criteria for evaluation of a hybrid simulation approach for short-term decision-support in healthcare.

This review aims to address the three objectives of RQ 1. The structure of the review is illustrated in Figure 2.2. Objective 1: *To understand the need for short-term decision-support in healthcare, in particular emergency care* is addressed through the pathway illustrated in Figure 2.2: *Healthcare as a sociotechnical system (Section 2.3) → Organisational decision-making (Section 2.4) → Situation Awareness (Section 2.4.2).* To explore the need for short-term decision-support in healthcare, in particular emergency healthcare, the role of SA in a sociotechnical system and how it contributes to decision-making will be examined. These topics will be positioned within healthcare, and more specifically within ED, which has particular challenges with regard to short-term decision-support.

Objective 2: T*o explore how analytics methods can be used for short-term decision-support* is addressed primarily through Section 2.5, *Data analytics and simulation*, and through the subsequent pathways illustrated in Figure 2.2. Data analytics as an approach to decision-support will be reviewed, and an evaluation of how these methods have been used alongside simulation in modelling and simulation (M&S) studies is undertaken, with the aim of understanding how these methods can enhance a modelling approach to addressing a research problem.

Objective 3: T*o evaluate simulation approaches used in healthcare for decision-support and to identify how simulation is used for short-term decision-support* is addressed through the pathway: *Simulation in healthcare (Section 2.6) → Real-time simulation (Section 2.6.4) → Hybrid Systems Modelling (Section 2.7)*. A review of simulation in healthcare, and the challenges and forward directions of modelling and simulation, including real-time simulation as used in healthcare and more widely in other domains, is undertaken. In this review, these pathways interact and converge to answer the first research question, and to understand the gaps in the literature to be addressed by this research. The review will conclude with a summary of the criteria for evaluation of a real-time hybrid modelling approach for short-term decision-support in healthcare.

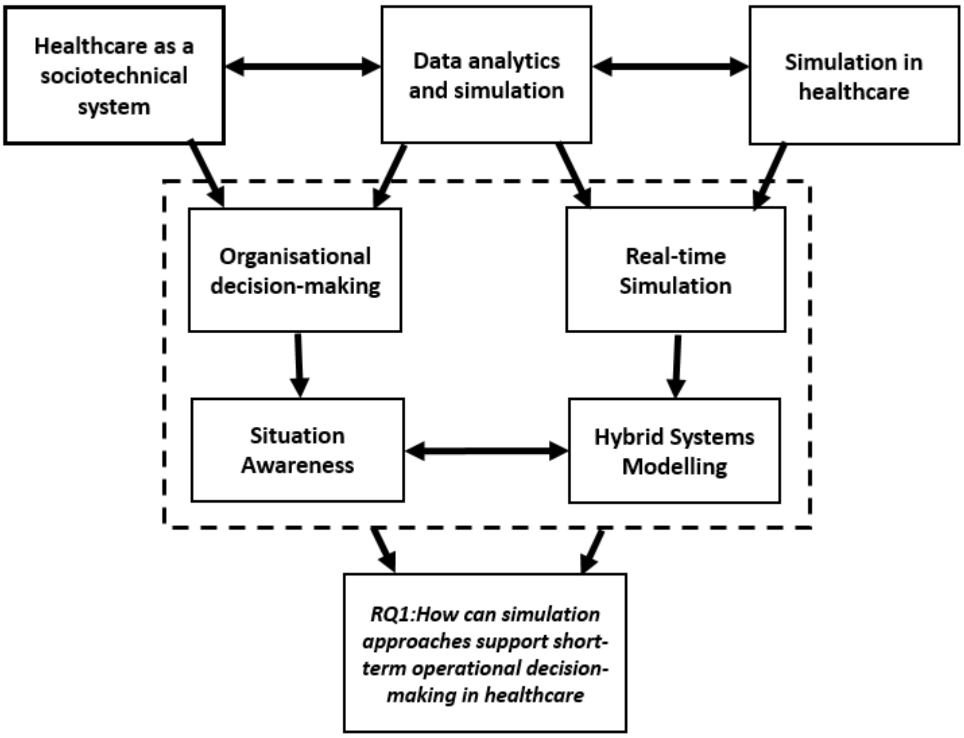


Figure 2-2 Structure of the literature review

The concepts in Figure 2.2 have been introduced in Chapter 1. The subsequent sections will go in to more detail to define the concepts for healthcare, and indicate their relevance to the problem situation of short-term decision-support in healthcare, in particular emergency healthcare, and a proposed solution, an integrated hybrid model using real-time data, predictive analytics, and simulation.

2.3 Healthcare as a sociotechnical system

2.3.1 *Quality Improvement in healthcare*

The recent NHS funding deal had hoped to ease ongoing pressures, however it was arguably not enough to restore performance against waiting times standards (The Kings Fund, 2019). Additionally, there are suggestions that Clinical Commissioning Groups regionally will take significant financial hits as a result of the COVID-19 virus, despite the large allocation of funds available to support the NHS during and following the crisis (HSJ, 2020). As NHS healthcare services have become increasingly constrained financially, processes are more tightly coupled and system resilience has decreased (Ham, 2017). Quality and safety can be conceived on a continuum, and it is increasingly important that quality is a driver of both strategic healthcare service improvement activities and day-to-day operational decision-making. Vincent & Amalberti (2016) describe almost all current safety initiatives as ‘optimising strategies’: that is, attempts to either improve the reliability of health processes, or initiatives to improve the wider system. They argue that in the real world, strategies to improve safety and quality should be aimed at managing risk in the often complex and adverse daily working conditions of healthcare. Where care is optimal in terms of adherence to standards, quality is achieved. However while standards of safety must be met, this definition of quality ignores a fundamental perceptual aspect to ‘quality’, which belongs to the recipient of the service: the patient (Coulter, 2015). The Health Foundation adopted the principles of quality from the Institute of Medicine for their definition: ‘making healthcare safe, effective, patient-centred, timely, efficient and equitable’ (The Health Foundation, 2013), shifting the emphasis toward perceived quality and value for patients. Batalden and Davidoff (2007) and Batalden et al. (2016) emphasised that quality improvements require the combined efforts of all stakeholders, including patients and their families, health professionals and researchers.

Even prior to the current crisis, there is evidence that care in the NHS often fell below the standard expected (Amalberti & Vincent, 2020), and patient involvement has variously been proposed as a vehicle for improving quality and safety, accountability, healthcare delivery, health equity and maintaining sustainability (Hardyman et al., 2015). Ham, Berwick and Dixon (2016) suggested that the practice of QI requires an understanding of systems thinking; quantitative data to understand variation; a participatory approach; and a determined focus on the needs of the patient**.** The Health Foundation (2013) also underline the importance of data and measurement for improvement, understanding processes, improving reliability, and understanding demand, capacity and flow. These properties overlap considerably with the goals and objectives of many M&S studies in healthcare. In the Berwick report for the National Advisory Group on the Safety of Patients in England (Berwick, 2013), it is emphasised that while the pursuit of operational targets is important, it should not displace the primary goal of better care. Arguably any attempt at an intervention in complex healthcare systems should operate according to these principles.

2.3.2 *A sociotechnical system approach in ED*

The term ‘sociotechnical systems’ was first used by Emery and Trist (1960) to describe systems that involve a complex interaction between people, technology and the environment in the work system. It has become popular to take a ‘systems view’ of healthcare processes when focusing on safety and quality, however there are a number of factors specific to ED that need to be considered when attempting to do so. Carayon (2016) outlined a range of human factors issues in ED which position it as a complex sociotechnical system. These include the event-driven nature of the work, hence work space and resources are difficult to predict hour-to-hour. Patient type is highly variable, increasing the difficulty in identifying workload or services needed. The work in ED requires many individual tasks which are variable in number; involve a number of specialities and skill levels; and ED requires transitions or shift handovers of care. Finally, ED systems and processes are closely coupled with other system components; work tools are unstandardized requiring a combination of handwritten notes, digital data, status boards and human memory; and many ED workspaces are not designed for the work taking place there.

Given this, it is important to use an approach to modelling sociotechnical systems that abstracts away from the specifics of particular methods and obtains systems-level understanding, for example starting with insight into what matters in terms of patient experience, staff satisfaction, efficiency gains, or cost savings. In complex sociotechnical systems, non-linear outcomes can give unexpected effects, for example data and information overload can act as a distraction and reduce comprehension of a situation (Murphy et al., 2019). However appropriate information is necessary to support task- and systems-level understanding. The implications are that designing decision-support tools means designing to support the ability to gain and maintain awareness of a situation in a dynamic environment. This can drive effective dynamic decision-making, while ensuring awareness of unexpected uses or effects to maximise the intended value as an integral part of ED activity, and to minimise the risks in practice.

2.3.3 *ED activity*

EDs provide care for acutely ill or injured patients arriving at random time intervals, with unpredictable levels of urgency and/or complexity. Patients wait for assessment and treatment according to triaged level of urgency in distinct queues. Triage categories are usually 1-4/5, with 1 being life-threatening and 4/5 being non-urgent. EDs typically provide a service 24 hours, 7 days a week. Patterns of demand are characteristically seasonal over the day, week and year. Patterns of patient urgency or complexity are less demonstrable, although more complex patients present in the winter months (Thornton, 2017). From 2004, the UK NHS constitution mandated that 98% of patients attending ED should be seen, treated and either transferred, admitted or discharged within 4 hours. In 2010 this target was relaxed to 95%, associated with declining performance. In recent years, performance against the 4-hour target has deteriorated in EDs across the country, with the national average as low as 69% in December 2019 (NHS England, 2020); the number of total ED attendances is also increasing, up 20% since 2009/10 (NHS Digital, 2019b).

A key focus of improvement for ED managers is reducing wait times and increasing throughput to reduce crowding and improve ED performance (Paul et al., 2010; Gul & Guneri, 2015). Crowding and long wait times are a primary cause of patient dissatisfaction in ED (Jurishica, 2005; Soremekun et al., 2011; Abo-Hamad & Arisha, 2013; Komashie et al., 2015). Crowding is also associated with poor clinical and operational outcomes, and perceived quality of service (Marmor et al., 2009).

Many solutions to the problem of crowding focus on increasing resources and capacity, however in the public sector this is often impossible. The focus therefore shifts to optimising existing capacity and resources in ED. Another consideration is that being a service system, both customers and providers can respond to changing work environments, and as performance targets curb the buffer to cope with changing demand, supply must adjust to meet demand. Under these circumstances, performance and behaviour of staff must adapt (Oliva & Sterman, 2001; Chahal et al., 2013). For example, Komashie et al. (2015) used Little's Law to investigate the relationship between queueing time and patient satisfaction, and service time and staff satisfaction. They hypothesised that as operational targets to limit waiting time will necessarily reduce service times where demand remains unchanged, the resultant impact on staff resources will ultimately result in reduced staff satisfaction. The proposed link to patient safety is that dissatisfied or stressed staff are more likely to cut corners through 'coping methods' to shorten service times or to meet operational targets (Bevan & Hood, 2006). Using Queuing Theory, Komashie et al. (2015) concluded that managing queues through targets is problematic, while synergising both staff and patient satisfaction is necessary for healthcare quality by keeping ideal service times close to actual service times for staff, and ideal wait times close to actual wait times for patients. However this requires transparency of actual waits, and also assumes that patient demand will remain stable.

2.3.4 *Adaptive capacity in ED*

Alongside demand variability, with growing demand for emergency care, EDs must anticipate crowding, and be reactive and adaptive with their delivery to have the required resilience to continue to deliver services safely during busy times. ED's are complex and can never be fully specified or controlled by rigid procedures and protocols, but require the ability to adapt to variability in the internal and external environment when needed (Dekker, Cilliers & Hofmeyr, 2011; Amalberti & Vincent, 2020).

Hollnagel (2009; 2011a,b) defined a system as *resilient* if it is able to adjust its functioning prior to, during, or following changes, disturbances, or opportunities, and thereby sustain required operations under both expected and unexpected conditions. Performance measures for resilience include rapidity of recovery, resource utilisation, performance stability and team situation awareness (Son et al., 2017). Back et al. (2017, p660) adapted the definition for healthcare as '*the intrinsic ability of a healthcare system…to adjust its functioning prior to, during or following events (changes, disturbances and opportunities), and thereby sustain required operations under both expected and unexpected conditions*.' System resilience sees safety as a process consisting of: knowledge about what to expect (*anticipation*); competence in knowing what to look for (*attention*); and knowing what to do and the resources to do it (*rational response*) (Hollnagel, 2009; 2011a). In ED, resilience has been measured using ED-RAG (Chuang et al., 2020) an adapted version of the Resilience Assessment Grid (RAG) questionnaire validated for use in other domains (Hollnagel, 2010; 2011b). This measures the potential for resilience, using features such as the ability to be both reactive and proactive, and the ability to learn. Kadri, Chaabane, Bekrar & Tahon (2015) applied the concept of system resilience in relation to emergency healthcare, describing it as the ability to *anticipate*, to *react* and to *mobilise resources* for rebuilding and recovering after a degraded or critical state. This can be due to epidemics or crises, but can also occur as a result of unpredictable fluctuations in demand. Expected fluctuations can be defined as seasonal variations monthly, daily and hourly which can reasonably be predicted. When care is operating close to the boundaries of capacity as has been the case in the UK for some years now, the risk of a critical event occurring is high. Anticipation is a key element of adaptive processes, by detecting as early as possible that a critical event is imminent. Proactive adjustment means that the system can change from a state of normal operation to a state of heightened readiness before the onset of an event. This involves monitoring changes in ED functioning and having sufficient time to implement corrective actions. In a state of readiness, resources are allocated to match the needs of the expected event (Hollnagel, 2011b).

Kadri et al. (2015) evaluated the effect of different corrective actions on strain indicators (for example, wait time until first assessment; number of patients in department when new patient arrives; ratio of number of patients to number of physicians), while a follow-up paper (Kadri, Chaabane & Tahon, 2016) described a set of corrective actions, which include transferring and rescheduling care activities; re-defining or re-allocating tasks to staff, adapting or adding medical personnel; and adapting or reassigning treatment areas. However these reactions must occur in real-time, while the ED is functioning. For staff, this involves determining the state of the ED, assessing the impact of corrective actions and launching corrective actions, while continuing to treat patients. Kadri et al. (2016) modelled a range of corrective scenarios, however the capability to improve anticipation of degraded situations would allow corrective actions to be initiated earlier, and recovery to be faster. In contrast, Ahalt et al. (2018) built an ED DES to evaluate three *crowding scores* (EDWIN, NEDOCS and READI), which are metrics used to quantify crowding to anticipate imminent crowding problems. The formulas variously capture aspects of patient severity; capacity in terms of number of doctors, ED trolleys and in-patient beds; wait times and arrivals; and service rate. While crowding scores can predict 'impending crowdedness', they cannot predict patient flow, making planning beyond the immediate short-term difficult. Managing patient demand at source is an alternative approach to coping with variable demand/capacity mismatch where capacity is fixed, operational targets are unyielding, and there are potential threats to patient safety as the adaptive response cannot compensate indefinitely or completely.

2.3.5 *Demand and demand management in ED*

The Kings Fund (2010) identified a number of ways that emergency demand can be reduced through demand management, such as pro-actively managing those at greatest risk of admission through risk stratification; redesigning primary and community care including integrating health and social care; and giving ambulance services more clinical responsibility. While these are aimed at reducing emergency admissions, for those with minor conditions, efforts have been made to empower self-care and pharmacy advice; social marketing campaigns have been used to encourage more appropriate use of EDs; there has been a mandated increase in GP numbers and access; and hospitals have been funded to stream people away from ED to more appropriate facilities. However, the evidence supporting most of the above approaches remains limited (Lee et al., 2013; Care Quality Commission, 2018).

Understanding factors which influence demand is important for managing demand. A significant amount of research has been undertaken to explore the characteristics of patients with low-urgency conditions who attend urgent or emergency care services. A report commissioned by the BMA and NIHR (Mason et al., 2017) investigated reasons for non-urgent ED attendance longitudinally from 1997, 2006 and 2016. They found an increasing unwillingness or inability of patients to manage their own risk, and an increasing perception that health problems are serious, with a desire for rapid reassurance. Fitzgerald et al. (2015) also found that the main reason for attending ED is perceived severity of illness, followed by lack of knowledge of alternatives, and that roughly a third of attendees are referred by another health professional. Mason et al. (2017) similarly found that an increasing number of patients are referred to ED from other healthcare providers, which they attributed to risk aversion. They also found an increased awareness of other services but confusion or reluctance to use these services, perhaps due to lack of knowledge about which clinical problems can be treated where.

A systematic review by Liscott (2016) found a high number of diverse and complex contributors to avoidable ED attendance in the UK. In common with the work of Mason et al. (2017) they found that perception of illness was a significant factor to avoidable ED attendance, related to lack of knowledge/education, and decision-making anxiety resulting in risk-averse decisions (Booker, Simmons & Purdy, 2014). Interpersonal factors were found to be a significant influence, with carers and relatives influencing a reduced tolerance for risk both for ED attendance and for self-management. In a report for the DoH, Rowe et al. (2015) found that parents of children aged under five years are risk averse, and will attend ED to be cautious. However their perceptions of what is available to them, and how appropriate these services are perceived to be is key to influencing their decision to attend ED. Banks (2010) reported that treatment-seeking behaviour is repetitive and reinforcing, such that while a large percentage of patients with minor ailments will self-manage, past experience will influence attendance behaviour. In Liscott's (2016) review, demographic factors showed mixed evidence, most likely interacting with other health and social variables, while socioeconomic evidence tends to point to social deprivation contributing to increased use of emergency services. Perceived lack of access to community services and GPs were a contributory factor, particularly out-of-hours. The Keogh Review for NHS England (Keogh, 2013) and Turner et al. (2013) suggested that supplier-induced demand may be a problem where access to multiple services is good. Several studies suggested that perceived ease-of-access and 24-hour service in ED may be a contributor to inappropriate use, particularly where access is perceived to be poor in alternative centres (Patton & Thokore, 2012; Smith & McNally, 2014). However Agarwal et al. (2011) found that confusion over alternatives may be a main contributor to inappropriate use of services. Recognising that the public find the range of alternatives to be confusing – Minor

Injury Units, Walk-in Centres, Urgent Care Centres – NHS England (2020a) announced the introduction of standardised Urgent Treatment Centres (UTC) by the end of 2020 as an attempt to simplify the options so that ED is not the default choice, aiming to reduce ED demand at source by providing a single alternative.

Thus, demand management can take two forms. Firstly, patients can be provided with knowledge that can support decisions about the most appropriate place to attend through education and social marketing, or based on the provision of new information such as current wait times (e.g. Mustafee et al. 2017b). However whether this information supports decision-making, and if so, in which patients, is unknown. Secondly, demand management can take the form of redirecting appropriate patients to alternative services as queues become unmanageable. Xu and Chan (2016) found through analytical investigation that proactive patient diversion using demand predictions could outperform diversion based on real-time information alone. This is an example of data-analytic decision-support, which extends traditional Business Intelligence or descriptive analytics to potentially offer further insights from the data using predictive analytics (Mustafee, Powell & Harper, 2018). The role of analytics in organisational decision-making will be explored in more detail in Section 2.5. The next section examines organisational decision-making, and the role of situation awareness (SA) - a constantly evolving understanding of the state of the environment - which drives decision-making and performance in complex systems. The theory of SA will inform the development of the implementation framework in Chapter 4.

2.4 Organisational Decision Making

2.4.1 *Knowledge as value*

Decision-making refers to making choices among alternative courses of action, and individuals throughout organisations use the information available to them to make a wide range of decisions at different levels. Strategic decisions set the course of an organisation over the long-term, tactical decisions determine how things will get done in the medium-term, while operational decisions are made in the short-term to keep operations running day-by-day and hour-by-hour. In complex systems, operational decision-making can be complicated by uncertainty, yet many decisions are critical to the success of an organisation's strategy.

The goal of simulation and other analytic applications is to provide decision-support for enhancing organisational performance through the analysis of data (Davenport et al., 2007; Evans et al., 2012). A commonly cited linear model of data analytics describes the transformation of data into information, information into knowledge, and knowledge into value (e.g. Davenport, Harris, DeLong & Jacobson, 2001; Acito & Khatri, 2014; Kuiler, 2014; Sato & Huang, 2015; Wang & Hajli, 2017). It is commonly accepted that information has no value until it is assessed and interpreted alongside existing knowledge, either tacit or explicit (e.g. Rasmussen & Ulrich, 2015). Explicit knowledge can be specified into formal rules and procedures, while tacit (or implicit) knowledge is associated with experience and expertise (Loebbecke, van Fenema & Powell, 2016).

Knowledge is variously defined as: the understanding gained from the analysis of information (Kuiler, 2014); the processing of information in the mind of an individual (Huber, 1991); and information combined with experience, context, interpretation, and reflection (Albert & Bradley, 1997). Alavi and Leidner's (1999) definition of knowledge is more specific: '*Knowledge is a justified personal belief that increases an individual's capacity to take effective action'*, however Connell et al. (2003) argued that such a personalised view of knowledge ignores the systemic context within which the knowledge is defined. A knowledge management approach, where knowledge is viewed as a systemic property of the organisational system to which it belongs rather than within the individual, supports a sociotechnical perspective. Here, the system can be viewed as a whole, with information held by people, artefacts, and their interactions (Stanton, Salmon & Walker, 2015). Boisot and Canals (2004) saw data, information and knowledge as possessing specific types of utility: data utility in that it can carry information about the physical world; information utility in that it can modify an expectation or state of knowledge; and knowledge utility in that it allows an agent to act in an adaptive way upon and within the physical world. Once enough awareness of the situation has been gained, a match between past experience and knowledge about the current situation can be sought to determine the appropriate course of action (Salas et al., 2010).

Correspondingly, Sharma et al (2014) argued that to capture the value that analytic activities can have on organisational performance, more attention needs to be paid to the organisational and behavioural aspects of decision-making. An example of this is how simulation or other forms of data analysis can work together with human sensemaking – the creation of mental models - to improve the generation of knowledge (e.g. Jolaoso et al., 2015). Dreyfus (1981) examined the intuitive thought processes of management decision-making, and established inherent limitations on simulation modelling. He found that the most critical factor in successful decision-making is the extent of the decision-makers' familiarity with, and situational understanding of, the problem situation. However fast, intuitive decision-making or slow, deliberative decision-making may take place in response to new insights, understanding or knowledge (Kahneman & Frederick, 2005). It is generally agreed that expertise increases the use of automatic decision-making; however some researchers have argued that objective accuracy of expert decisions are low overall, with better decisions unrelated to length of experience (Ericsson, 2007; Moxley et al., 2012). This suggests that even for experts, new information may support fast or automatic decisions, particularly where decisions need to be made in the short-term. In the study of Naturalistic Decision Making, defined as 'the way people use their experience to make decisions in field settings' (Zsambok & Klein, 2014) one influential position aligned with organisational decision-making is that of Beach (1997). He stated that values and beliefs, specific organisational and individual goals, and operational plans for reaching the goals, will guide and limit decision–making. This merges goal-orientated individual behaviour with the decisions and goals of other organisational stakeholders. Organisational decision-making is often challenged by shifting or competing goals and uncertain, dynamic environments. Other factors relevant to ED include ambiguity, a longitudinal context, incentives, repeated decisions and conflict (Gore et al., 2006). Naturalistic Decision Making is concerned with how people make decisions in complex, real world, uncertain contexts that can require real-time decisions in urgent situations with significant implications for errors. In decision-situations that have low immediate feedback, more information may be required to gain understanding. Analytics, including real-time data and visualisations, can have an important role in contributing to awareness of the current state of a situation by updating users' immediate knowledge and experience to make fast decisions that can inform adaptive action (Riveiro, Flakman & Ziemke, 2008). This is achieved by enhancing situation awareness, a knowledge state which is considered to be essential for decision-making and performance (Endsley, 2016). This will be explored further in the following subsections with a view to understanding how it might explain the need for short-term decision-support, contribute to system performance, and its role in evaluating the success of a simulation study for short-term decision support.

2.4.2 *Situation Awareness in sociotechnical systems*

Situation awareness (SA) is a concept in cognitive psychology and human factors which describes the degree to which a decision-maker is aware of events and elements in their environment, both spatially and temporally, and the effect of actions on goals and objectives now and in the future. Endsley (2016) and Endsley and Garland (2000) described it succinctly as 'knowing what is going on around you', and more expansively as the '*perception of the elements of the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future'* (Endsley, 1995). It is usually most relevant in a highly dynamic environment (Chiappe et al., 2015). SA provides the primary basis for subsequent decision-making and Endsley (2000) stressed that it is a state of knowledge, not the processes used to achieve that knowledge. This is in contrast to the concept of *sensemaking*, which has been defined as 'how people make sense out of their experience in the world', and is sometimes considered synonymous with the process of creating a mental model (Weick, 1995; Brock et al., 2008; Weick, Sutcliffe & Obstfeld, 2016). Both sensemaking and mental models are forms of *situational assessment*, which is a necessary step toward SA. Klein, Moon and Hoffman (2006a) differentiated sensemaking from SA in that SA is about the state of knowledge which is achieved through data, inferences or predictions, while sensemaking is the process of achieving this knowledge.

SA is an important constituent in decision-making processes (Endsley, 2000; Nguyen et al., 2018), and gaining understanding through interpretation is an essential activity for managers (Dreyfus, 1981). Individual SA is an operator's understanding of 'what is going on' while interacting with a complex, dynamic system. It occurs at three levels: the perception of elements in the environment, comprehension of their meaning, and the projection of their status into the near future (Endsley, 1995). These can be mapped onto the definition of system resilience discussed in Section 2.3.4, which sees safety as a process consisting of: knowledge about what to expect (*anticipation*); competence in knowing what to look for (*attention*); and knowing what to do and the resources to do it (*rational response*) (Hollnagel, 2009; 2011a). Theories and models of SA have been studied in a wide range of domains and organisational levels (e.g. Stanton et al., 2017), and are considered to be an important part of system resilience.

Endsley's (1995) influential theoretical model of SA based on its role in dynamic decision-making explored the relationship between SA and a variety of individual and environmental factors, including attention and available memory. Environmental limiting factors to SA include workload, stress, system complexity and environmental stressors, in particular, the effects of these factors on the ability to process information and make effective and timely decisions. Stress and anxiety reduce the capacity of available memory, such that individuals may be more likely to rely on external sources of information than internal memory storage. This impacts both on the decision itself, and the ability to adopt an effective decision-making strategy (Chiappe et al., 2012). Further, Endsley (1995) proposed that performance will be impeded where SA is incomplete or inaccurate. The competing demands of tasks for attention can exceed the operator's limited resources (e.g. Riveiro et al., 2008; DeWinter, Happee, Martens & Stanton, 2014). However, some researchers have suggested that given the increasing complexity of sociotechnical systems, the study of individual information processing is no longer relevant (Salmon et al., 2010; Chiappe, Strybel & Vu, 2012; Stanton et al., 2015; Stanton, 2016). Hence in in the last few decades, the focus of research has shifted from the unit of the individual, to that of whole systems (Stanton et al., 2006; Stanton, Salmon & Walker, 2015; Stanton et al., 2017).

Sociotechnical systems (STS) describe a combination of people and technical elements that interact in such a way as to support organisational activities and goals. The centre of STS are teams and team working, where multiple stakeholders with different goals are governed by organisational policies, rules, cultures, and regulatory policies. The technical elements are part of the STS and are considered important constraints and enablers of behaviour, while the system components interact with complex, non-linear and non-deterministic behaviours (Stanton et al., 2017). Distributed SA (DSA) views SA through a systems perspective, rather than a cognitive psychology lens, such that the whole system holds information, and humans within it can compensate and adapt for each other to maintain safe operation. This expanded definition of DSA is thus ‘the shared understanding of a situation among team members at one point in time’ (Salas et al., 1995; Stanton et al., 2017). Stanton et al. (2015) argued that either humans or technology can own this information, however the right information must be activated and passed to the right agent at the right time. Viewing a system through a STS lens, with a focus on the needs of the users of the system, supports effective performance, for example by paying attention to what type of information is needed to support system goals (Jones, 2015). This is relevant when designing decision-support tools, as the impact on the wider system needs to be understood.

2.4.2.1 *Understanding the relationship between SA, workload and performance*

Operators exert effort in a variety of ways. Physical effort is easy to conceptualise and measure, but mental effort is more conceptual, though both mediate between workload and performance (Hart & Staveland, 1988). Mental workload has been defined by Parasuraman et al. (2008, pp 145-146) as ‘t*he relationship between the function relating the mental resources demanded by a task and those resources available to be supplied by the human operator’*, and by Hart and Staveland (1988, p2) as ‘*a hypothetical construct that represents the costs incurred by a human operator to achieve a particular level of performance’*. Workload imposed upon an operator results from the task objectives, duration and structure, as well as the available resources, and may be modified by a range of individual and environmental factors. While there are many parallels between workload and performance, as task load increases workload will increase but performance can remain stable as a result of a range of adaptive strategies to maintain performance under increasing task load (Parasuraman & Hancock, 2001). However, at some point sustained high workload may prevent the operator from responding effectively to an unexpected increase in task load demand.

Performance is the degree of success in meeting task requirements as a result of the adoption of different strategies or different levels of effort (Hart & Staveland, 198). SA differs from performance in that it represents a continuous diagnosis of the system state, and is value-free, while performance results from a decision about which actions to take as a result of this diagnosis (Parasuramen et al., 2008). Many studies have demonstrated how the concepts of workload, SA and performance relate to each other at the individual level. In general, workload is seen to have a negative effect on SA which is positively correlated with performance (e.g. Nählinder & Berggren, 2002; Svensson & Wilson, 2002; Nählinder, Berggren & Svensson, 2004; Kiani et al., 2015; Naderpour, Lu & Zhang, 2016). Additionally, workload has a negative effect on both teamwork and SA (Berggren, Prytz, Johansson & Nählinder, 2011). In other words, task demands increase mental workload, which has a negative effect on SA, and in turn a negative effect on performance. For system design, these distinctions are important, as designs which support or improve task performance are different to those which support SA. As the cognitive nature of tasks asked of workers increases, and operational targets makes controlling task-load demand difficult, the understanding of SA in sociotechnical systems is becoming particularly important. An increasing interest in SA in healthcare reflects this, as an understanding that decision-making under cognitively difficult situations has a direct impact on patient and system safety.

2.4.2.2 *Situation awareness in healthcare*

SA has been identified as an important non-technical skill in healthcare clinical practice, and is the focus of a large body of research (e.g. Schulz et al., 2013; Wright & Endsley, 2017). Being able to perceive and comprehend a patient or system state, and make projections about the expected future development is crucial for safety. Healthcare processes operate using teams, and where the context is time-pressured and high-risk, distributed SA becomes more important. For example in operating theatres, where all members of an operating team need to share understanding of the current state and respond appropriately, this is critical (Fioratou et al., 2010; Schulz et al., 2013; Gillespie et al., 2013). A major portion of the job of a healthcare provider involves developing SA and keeping it up to date in a rapidly changing environment. Within a healthcare team, successful performance requires that team members maintain individual SA as well as shared SA. Specifically, shared SA requires team members to have an understanding of the type of information needed by others, knowledge of the devices used to distribute SA (e.g. visual dashboards), shared team processes to facilitate sharing of relevant information (e.g. communication, coordination, cooperation), and shared mechanisms such as a shared mental model.

SA can contribute to system resilience by supporting anticipation and attention, and informing adaptive action. In healthcare, escalation policies, which specify thresholds for increasing operational pressure, and responses to maintain patient flow, are the major codified organisational response for maintaining ED resilience. The effectiveness of ED escalation policies was investigated longitudinally by Back et al. (2017) using a mixed-method study. The aim of escalation policies are to increase capacity, reduce demand or increase efficiency through response predictability across staff, however the study found that in practice, escalation actions differed from those in policy. They found many examples of successful staff adaptive behaviour, such as pre-empting the need for escalation, using efficient practices to expedite patient flow, and flexing staff to areas in the department under pressure that were not specified in policy. Successful adaptive practice requires the ability to maintain awareness of the state of the wider system, and a delicate balance between continuing clinical work and interrupting workflow to perform planning activities (Back et al., 2017). It is critical to both staff morale/satisfaction (Kosnik, 2013; Johnston et al., 2016) and patient safety (Kosnik, 2013; Back et al., 2017).

Levin et al. (2012) stated that although direct links have not been established, there is growing evidence of a relationship between ED crowding and patient safety. This occurs when the system decompensates, that is, exhausts its capacity to adapt. Staff manage pressures by making in situ adaptations and goal trade-offs toward safe, quality outcomes, but this requires awareness of the situation to respond in an appropriate and timely way. Levin et al. (2012) investigated factors which interfered with SA in an emergency department and found that the number of patients managed (i.e. high taskload) contributed most to a reduction in SA and its potential effects on patient safety (i.e. performance). After the response mechanism is exhausted, the controlled parameter suddenly collapses or decompensates (Woods & Branlat, 2011).

Information technology (IT) has been implemented in health care environments, often as a replacement for paper-based or other manual tools for improving SA. One example of this is the replacement of large dry-erase boards, used in ED for tracking patient locations and clinical care, with real-time patient-tracking systems. The design of these systems has important implications for aspects of ED work, including changes to workload and SA of staff (Chahal et al., 2009; Pennathur et al., 2011; McGeorge et al., 2015). Such tools need to be implemented with care, as unanticipated effects can result from the type or presentation of information, including technology-induced errors (Peute et al., 2013; McGeorge et al., 2015). Similarly, IT systems that provide ambiguous information can actually reduce human decision quality and speed (Endsley, 2016). As a result, research on health IT design and evaluation has provided insights into factors contributing to successful system design, safety-critical aspects, system user-friendliness and usability issues. The majority of these systems compose dashboards displaying data and visualisations, which may be in real-time or near real-time. Real-time applications in healthcare have been used in health monitoring systems using physiological sensors on patients (Simpao et al., 2014; Wan et al., 2013); clinical decision-support (Mane et al., 2012), and dashboards for operational performance monitoring and compliance (Simpao et al., 2014; Weiner et al., 2016). It is increasingly necessary for healthcare organizations to improve performance by creating a data-driven decision-making culture, and to facilitate transparency and accountability. However healthcare IT applications focus on supporting evidence-based decision-making by healthcare service providers and managers (Spruit, Vroon & Batenburg, 2014; Simpao et al., 2014; Kao et al., 2016), while patient decisions, as end-users and an integral part of the system, tend not to be considered.

2.5 Data Analytics

Healthcare presents challenges when designing relevant decision-support processes. While Data Analytics (DA) is playing an important role in improving the delivery of healthcare services (de la Torre Diez et al., 2016), it is arguably not yet being fully exploited for enhanced effectiveness and efficiency of delivery (Wang & Hajli, 2017, 2019; Mehta et al., 2019). The most rapidly growing application of DA in healthcare is clinical decision-support, for example disease progression models, adverse drug events and risk prediction analytics (Raghupathi & Raghupathi, 2013; Tomar & Agarwal 2013; Jothi et al., 2015; Sukanya et al., 2017; Galetsi & Katsaliaki, 2019a, b; Mehta et al., 2019). For example, Mehta et al. (2019) analysed 2421 articles from 2013-2019, where 61% of papers focused on clinical applications, and 17% had an organisational focus. However the opportunities arising from data analytics for organisational benefits are considered to be far-reaching (Günther et al., 2017).

2.5.1 *Defining Analytics for healthcare*

Definitions of analytics vary, although while there are different perceptions about the nature and scope of analytics, there is a general agreement that it involves data-driven decision-making (Holsapple et al., 2014; Delen & Zolbanin, 2018). A commonly used definition is that given by Davenport and Harris (2007, p. 7):

'..the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions'.

While this definition seems broad, it was explicitly positioned within the business domain, with the goal of enhancing profit, market share and customer loyalty (Davenport & Harris, 2007). However DA can be used in any domain, for varying outcomes, while business analytics, indicating the domain of application, has been applied to organisational or operational outcomes such as improving patient flow and capacity-planning. These outcomes are often more relevant for healthcare than financial performance or customer retention, hence the above definition of DA, with its broad characterisation within the business domain, is adopted for this thesis.

The rationale for using DA in business to support decision-making overlaps considerably with the rationale for using OR in business domains (Evans, 2012; Hersh, 2014; Ranyard et al., 2015). For example, INFORMS define OR as '*the application of scientific and mathematical methods to the study and analysis of problems involving complex systems*' and BA as '*the scientific process of transforming data into insight for making better decisions*' (Robinson et al, 2010). However despite maintaining its position as a separate discipline, DA and its methods are no longer viewed as distinct for the OR community. Likewise, data analysts are adopting methods and techniques more traditionally viewed as OR tools such as simulation, particularly in the analysis of complex systems (Haas et al., 2011, Marshall et al., 2016). Indeed, Galetsi and Katsaliaki (2019a) defined analytics for healthcare as combining Information Systems, OR and statistics.

Similarly, Holsapple et al. (2014) took a more general perspective, offering a rationale for the application of DA as gaining value from "*supporting knowledge acquisition, insight generation, problem finding, and problem solving to assist decision-making*", using a range of techniques in situations that may be either well-structured or complex and messy. This complexity increases where there is rapid change, either internally or externally, and flexible solutions are required (Delen & Demirkan, 2013). The current interest in DA and its wide range of applications reflects the complex situations that organisations such as healthcare find themselves in, however it is also driven by advances in technology which allows data to be generated and collected quickly and efficiently.

The term *Big Data* describes large, heterogenous digital data that has been defined according to the 3V model - *Volume, Velocity and Variety* (Sagiroglu & Sinac, 2013; Katal, Wazid & Goudar, 2013; Larson & Chang, 2016; Marshall et al., 2016) with more recent additional descriptors *Veracity* and *Value* (Shao, Shin & Jain, 2014; Najafabadi & Villanustre, 2015; Sanjay & Alamma, 2016). Healthcare data exhibits all of these characteristics. *Volume* refers to the size of the datasets, which can enlarge rapidly; *Velocity* refers to the generation of data in real or near real-time, such that data collection and analysis must be performed at a much faster rate to maximise value from it; *Variety* deals with the various types of data, from structured to unstructured; *Value* refers to the trade-off between the costs associated with generating, collecting and analysing data, and the potential value of that data as a commodity or for providing competitive advantage; *Veracity* refers to data accuracy, and automated methods of checking for this (Sanjay & Alamma, 2016). Healthcare analytics publications have proliferated in the last five years (Galetsi & Katsaliaki, 2019a), as the value in the portability and interconnectivity of data is increasingly realised (Günther et al., 2017). However as the volume of healthcare data continues to multiply globally, the benefits and value created by DA in healthcare remains relatively unexplored, opening up the opportunity to focus on tools and applications used for the analysis of healthcare data, along with the barriers to their extensive use.

One difficulty is determining how to measure the outcome to define what value has been achieved. Mustafee et al. (2017b) suggested that this can sometimes be based on judgement rather than any statistical or scientific measurements, which may result in additional challenges that need to be managed. Additionally, the culture of health can challenge change based on data; senior management support and involvement are critical to success (Koster, Stewart & Kolker, 2016; Kao et al., 2016; Chen et al., 2020). How organisations facilitate the interests of different stakeholders, secure stakeholder interests at the work-practice level, and understand in which contexts different stakeholders succeed or fail to gain value from data are topics which have remained unexplored in research (Günther et al., 2017). Further, hybrid predictive and prescriptive approaches potentially bring their own challenges, in particular around validation of the models, real-world implementation, and evaluating the effects of implementation in a clinical setting (Janke et al., 2016).

Nonetheless, analytics can reduce subjectivity in decision-making (Sharma et al., 2014), and benefits have been demonstrated in healthcare (de la Torre Diez et al., 2016; Salleh et al., 2017). The most widely used functional categorisation for DA loosely groups techniques as *descriptive* (what happened?); *diagnostic* (why has it happened?); *predictive* (what is likely to happen?); and *prescriptive* (what should be done about it?) (e.g. Raghupathi & Raghupathi, 2013; Khalifa & Zabani, 2016). These categorisations are generally considered to be hierarchical (e.g. Kiron et al., 2012), such that descriptive techniques present historical information in real-time, standard or ad hoc reports, queries and alerts to answer questions around what happened, where it happened and what actions may be needed. Diagnostic, predictive and prescriptive techniques support questions about why an event occurred, what will happen next and what is the best that can happen, potentially providing a higher level of insight and intelligence.

2.5.2 *Positioning Analytics for OR*

While the functional categories of quantitative techniques described above are widely used, it is increasingly considered advantageous to include qualitative methods in the analytics toolbox, with a view to maximising the value that can be gained from an analytics approach. For example, from the perspective of Supply Chain Management (SCM), Waller and Fawcett (2013) defined analytics as '*the application of quantitative and qualitative methods from a variety of disciplines in combination with SCM theory to solve relevant SCM problems and predict outcomes, taking into account data quality and availability issues'*. Similarly, in Human Resources, Rasmussen and Ulrich (2015) contended that analytics is not about data *per se*, but about data for informed decision-making. This means defining the problem, having the right data – which could be qualitative or quantitative - asking the right questions, and interpreting the results and implications the right way. They further emphasised that to gain impact from data requires a focus on intervention and change. This is easier if the analytics approach '*includes qualitative data, intuition, experience and — most of all — if it* *works on co-creating a coherent story with key stakeholders'* (Rasumussen & Ulrich, 2015, p 239). Similarly, Holsapple et al. (2014) argued that analytics is far more than quantitative methods, as the real world is full of messy problems which have qualitative aspects.

From an OR standpoint, Mortenson et al. (2015) conceptualised DA through a historical perspective, suggesting a research agenda incorporating DA methods such as unstructured approaches, real-time analytics and other methods, that is, engaging with the surrounding ecosystem such that the most relevant methods are used to solve any specific problem. Many OR methods have been developed to deal with these problems, for example as summarised by Mingers (2011). These are particularly useful where decision problems are unstructured, and often more than one method is required to fully capture a problem situation.

Data quality and analytics capability are increasingly understood to provide organisational advantage, including in healthcare (Chae et al., 2014), and data-driven decision-making is rapidly gaining traction in the healthcare domain (Batarseh & Latif, 2016; de la Torre Diez et al, 2016; Galetsi & Katsaliaki 2019a,b,c). It is therefore of interest to understand how DA approaches might enhance simulation studies for operational decision-support.

2.6 Modelling and Simulation in healthcare

2.6.1 *Simulation Methods*

Modelling and simulation (M&S) describes the use of a simplified, typically dynamic representation of a real or proposed system, often accompanied by an animation to facilitate visualisation, communication and decision-making. The purpose of M&S is to provide insight and understanding into the physical processes of a system. Potential changes to the system can first be simulated to predict their impact on system performance (Fishwick, 1995), and the knowledge gained may be of great value toward suggesting improvements in the system under investigation (Banks et al., 2001). These systems analysis tools are usually developed in response to a specifically identified problem, and running the simulation provides information about the interactions of the system over time. This allows for exploration of the consequences of different decision scenarios through experimentation with the model using 'what-if scenarios', rather than experimenting with the real system, which may be costly or unsafe (Chahal, Eldabi & Young, 2013). Simulation modelling is appropriate for finding solutions in complex systems such as healthcare, where there can be a large number of parameters, behaviour is non-linear such that cause-effect relationships can be difficult to establish, and variability is important. It can capture complex system behaviours, and is considered a powerful and cost-effective set of methods for quality and performance improvement in healthcare, to understand how complex systems operate and meet operational targets, and how they can be improved (Katsaliaki & Mustafee, 2011; Gul & Guneri, 2015; Brailsford, Carter & Jacobson, 2017). This makes it a suitable set of methods for planning, research, education and decision-support.

The most commonly used modelling and simulation (M&S) methods in OR are Discrete-event simulation (DES), System Dynamics (SD), Agent-based simulation (ABS) and Monte-Carlo simulation (MCS) (Katsaliaki & Mustafee, 2011). DES started in manufacturing, and has evolved toward much broader applications. It models queuing systems over time, by representing entities, which flow through a network of queues and servers, with resources shared by the servers. DES is particularly suitable for modelling healthcare systems, as it can account for stochasticity, while entities in the simulation can specify patients with varying characteristics. It is useful in highly stochastic settings where the visual interface may be important (Seila & Brailsford, 2009). In particular, any system that involves the flow of objects naturally lends itself to DES modelling. However other modelling approaches have been successfully applied in healthcare, which also features 'dynamic complexity', where SD is an ideal approach for this (Brailsford et al., 2017). The underlying principle of DES is that changes to the system over time are due to the variability within and adjacent to the system being modelled, while the principles of SD are that the structure of the system is responsible for its changes over time (Morecroft & Robinson, 2006). SD is an analytical modelling method which combines qualitative and quantitative elements. In SD, the problem is defined dynamically, and the structure of the system is conceptualised as non-linear behaviour using stocks and flows, internal feedback loops and time delays. It is particularly useful for modelling strategic-level problems, such as policy implications, and has been used for modelling emergency care systems (Lane, Monefield & Rosenhead, 2000; Lattimer et al., 2004; Kang et al., 2014) and healthcare more widely (Brailsford, 2008; Lyons &

Duggan, 2015; Rashwan, Abo-Hamad & Arisha, 2015). While SD explicitly accounts for qualitative features of a system, ABS takes the agent perspective when viewing any system. Modelling the agent, its behaviours and interactions with other agents and the environment can produce a more accurate representation of the world (Macal, 2016). However challenges include calibration and validation, particularly due to the complexity of ABS models. ABS and DES have some features in common, in particular the ability to model complex, non-linear states, to deal with stochasticity and to model individual patients, with the key differences being that ABS can model interactions between entities, while DES can model queues (Gul & Guneri, 2015). Where crowding is the problem under investigation, DES is an appropriate choice of methodology.

Simulation modelling, in particular DES has been used extensively in healthcare for decades (e.g. Jun et al, 1999; Fone et al., 2003) for a wide array of problems. These include designing policies and strategies, implementation and delivery of services and targets, monitoring and evaluation, resource allocation, cost-benefit analysis, patient flow and risk assessments (Brailsford, Harper, Patel & Pitt, 2009; Royston, 2009; Almagooshi, 2015). The body of research in simulation modelling in the healthcare domain is vast, and expanding rapidly. This is evidenced by the number of reviews of the use of simulation in healthcare, including those by Eldabi, Paul and Young (2007), Brailsford et al., (2009), Gunal and Pidd (2010), Mustafee et al. (2010), Katsaliaki and Mustafee (2011), Aboueljinane et al. (2013), Bhattacharjee and Ray (2014), Almagooshi (2015), Mohiuddin et al. (2017), Salleh et al. (2017), Palmer et al. (2018) and Salmon et al. (2018), amongst others. Despite a rapid escalation in the successful application of these methods in healthcare, challenges continue to exist, which are explored in the following sections.

2.6.2 *Challenges for healthcare simulation modelling*

Simulation, in particular DES, has been used extensively for decision-support in healthcare Operations Research (OR) for decades, establishing both successful and impressive progress, and a degree of frustration. An early review by England and Roberts (1978) examined several hundred healthcare simulation studies across a wide array of applications. They found that relatively few of these reported significant effects on the healthcare system being studied. Multiple reasons were posited for this, including models that failed to capture human system behaviour adequately, and lack of incentive to implement change compared with engineering and production applications, where lowering cost and maintaining or improving quality is considered a fundamental goal (England & Roberts, 1978). In the forty years since this publication, increasing requirements in health to lower costs and improve safety, alongside computing and technological advancements and a rapid evolution of methods have seen a significant and steady rise in simulation studies, which has considerably changed the landscape for healthcare simulation applications. However, whilst the results of simulation studies are frequently encouraging, reviews continue to report low evidence of real-world impact (Brailsford et al., 2009, 2018; Katsaliaki & Mustafee, 2011; Jahangirian, 2016).

Multiple reviews of healthcare simulation have been published in the last decade, including the umbrella review by Salleh et al. (2017) who found that of the 37 healthcare simulation reviews they examined, 30 were published since 2010, reflecting the increasing use of the method in healthcare. Yet Royston (2009) and more recently Monks (2015) pointed out that OR as a whole does not seem very visible to healthcare managers or clinicians, with applications scattered despite the potential effectiveness of the methods for improving organisational performance. The evidence for successful implementation in healthcare may be more readily available in the grey literature (Brailsford et al., 2009; Van Lent, VanBerkel & Van Harten, 2012), however it has also been elaborated that lack of implementation of the results of a simulation model do not necessarily equate to failure of the study (Ormerod, 2001; Crowe et al., 2017).

Several authors have addressed the question of what constitutes *success* or *failure* in a simulation study (Gogi et al., 2016; Jahangirian et al., 2017), with temporal, perceptual and contextual factors all contributing to a notion of success. The healthcare domain is a particular focus of research for challenges, failures and successes in simulation studies. Harper and Pitt (2004), Brailsford (2005), Jahangirian et al. (2012), Tako and Robinson (2012), Brailsford et al. (2013) and Klein and Young (2015) have all investigated the research and implementation challenges of healthcare simulation studies. Process is usually considered more important than content (Robinson, 2002), as OR is a collaborative discipline, and modellers engage with stakeholders in the system to define and develop solutions to problems (Monks, 2015). For example, five critical success factors proposed by Jahangirian et al. (2017) for evaluating the success of a simulation study included communication with stakeholders, competence of the modeller, and responsiveness to the needs of stakeholders. Jahangirian et al. further developed these factors into a measurable set of KPIs that captures the relationship between the modeller, the problem-situation and the stakeholders as a measure of the success of the study. It has been argued that simulation studies help stakeholders to gain insights into problem-situations, and subsequently to develop effective solutions that are not necessarily generated by implementation of the model results (e.g. Connell, 2001; Ormerod, 2001; Monks & Meskarian, 2017; Kotiadis & Tako, 2018). Gogi et al. (2016) explored the role of DES models in generating insights with supporting empirical evidence. Likewise, Connell (2001) presented a four-quadrant evaluation of the *contribution of the approach* and the *contribution of the outcome*, together with '*gaining insight'* and '*managing change'*. These lend support to the proposition that implementation is not necessarily the only measure of success, but that the process of conducting the study in collaboration with stakeholders can result in successful outcomes. However not all researchers agree. Royston (2013) strongly argued the case for OR as a whole to focus on implementation of results of OR studies through 'design thinking'. While translating research into real-world results is a concern in all sciences, Royston asserts that for OR this is a particularly serious issue given that improvement is the goal of the discipline, hence all studies should be thinking about the realities of implementation. This requires a focus on synthesis, rather than analysis and a solution-focus rather than a problem-focus. Monks (2015) was in broad agreement, linking OR with implementation sciences, that is, the study of methods to increase the uptake of research findings in healthcare. Ultimately, it remains difficult to determine the value of OR interventions, in particular simulation, with few published studies evaluating the approach, the implementation, or impact. Salmon et al. (2018), in their review of ED M&S studies, highlighted a number of papers with a variety of methods that failed to follow through to evaluate the benefit of the study, concluding that an implementation or evaluation plan should form part of the overall study.

Outside of the issue of real-world benefit, conducting simulation studies in healthcare presents challenges. A qualitative study by Brailsford et al. (2013) found that barriers to the uptake of simulation modelling in healthcare included time and capacity issues, lack of senior management and expert support, and data and IT issues. For modellers, challenges include gaining buy-in and credibility, conflicting political objectives and data issues (Harper & Pitt, 2004). In order to build credibility, the relationship between the modeller, key staff and the problem situation is considered fundamental. Tako and Robinson (2015) surveyed expert modellers and found that simulation modelling in healthcare differs to other sectors due to its complexity, the messiness of its problems, difficulties with access, political influence, lack of client time and resistance to change. They suggested a range of mitigating strategies including developing innovative ways of managing projects to account for these barriers, and using qualitative methodologies. Eldabi, Paul and Young (2007) also recommended qualitative methods to facilitate participation, but suggested that there may be a communication gap, in that few modellers really understand healthcare and few clinicians or healthcare managers understand simulation. Nonetheless, while wide-scale adoption of the approach remains limited, there is a greater demand than ever for evidence-based operational change in healthcare (Brailsford et al, 2017).

2.6.3 *The future of healthcare simulation modelling*

The scale and scope of ED M&S is enormous, and a recent review of ED simulation by Salmon et al. (2018) found that the number of ED articles is increasing by approximately 25 papers per year. Due to the stochasticity and queue-based structure of healthcare, DES is the most commonly used simulation method in ED, and crowding is the most common problem investigated (Gul & Guneri, 2015; Mohiuddin et al., 2017). From early reviews (Jun et al., 1999), it was found that many simulation models investigated individual units within multi-facility hospitals but failed to capture the interaction of major services and the analysis of the system as a whole. Salmon et al. (2018) found that simulation studies that considered ED as part of a larger system used a greater variety of modelling approaches, including hybrid simulation. They contended that more effort is required to examine external influences such as downstream constraints, including available beds or domiciliary care. While this brings additional challenges associated with wider stakeholder engagement, clarity and communication of purpose, and project management, it will increase the relevance of the study for stakeholders, hence confidence in the solution.

Representing and predicting human behaviour is seen as a significant unanswered modelling challenge (Taylor et al., 2015), and a renewed interest in behavioural OR evidences this (e.g. Franco & Hämäläinen, 2015; 2016). Modelling human behaviour provides an ability to avoid abstracting away peoples' differences and increase the accuracy of prediction, particularly where behaviour affects output variables to an extent that they require consideration (Greasley & Owen, 2018; Wang et al., 2020). Also with a behavioural focus, the human and social influences on the process of modelling are of renewed interest, in particular the role and impact of these aspects related to the use of M&S to support problem-solving and decision-making (Wezel & Winterfeldt, 2016). As this can be critical in determining outcomes of an M&S study, Brocklesby (2016), amongst others, have argued for reporting accounts of the modelling process that provide insights into conditions of success or failure of the study. In social domains such as healthcare, with well-documented barriers to the successful implementation of M&S studies, this is a particularly interesting area for research given the fundamental concern of OR with human problem-solving and decision-making in practice. It is widely accepted that effective stakeholder engagement influences the outcomes of a modelling study in a social organisation (Long & Meadows, 2018) however the healthcare domain in the UK continues to be both under pressure, and reactive in its decision-making, increasing the difficulty of participatory approaches.

In a resource-limited environment, planning and effective decision-making continues be of critical importance. Data-driven approaches offer potential solutions. For example, process mining uncovers process knowledge from the analysis of event logs, reducing the time taken in the conceptual modelling stage of a DES study, and errors of bias (Abohamad et al. 2017). Similarly, Abdelbari & Shafi (2017) demonstrated how causal loop diagrams for SD conceptual models can be effectively learned using a neural network methodology directly from system observations. Elbattah and Molloy (2016) used machine learning to reduce the uncertainty underlying DES, and realise a higher level of model realism. Given historical patient records, machine learning models made validated predictions on important outcomes related to patient discharge for every patient generated by the simulation model, specifically length of stay and destination. With increasing awareness of the importance of data, advances in computer hardware and software, as well as increasing research focus on the use of data to support improvement, efforts lend themselves to more innovative uses of simulation methods applied to real-world problems that support system insights from data. These approaches are discussed in more detail in Section 2.7.5. Greasley and Edwards (2019) recommended multi-disciplinary collaborations between OR and data scientists to share expertise, as each field brings complementary skills.

As well as enhancing the scope of simulation using different methods, enhancing the modelling approach by integrating theories and knowledge from other disciplines can support study outcomes. For example, how to improve or maintain health is an important component of the social aspect in the sustainability triple bottom line model (Moon, 2017). Increasing demand for services combined with financial pressure is a challenge to sustainability in healthcare, which may have a spiralling negative effect on society and patient care, and Fakhimi et al. (2015) argued that simulation models should explicitly consider this dimension. Dode et al. (2016) proposed a methodology to apply Human Factors (HF) principles in DES to allow the application of HF early in an engineering design process, before operators are put at risk. This takes into account the mechanical and mental loading that people are subject to when performing a specific task, showing that it is possible to design production systems that are more productive and less hazardous for the system operator. They argued that models that fail to include human aspects may provide unreliable results in terms of productivity and quality estimates, where HF may be a 'missing link' in DES by accounting for aspects such as cumulative load, psychosocial factors, injury, task unfamiliarity, or learning (Perez et al., 2014). The complementary roles of DES and QI improvement methodology have been investigated. In healthcare, this has been applied in the SimLean methodology (Robinson et al., 2012) based on the proposition that the symbiosis of lean and DES approaches has unrealised potential. Also in healthcare, Komashie et al. (2015) modelled QI (patient and staff satisfaction) using queuing theory. In a social system consisting of people on both the demand and supply sides, it is apparent that there is still work to be done in these areas. Finally, Monks (2015) recommended that OR and implementation science could be mutually beneficial disciplines in healthcare implementation problems. Proposed roles for OR are structuring implementation problems, prospective evaluation of improvement interventions, and strategic reconfiguration, while challenges include reaching mutual understanding and the lack of evaluation and evidence supporting OR interventions. Healthcare differs from other domains with its permeable boundaries between organisations, and multiple levels of funding and compliance demands, hence Long and Meadows (2018) recommended exploring models of stakeholder engagement and implementation designed specifically for this complex environment. Inter-disciplinary research with implementation science might be one path toward supporting OR in both implementing the results of the work, and addressing healthcare problems that represent relevant real-world priorities.

One relevant real-world problem in healthcare is its rapid operational change, and the need to constantly adjust behaviour and activities according to variable and rising demand. This is particularly the case in ED, where the pace and unpredictability adds a specific challenge for simulation modelling studies. In the age of Industry 4.0, based mainly on the concept of Cyber-Physical Systems, that is, the integration of computing, communications and control (Aceto et al., 2020), there is an increasing challenge to positively impact the access, efficiency and quality of healthcare processes. Industry 4.0 is revolutionising the manufacturing sector (Bonci et al., 2016; Rodič, 2017), and in healthcare, the rapid technological evolution of Internet of Things (IoT), big data, and cloud computing is beginning to have a similar impact (Aceto et al., 2020). The presence of wireless and mobile technologies, medical software mobile apps, low-cost wireless sensors, wearable IoT, and the technologies designed to extract value from large volumes and a wide variety of data has the potential to influence healthcare at many levels. However as the world becomes more connected, with increasing dependence on machines and simulations to make decisions on our behalf, it is critical that the data from sensors, artefacts, and devices is trustworthy and secure. This raises concerns regarding privacy, security and trust (Aceto et al., 2020; Onggo et al., 2020).

In manufacturing, Gualtar (2018) proposed future research for DES to comply with Industry 4.0. Simulation has been mostly used for the development of standalone solutions with a limited scope and lifetime, however increasingly simulation development is shifting to integrate simulation models into decision-support tools for recurrent use (Rodič, 2017). One challenge is automated data exchange, in particular receiving data from sensors, machines or other data acquisition systems, and integrating it into the simulation model. Gualtar (2018) outlined a range of approaches, from the use of intermediary artefacts to direct integration of the simulation model with the data sources, enabling real-time reconfiguration and re-run of the simulation with updated data. A key constituent of Industry 4.0 is the ‘digital twin’, a real-time digital representation of a manufacturing facility (Rodič, 2017). These representations can be enhanced using DES to test scenarios in real-time. DES supports short-term interventions in the system by providing insight into complex systems, and the digital twin concept has resulted from significant advances both in data collection and M&S, resulting in the simulation being a core functionality supporting operational performance by direct integration of operational data (Weyer et al., 2016). It contains all information that is needed by various stakeholders, but this represents a fundamental challenge, requiring digital continuity, real-world synchronisation and multi-disciplinarity (Weyer et al., 2016). Additionally, speed can be an issue with continuous experimentation, especially for detailed models which require updating (Taylor et al., 2019). In the healthcare domain, specific challenges exist, as summarised by Jimenez et al. (2020). Cybersecurity of cloud computing presents risks including data breaches, challenges in data privacy and integrity; IoT challenges include low-speed processors, limited memory, compatibility, and security; software challenges include usability and reliability; certification and regulatory approval processes for medical devices is an ongoing challenge; and security and privacy of interoperable technologies, which could include attacks on data, the device, or the institution are all significant risks. For these reasons, digital twin research in healthcare will lag that of manufacturing, to ensure the status and assurance of system and patient data. Nonetheless, there is increasing research interest in the advantages to real-time simulation for short-term decision-support, and a number of researchers have started to address the challenges. The next section defines real-time simulation, and reviews its use in healthcare.

2.6.4 *Real-time simulation*

2.6.4.1 *Introduction*

Simulation-based methods are important tools for decision-support in domains such as traffic control, supply chain modelling, manufacturing, and healthcare.

While traditional simulation analysis using probability distributions from historic data can be used to generate and test scenarios, it can be time-consuming to keep it updated and validated for recurrent use. Additionally, using historical data means that the model can be inaccurate in the short-term (Bahrani et al., 2013). This is particularly a problem in dynamic systems where historical data becomes out-of-date (Tavakoli et al., 2008).

Real-time simulation has been proposed as a solution to the above problems (Tavakoli et al., 2008; Turner, 2011; Bahrani et al., 2013) whereby a simulation model is integrated with an automated data acquisition system (Uhlemann et al., 2017). The purpose of the real-time simulation is to serve as a means of projecting the development of a situation in an existing system over a short time period, supporting short-term operational decisions. It is particularly useful for dynamic, goal-directed decisions in systems that continuously make decisions in real-time (Dalal et al., 2003). The real-time simulation is initialised and driven by real-time or near real-time data. This data links the information system with the simulation model, to provide actual performance, and can add flexibility to the monitoring of operational systems (Altaf et al, 2016). It requires both a validated simulation model of the physical system and real-time inputs. Subsequently, the model and its multiple runs must be completed in a short time-frame in order to be used in ensuing decision-making processes (Hanisch, 2005).

2.6.4.2 *Defining real-time simulation*

The execution of real-time simulation has been in use in manufacturing systems for decades, with Annan and Banks (1992) describing one of the earliest unifying frameworks for connecting the real-world system and the control system, termed ‘knowledge-based on-line simulation’. They defined 'on-line simulation' as a computerised system capable of performing both deterministic and stochastic simulations in real- or near real-time, for evaluating alternative control policies on a shop-floor. For dynamic scheduling decisions, Rogers and Gordon (1993) surveyed the literature on the use of simulation for supporting real-time decision-making in manufacturing. For these authors, ‘real-time simulation’ was defined as ‘fast enough to be useful’. In the following years, a number of studies investigated real-time simulation in the manufacturing domain (e.g. Drake & Smith, 1996; Ruiz-Torres & Nakatani, 1998; Lee & Fishwick, 1999; Son & Wysk, 2001). In 2002, the Grand Challenges for Modelling and Simulation (Fujimoto et al., 2002) proposed the use of the term 'symbiotic simulation' to describe a simulation which can dynamically accept and respond to real-time data from the physical system to improve the accuracy of the model. This conceptualisation emphasised the mutual benefit between the simulation and the physical system, such that the execution of the simulation and the real-time interaction with the physical system is continuous. Aydt et al. (2008a) subsequently relaxed this definition, proposing that symbiotic simulation is "*a close association between a simulation system and a physical system, which is beneficial to at least one of them*". In a closed-loop symbiotic simulation system there is a control feedback from the simulation to the real system. In an open-loop symbiotic simulation system there is no such feedback. The term Dynamic Data-Driven Application Systems (DDDAS) describes a similar concept, and was used in support of emergency medical treatment decisions by Gaynor et al. (2005). However in the manufacturing domain, the term 'symbiotic simulation' continued to dominate (Low et al, 2007; Aydt et al, 2008b; 2009a,b; Fanchao et al., 2009). As computing technology has evolved, so have simulation tools. From 2010 onward, there has been an increase in the research area of real-time simulation and its various challenges. A 2010 review of 'real-time' simulation identified applications in power generation, automotives, transport, aerospace, and education (Bélanger et al., 2010). Around the same time, the approach began to be proposed in healthcare (Tavakoli, Mousavi & Komashie, 2008; Marmor et al., 2009; Mousavi et al., 2011). Onggo (2019) and Onggo et al. (2020) defined real-time simulation as 'as fast as clock time', and prefer the term 'symbiotic simulation system'. Although the above terms all continue to be in use, for this thesis the term 'real-time simulation' is used, to describe a validated simulation model which is triggered, and initialised using real-time or near real-time data from an ED operational system, for short-term decision-support. This enables a wider conceptualisation and application of the data, including descriptive and predictive analytics, in combination with simulation.

2.6.4.3 *Real-time Simulation in healthcare*

To date, few healthcare applications of real-time simulation have been published, with research into the application of real-time simulation lagging behind that of other industries. Tavakoli et al. (2008) proposed a generic framework supporting the application of real-time simulation adapted from manufacturing to healthcare, using an existing ED model. As in manufacturing, they proposed that entities (patients) required Radio Frequency Identification (RFID) to track their journey. Mousavi, Komashie and Tavakoli (2011) developed the Simulation-Based Real-time Performance Monitoring (SIMMON) framework proposed for continuous system QI monitoring and shorter lead times for response and improvement, by providing quality measures such as staff satisfaction. Again RFID technology was required, and like the framework proposed by Tavakoli et al. (2008), this novel approach for timely QI responses remained untested. The use of RFID for ED simulation has yet to be realised, however Espinoza et al. (2014) noted that in manufacturing processes, resources are typically stationary, while entities move. Here, the use of RFID or other Workflow Management Systems allows easy monitoring of the state of resources and the system at any given time. However in ED, patients are often stationary while resources are moveable. One practical challenge of ED M&S is that of untangling actual patient pathways which loop back on themselves and cross department and system boundaries (Young et al., 2004).

An alternative to using RFID is to access data directly from data management systems, where appropriate data exists. This is significantly limited by the need to gain access to the data in real-time, and of ensuring its quality and availability. However Espinoza et al. (2014) compared a minimal real-time ED data set in a real-time simulation model with an ideal real-time data input scenario, and found no significant difference between the two, suggesting that for many decision-making scenarios, a simple automated data system is useful for representing the system state realistically for decision-support. Also using operational data, a DES model for real-time estimation of the current operational state of ED, and a short-term prediction of future states was investigated by Marmor et al. (2009). The researchers used real-time simulation to estimate the current operational state, and to create short-term predictions regarding future ED states, where the estimation and prediction is based on incomplete or even inaccurate data. Six patient-type arrivals were inputted into the model in real time, while discharges were adjusted according to historical data for each patient-type. Forecasting using long-term moving averages was done to predict hourly arrival rates, based on the last fifty of the same hour on the same weekday, while service time per patient type was estimated. Through a decision-support dashboard, decision-makers were able to plan resource allocation for the next several hours to handle resource scarcity. This was a comprehensive project that considered both the staff and the patient perspective, aiming to provide a real-time estimation of the current operational state, and a short-term prediction of future operational states for staff planning, yet remained at the pilot stage.

A real-time simulation prototype using optimisation was proposed by Tan, Tan and Lau (2013) to adjust the number of doctors based on current and historical information about patient arrival, to enable the ED to better cope with demand surges. A similar approach was taken by Bahrani et al. (2013), who investigated the real-time impact on wait times of resource allocation scenarios. This prototype model requires manual initialisation at runtime, and manual setting of experimental scenarios, which include opening beds or adding staff. The prevalent scenario for addressing ED crowding is staff rescheduling (e.g. Badri and Hollingsworth, 1993; Beaulieu et al., 2000; Sinreich and Jabali, 2007). One criticism of this is that despite being formalised in ED escalation polices, tactics such as bringing in ward staff to assist or calling in additional doctors from home have been found to be either difficult in practice, or to significantly increase staff workload through patient handovers and inadequate skill levels (Back et al., 2017). This emphasises the need to investigate appropriate scenarios for real-world impact.

Adra (2016) outlined how real-time simulation can be used for descriptive (real-time visibility), predictive, and prescriptive purposes. For real-time forecasting of ED operating conditions, Hoot et al. (2009) developed and validated a DES model, whose purpose was for prediction of a range of operational indicators. An advantage to this model is that it has two outputs, one for discharges, and one for those waiting to be admitted, accounting for downstream delays in ED which contribute to crowding. Onggo (2019) proposed a framework for symbiotic simulation that could be used for control, prediction or prescription, in the context of Industry 4.0, emphasising real-time or near real-time SA by making use of cyber-physical systems and enabling technologies. Its potential application in urgent and emergency care was discussed in Onggo et al. (2018). Using a modular simulation model and a process analyser tool allowing design-of-experiments, Augusto, Murgier and Viallon (2018) proposed a prescriptive framework for real-time simulation in ED planning. The design-of-experiments features main decisions to take in order to reduce patient length-of-stay and service occupancy for the management of the service and the activation, if required, of exceptional measures in the event of crisis. This allows managers to choose the best decision from a wide range of scenarios to optimise operations.

Due to the time-critical nature of ED, the majority of the above proposed and prototyped studies were situated in the ED domain. Oakley et al. (2020) used a real-time (symbiotic) proof-of-concept DES model for hospital bed management of elective and emergency patients beyond the ED. They focussed on validation and application, and demonstrated the applicability of such an approach to support bed planning, using a tranche of hospital operational data. While interest in real-time simulation in healthcare continues to rise as the feasibility of leveraging data increases, a number of proposed, conceptualised and prototyped models have demonstrated the applicability of the approach. However in practice, gaining access to healthcare operational data or sensor data in real-time appears to be a significant obstacle. Real-time simulation is one example of hybrid systems modelling. This is defined and discussed in the next section.

2.7 Hybrid systems modelling

2.7.1 *Definition*

Mustafee and Powell (2018) differentiated between *Hybrid Simulation* (HS) which applies two or more simulation techniques at the implementation/model development stage of a simulation study, and *Hybrid Modelling* (HM), which combines simulation with methods and techniques from OR or other disciplines to any stage of a simulation study.

Using Minger's and Brocklesby's definitions of paradigm, methodology, technique and tool, Mustafee and Powell (2018) identified four cases and proposed a unifying conceptual representation that categorises these types as Types A, B, C and D. (Figure 2.3).

- *Type A (Multi-Methodology Hybrid Simulation)* combines continuous and discrete simulation methods, and there are numerous examples of studies that have combined SD-DES and SD-ABS.

- *Type B (Multi-Technique Hybrid Simulation)* combine discrete methods, e.g. DES-ABS or continuous methods e.g. Computational Fluid Dynamics Modelling (CFD) with SD.
- *Type C (Multi-Methodology, Multi-Technique Hybrid Simulation)* is mix of both of the above, e.g. ABS-DES-SD.
- *Type D (Hybrid Systems Model)* combines a simulation method with either a quantitative method such as optimisation or machine learning, or a qualitative method (*Type D1: Multi-Paradigm Hybrid Systems Model*).

Models of Type D and Type D.1 specifically refer to HMs applied to both quantitative and qualitative OR.

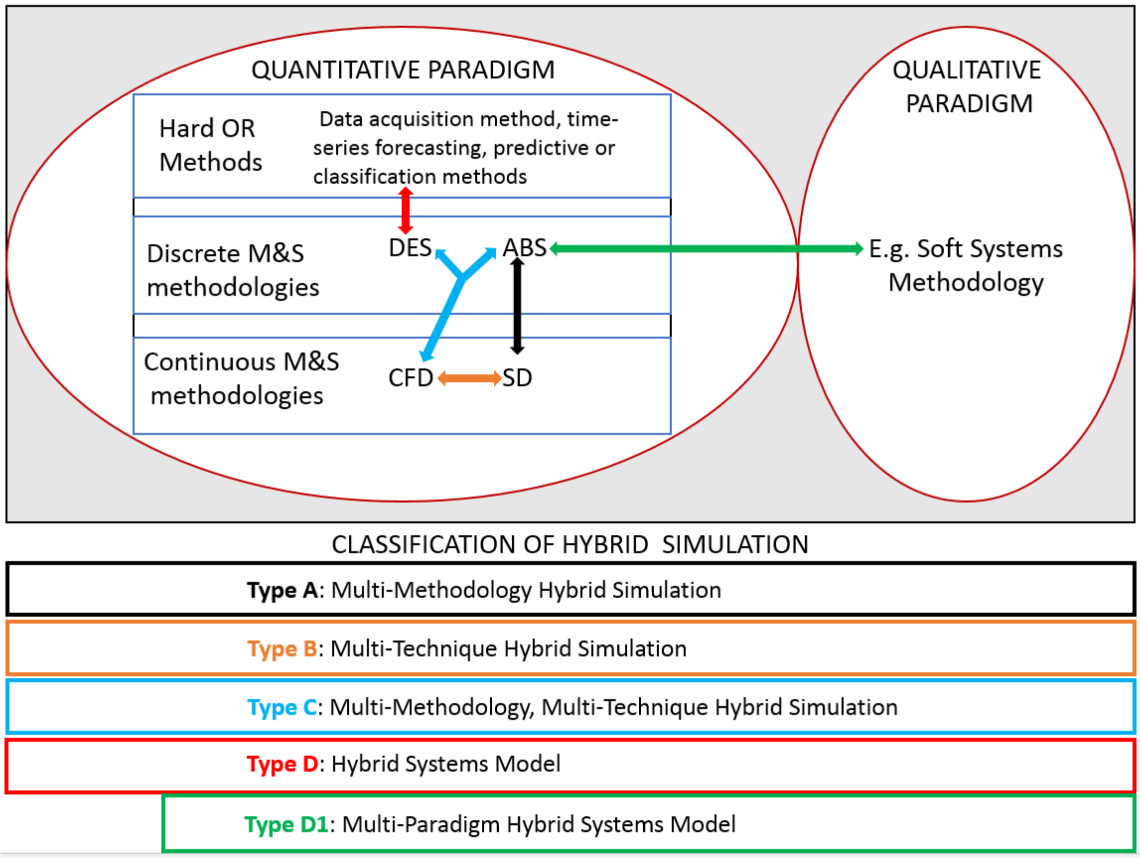


Figure 2-3 Unifying HS-HM Conceptual Representation using Classification of Hybrid Simulation (Types A-D) with examples. Adapted from Mustafee & Powell, 2018

A modelling and simulation (M&S) study consists of several stages, and a hybrid M&S study applies HM to one or more stages of an M&S study. Figure 2.4, taken from Powell & Mustafee, (2016), depicts four simulation methods that are frequently used separately, or combined as hybrid simulation in OR. Examples of other techniques from OR or other disciplines that can be applied to other stages are depicted using the M&S stages Model Conceptualisation, Model Formalism, Input Data Analysis, Output Data Analysis and Simulation Experimentation.

Examples of Type D1 HMs include Soft Systems Methodology and other Problem Structuring Methods which can be used in the problem formulation/conceptual modelling stage of an M&S study to enhance requirements capture. Similarly, Greasley & Edwards (2019) conducted a review of studies which have used big data analytics to enhance a DES study, mapping the studies to different stages of the M&S lifecycle. These are examples of Type D HMs (Mustafee & Powell, 2018).

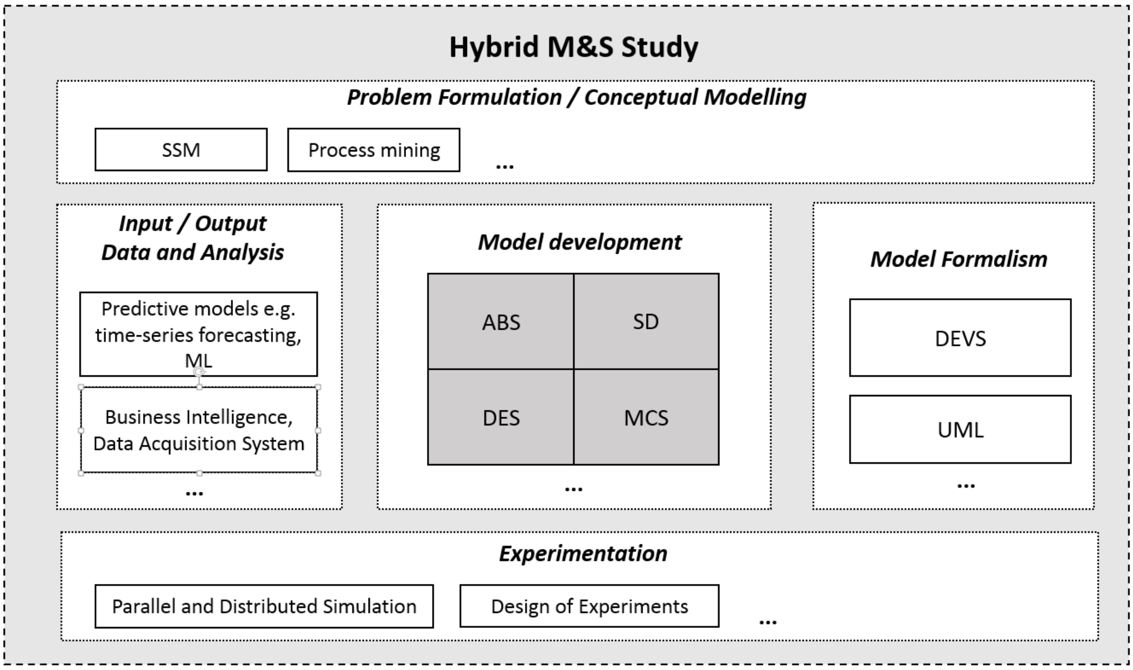


Figure 2-4 Conceptual representation of a hybrid M&S study (… denotes other methods) Adapted from Powell and Mustafee (2016).

The rationale for combining approaches has been described as synergising the methodological strengths of each method or to better capture the breadth of a problem situation (Sachdeva et al., 2007; Mustafee & Powell, 2018). Real-time simulation is one example of hybrid modelling, whereby a simulation model is combined with a data acquisition system to initialise the model. From their search criteria, Greasley and Edwards (2019) identified one example of a real-time simulation, which they categorised into the ‘model-building phase’. They note that real-time DES models pose particular challenges, with models needing to re-adjust at initialisation and consistently perform validation, analysis, and optimisation. Distributed simulation architectures may be needed to provide speed of execution (Taylor, 2019), and an architectural framework for the interaction between the physical and simulated system is needed (Onggo et al., 2018). In Figure 2.4, a real-time data acquisition system is categorised as 'input data analysis', as the data is acquired, visualised, processed, and inputted into a validated simulation model.

2.7.2 *Application and challenges for a HM Approach*

Established processes and conceptual frameworks supporting or providing guidance for HM approaches remain under-developed. Conceptual modelling is a critical stage of an M&S study, referring to the abstracting of a model from a real or proposed system, providing a specific set of steps that will guide the modeller on the translation of the conceptualisation into a formal model (Robinson, 2008). Zulkepli and Eldabi (2015) argued that the same is true for hybrid modelling, where the focus of previous attempts were at the software level. They presented a frameworks for HS that aimed at improving the conceptual modelling stage. This has relevance to HM, which can use a number of methods at any stage of an M&S study. For this reason, the conceptual modelling stage is critical, in particular the contribution and nature of communication between components of the hybridised model at the conceptual stage.

Zulkepli & Eldabi (2015) proposed a 3-phase framework for developing hybrid models. The three phases of the framework are: the conceptual phase; the modelling phase; and the model communication phase. The *conceptual phase* develops conceptual modules that represent the problem such that each cannot be divided into smaller chunks, while together they represent the system. The *modelling phase* is concerned with translating the conceptual model into a simulation model. The hybrid process between the two (or more) different techniques and software requires consideration. The final stage is concerned with the *communication between the models*. This involves identifying how variables from models communicate and change their impact on variables in other models. This framework can be aligned with Type D HMs, which share the same (or a complementary) philosophical paradigm, but can address different aspects of a problem situation. Modularising the conceptual model ensures the overall approach is not excessively complex, and keeps the purpose of the approach central at each stage. These modules are linked using their outputs i.e. output from one module will serve as an input to the next module, and how these variables influence other variables in the model requires specification.

Lynch et al. (2014) also focused on an explicit conceptual model component as a link between exploring the problem situation and building the model, in their HS framework. Within the conceptual phase is formulating the modelling question and determining whether a hybrid methodology is required based on the question. It has been argued that a sound rationale is required for hybridising approaches (Chahal et al., 2013). Howick and Ackerman (2011) found a number of rationales for mixing OR methods by analysing published case studies. These included dealing with complexity, supporting different stages of a project, combining the benefits or overcoming the weaknesses of specific methods, gaining credibility of the model, and considering the wider system. However they pointed out that there has been limited connections made between the rationale for taking a HM approach and the actual outcomes of the project.

Challenges exist when combining methods, which need to be understood. The danger of abduction risk was underlined by Lorenz and Jost (2006), where assumptions are layered in hybrid approaches, leading to the acceptance of wrong conclusions. The development of frameworks to support HS have highlighted that extending methods to incorporate the characteristics of other methods can lead to unnecessarily complex models with more assumptions and approximations (Chahal et al., 2013). A final challenge exists in validation and verification, which can be more complicated in hybrid approaches (Lynch et al., 2014; Viana, 2014; Eldabi et al., 2016). A generic verification framework may assist modellers in ensuring the challenges associated with competing model characteristics do not cause errors within their simulations (Lynch & Diallo, 2015). Lynch et al. (2014) provide some suggestions for validation of hybrid models, and an example in their use-case. Additionally, studies have investigated and strengthened aspects of validation of real-time simulations in healthcare (Hoot et al., 2009; Oakley et al., 2020).

There is a need for integration frameworks for HMs, with a well-defined set of guidelines for the integration of analytics models (e.g. forecasting, simulation) with real-time or near real-time data. The challenges can be conceptual, for example philosophical compatibility and parsimony; and technical, for example integration methodology, open access to software implementing different elements of the hybrid model (Onggo et al., 2018). This will support modellers to gain a better understanding of the complex system, allowing them to assess the problem from different dimensions and build a model that better meets the need of stakeholders (Zulkepli & Eldabi 2015; Mustafee et al., 2015a). The benefit of choosing any modelling approach is to enable the model to achieve its purpose more effectively, and Brailsford (in the position paper Mustafee et al., 2017a) argued that ultimately hybrid approaches will lead to more useful models that better represent the real-world problem asked of it, and provide better solutions.

2.7.3 *Applicability to healthcare*

Given the complexity of the healthcare domain, and the continuing focus on the gap between academia and practice, it is not surprising that a large number of hybrid simulation studies in healthcare exist. These are driven by the need to enhance the scope of the study and capture multiple aspects of a real-world problem. The review of HS in all domains by Brailsford et al. (2019) found that healthcare was the main area of application, and the popularity of HS in this application area was suggested to be due to the intrinsic complexity of the problems, reuse and adaptation of models to expand the scope of the model, alongside research interests of the authors. The authors stated that healthcare problems have multiple aspects, and it is rarely possible to capture all of them in one single model using only one method. There is likely to be a similar motivation for HM applications, and a growing number of these exist.

There are many examples of Type D HMs in healthcare. Royston (2013) argued that there is a mutual advantage in combining analytics and OR more explicitly together, including strengthening links to real-world concerns. Delen and Zolbanin (2018) agreed, stating that the use of analytics can provide more reliable information about the structure of relationships between variables, and generate more relevant research by using appropriate tools for a given problem. Analytics expands the range of tools available for OR applications (Ranyard, Fildes, & Hu, 2015), with both sets of methods used to convert data into actionable insight for timely and accurate decision-support (Sharda et al., 2017). Greasley and Edwards (2019) differentiate analytics as data-driven - with a focus on data and outputs, but potentially little knowledge of underlying processes - and simulation as model-driven, with considerable knowledge about underlying processes. In healthcare, many studies have exploited this synergy, outlined later in Sections 2.7.5.

Similarly, studies promoting the combined use of qualitative methods with quantitative OR methods in healthcare and other social systems have argued that reductionist approaches used in isolation may capture insufficient understanding of the nature and context of complex issues, and fail to secure buy-in from stakeholders (Sachdeva et al., 2007; Franco & Lord, 2011; Crowe et al., 2017). This discussion is extended in the next section.

2.7.4 *Hybrid modelling approach qualitative methods*

The term 'multi-methodology' in Management Science has been used to describe the combined use of two or more methodologies within a single intervention (e.g. Mingers & Brocklesby, 1997; Mingers, 2000). This aligns with Type D1 HM in the classification scheme proposed by Mustafee & Powell (2018). The justification for combining qualitative and quantitative methods in OR is to more effectively deal with the breadth and nuance of the real world (Mingers & Brocklesby, 1997; Franco & Lord, 2011). Quantitative OR methods are used not simply for analysis, but for decision-support, suggesting that a measure of study success is the degree to which the outcomes successfully support decision-making. However decision-makers must be able to trust the results of the analysis, and believe in the credibility of the model and the modeller, and a significant amount of research has demonstrated that the quality of stakeholder engagement influences this (e.g. Jahangirian et al., 2016; Crowe et al., 2017; Long & Meadows, 2018). Most simulation studies in complex systems such as healthcare will require some degree of stakeholder engagement in order to understand the problem-situation and validate the resultant model. For example, Crowe et al. (2017) judged the effectiveness of their approach by the material impact it had on the process of drawing conclusions from the work, rather than implementation of results. This is complicated by the challenge of determining whether to attribute success or failure of an M&S study to the choice of methods, contextual factors such as personal attributes of the researcher, or dynamics of the team involved (Connell, 2001; Henao & Franco, 2016).

Type D1 HMs in healthcare have included using SSM or cognitive mapping in the conceptual modelling phase to support stakeholder participation, encourage plurality of viewpoints, and support validity and credibility of the conceptual model (Powell & Mustafee, 2014; Kotiadis et al., 2014; Pessôa et al., 2015). Facilitated modelling supports the model building, scenario selection, and decision-making processes (Fokkinga et al., 2009; Franco & Montibeller, 2010; Kotiadis et al., 2013; Tako & Kotiadis, 2015; Proudlove et al., 2017); and facilitation can also be used to successfully support the implementation stage (Kotiadis & Tako, 2018). While these, and similar, approaches bring significant advantages in enabling a focused and common understanding on real-life issues, drawbacks include additional time, preparation, and skills required (Tako & Kotiadis, 2015).

2.7.5 *Hybrid modelling approach quantitative methods*

A growing interest in hybrid approaches has been explained by Jahangirian et al. (2010) as the common belief in the mutual impacts that different parts of systems have on each other. However it can also be used to gain new knowledge or insights, in particular with the current interest in data-driven knowledge. For example, analytics can be used with simulation to assist with conceptualisation of systems or problem formulation (Augusto et al., 2016; Elbattah & Molloy, 2017); for analysing input data into simulation models (Glowacka, Henry & May, 2009; Garg et al., 2009); at the experimentation stage (Elbattah & Molloy, 2016) and for analysing simulation output (Rabelo et al., 2014). The need for constructing reliable representations of real-world problems, and for fact-based decision-making in healthcare, makes this an area worthy of exploration and understanding.

Data-driven methods are increasingly used to leverage evidence-based insights from hospital operational data. For example, as touched upon in Section 2.6.3, various efforts have been made to combine Process Mining with DES for re-designing healthcare processes (e.g. Augusto et al., 2016; Rojas et al., 2016; Abohamad et al., 2017). The main advantage to this combined approach is that Process Mining aids the development of the conceptual phase of a DES model in a semi-automatic way by analysing the event log and discovering a structured process flow, which can then be used to develop the model directly from the data (Zhou et al., 2014; Wagner et al., 2016). A second advantage advanced by Abohamad et al. (2017) is a reduced need for domain experts in DES development, although they will still be required to verify the model, as inaccuracies or errors in data could have been recorded. Where healthcare staff have limited availability, and may not have full sight of a process, Process Mining can support the development of unbiased simulation models (Abohamad et al., 2017), and process flows with very low frequency can give useful insights for analysing exceptional behaviour (Turner et al., 2012; Abohamad et al., 2017).

With a similar aim and again touched upon in Section 2.6.3, Abdelbari and Shafi (2017) were interested in exploring the extent to which machine learning (ML) can be used to infer conceptual models as part of the SD modelling process. A recurrent neural network was used to automatically learn causal loop diagram-like structures directly from system data. The proposed data-driven approach aimed at complementing the development of a conceptual model by providing modellers with several probable model structures that can be accepted or considered for refinement. For input data analysis, several authors have demonstrated the potential for using machine learning or data mining algorithms for defining DES patient types or care pathways (e.g. Isken and Rajagopalan, 2002; Codrington-Virtue et al., 2006; Ceglowski, Churilov & Wasserthiel, 2007; Glowacka et al., 2009; Elbattah & Molloy, 2016; 2017). For example, Glowacka et al. (2009) used association rule mining to generate decision rules for patient non-attenders. This method embeds a subset of rules as conditional and probability statements in the DES model. This means that the variables do not need to be traded off against each other and the rule-based model is easy to explain to stakeholders, an important consideration. Elbattah and Molloy (2016) described an approach that combined data-driven ML and DES. The ML models made predictions about the inpatient length of stay and discharge destination of the simulation-generated patients. On a population basis, the simulation model provided demand predictions for healthcare resources related to discharge destinations. The significance and complexity of discharge planning has increased due to the rising challenge of population ageing, and this study found that the current distribution of nursing homes may not meet the needs of the ageing populations in some geographic areas of Ireland.

At the experimental stage of an M&S study, Delen et al. (2011) used simulation in combination with data mining, optimisation and GIS-based analytics to model a blood supply chain with a high level of complexity that the authors argued could not be handled by a single method alone. An implementation of this system is being actively used at different levels of a Defence supply chain. Uriarte et al. (2017) offered an approach to healthcare decision-making that combined DES, simulation-based multi-objective optimization, and data mining for the analysis of the results of the DES optimisation. They reported that hospital stakeholders agreed that the knowledge this approach offered is valuable, while the authors were clear that this combination of methods reduced the drawbacks of each technique when they are applied alone.

Visual analytics presents large-scale data in a visual form, allowing insight into the data, and interaction with the data to confirm or disregard those conclusions (Feldkamp et al., 2015). Visual Analytics has the potential to provide a useful additional tool when interpreting simulation output data, for example Feldkamp et al. (2017) used visual analytics at the experimental stage based on a binary decision tree that maps the relationship between simulation input and output factors. This approach is particularly useful for big data applications by synthesising large amounts of data to reveal patterns otherwise not readily seen.

For short-term decision-making, forecasting and other prediction methods have been used extensively in healthcare to support change based on predictions of a future state (e.g. Makridakis et al., 2018; Kraaijvanger et al., 2018; Zhou et al., 2018; Kaushik et al., 2020). Time-series analysis can provide accurate forecasts of future ED attendance for allocation of resources, such as optimum staff scheduling by day and time (Morzuch & Allen, 2006), however few studies indicate how demand forecasts can be used for planning. Boyle et al. (2012; 2016) acknowledged the need to capture real-world benefit from forecasting ED demand, such as identifying appropriate triggers for escalation responses. A logical extension of demand forecasting in ED is the use of DES to plan for recovery based on forecasts of overcrowding.

Such an approach was used by Park et al. (2008), who used a linear regression forecasting model to predict average daily arrivals. A DES model, as a separate component, was used to establish efficient ED staff scheduling to decrease non-value added patient wait-times and increase the quality of patient care. A similar approach was used by Lin and Chia (2017), who proposed a combined ARIMA forecasting approach with DES to forecast ED patient arrivals and subsequently reschedule doctors. Forecasts were made daily, allowing time to reconfigure staff rotas. In practice, this may be difficult to implement, however the results found improvements in wait times. This approach indicates the potential value in forecasting a parameter such as patient arrivals, combined with simulation to determine how best to align the system with the forecasted demand. The combined approach supports forward planning over a short time period. For rapid decision-making, this could be enhanced by the use of real-time data to improve the accuracy of the simulation model in the short-term.

Many examples exist of successful HM approaches which have used mixed methods to enhance the success of the simulation or the overall study in healthcare, however healthcare can also learn from other domains. For example in manufacturing, Aydt et al. (2008b) were interested in response times in real-time simulation, in particular where timing is critical in resolving a problem. They proposed the use of 'preventative what-if analysis' using forecasts of a critical condition indicator in the real system, which, once detected, triggers a what-if analysis. This means the real system can be reconfigured before the critical condition occurs, compared with a reactive approach, which causes a more drastic performance drop. Augusto et al. (2018) proposed a similar approach in their real-time M&S framework for supporting emergency units in times of crisis, to predict precise arrivals using data history. Their framework consisted of a modular simulation model connected with a process analyser tool, allowing a design-of-experiments based on requirements of healthcare stakeholders. Their proposed model intends to be automated via the hospital information system, to take into account the number of patients in the system, the occupancy of resources, and the history, to predict future arrivals, and to require minimal manual interaction.

From reviewing the literature, it seems that for a real-time HM to be useful for short-term decision-making, it must provide timely decision-support with limited interaction required by stakeholders, requiring a HM solution consisting of real-time data, predictive analytics, and simulation. This approach aims to support SA and short-term decision-making toward both adaptive actions and formal escalation responses. This is illustrated in Figure 2.5.The chapter summary will summarise this review toward addressing RQ1, and outline the gaps in the literature to be addressed by this research, and the criteria for determining the success of the proposed HM application.

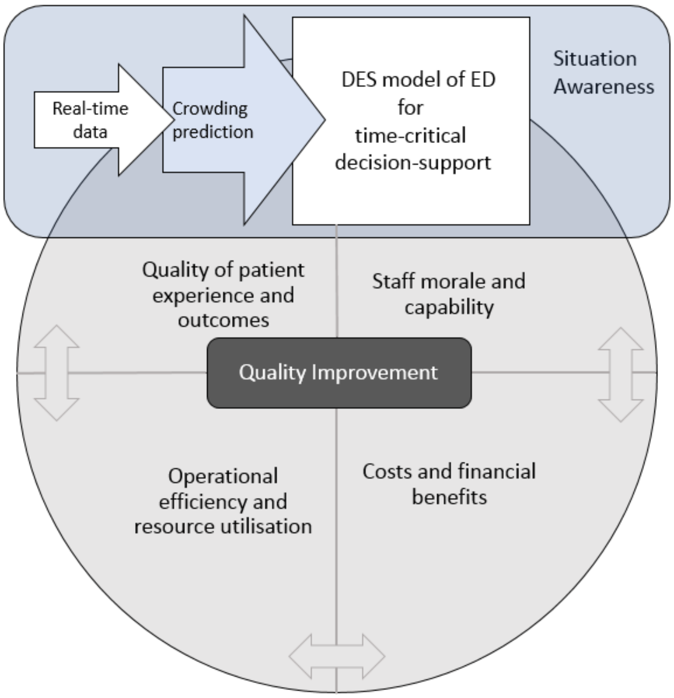


Figure 2-5 The relationship between real-time data, data analytics, SA and QI for healthcare short-term decision support

2.8 Chapter Summary

The purpose of this chapter has been to address the first research question:

RQ1: *How can simulation approaches support short-term operational decision making in healthcare?*

This has been achieved through determining the need for short-term decision support in healthcare, and examining how simulation and hybrid modelling approaches have been used for short-term operational decision-support in the healthcare context, and the opportunities these approaches offer.

In doing so, it has addressed the following objectives:

Objective 1: To understand the need for short-term decision support in healthcare, in particular emergency care.

Given the current public health situation, NHS healthcare services are likely to become progressively more constrained, yet quality must remain a driver of healthcare service interventions. Quality considers both safety and perceived value, and should underline all attempts to improve efficiency and productivity. Taking a systems view is important with emergency departments which operate as the ‘front door’ to the hospital. Workload can be unpredictable hour-to-hour, and crowding, where demand exceeds available capacity, has been an increasing problem. The effects of crowding are seen in both patient outcomes and staff morale (Morley et al., 2018), as well as in operational performance. A widely used conceptual framework of crowding is the input-throughput-output model (Asplin et al., 2003). Input relates to the demand for ED services and any condition that contributes to this; throughput relates to internal ED processes; and output factors are related to disposition of patients to discharge, admission, or transfer to another service. The multifactorial nature of the problem does not suggest a single solution, and the focus is on optimising existing capacity and resources available to ED. However when care is operating close to the boundaries of capacity, the risk of a critical event occurring is high. To enhance system resilience, staff must be enabled to accurately determine the state of ED, assess the impact of corrective actions and launch action. Clearly, the ability to improve anticipation of a degraded situation would allow corrective actions, such as demand management, to be initiated earlier.

Demand management can take two forms. Firstly, patients can be provided with additional information that can support decisions about the most appropriate place to attend; secondly, demand management can take the form of redirecting appropriate patients to alternative services as queues become unmanageable. Decisions regarding the redirection of patients are short-term decisions, and require situational assessment and situational awarenss (SA). However as workload and crowding increase, and staff stress-levels rise, SA reduces. Performance, which can be managed and adapted to a point, will ultimately decline (Nählinder et al., 2004). As operational targets make controlling task-load demand difficult, interventions which support SA and system resilience become increasingly important. With a focus on the needs of the users of the system, one approach to improving SA and supporting effective performance is paying attention to what type of information is needed when and by whom to support system goals.

Objective 2: To explore how analytics methods can be used for short-term decision support.

Even for experienced staff, there is a need to enhance SA to support effective decision–making and system resilience. Data-driven approaches are increasingly gaining traction in healthcare for this purpose, as the value in data becomes increasingly realised, alongside its rapidly growing quantity, quality and availability. The most widely used functional categorisation for data analytics loosely groups techniques hierarchically as descriptive, diagnostic, predictive and prescriptive methods. Additionally, data analysts have shown a rising recognition of the value in using qualitative methods alongside data-driven approaches, for example for diagnosing a system problem, or evaluating the impact of an intervention.

M&S is an example of a prescriptive method, which provides insight and understanding into the physical processes of a system, and the knowledge gained may be of great value toward suggesting improvements in the system under investigation. Hybrid modelling (HM) may further enhance its value, by better capturing a problem situation, or by aiming to combine the benefits or overcome the weaknesses of individual methods. M&S has been used extensively in healthcare, although evidence of its real-world benefit remains limited. Implementation is not the only value offered by M&S, and participatory and qualitative approaches may enhance an understanding of its value.

<u>Objective 3:</u> *To evaluate simulation approaches used in healthcare for decision-support and to identify how simulation is used in other domains for short-term decision-support.*

The scale and scope of ED M&S is enormous, however the pace and unpredictability of ED adds a specific challenge for simulation modelling studies. The majority of ED models provide tactical or strategic decision-support, and are parameterised using historical data, meaning they can be inaccurate for short-term simulation. Real-time simulation has been proposed as a solution to this problem in healthcare, whereby a simulation model is integrated with an automated data acquisition system to improve its accuracy in the short-term. Increasing interest in this area reflects the increasing availability, volume and velocity of data, however to date, published studies are conceptual or prototyped, with no study yet integrating real-time data in their model.

Real-time simulation is used to mitigate a critical situation once detected, however a HM using predictive analytics to forecast a future critical situation would allow health-care practitioners to proactively use real-time simulation to recover before the onset of the critical situation. In comparison, a reactive approach can cause a more drastic performance drop. This is an example of a HM which works to maximise the synergy between methods to support SA and short-term decision-making. The predictions provide information which allows anticipation of a degraded situation, while the simulation can support assessment of the impact of corrective actions, and the launching of action before the event has occurred, enhancing system resilience.

The HM can be evaluated in the light of a set of assessment criteria.

<u>*Objective 4:*</u> *To determine the criteria for evaluation of a hybrid simulation approach for short-term decision-support in healthcare.*

The review of the literature has shown that there is a need for short-term decision-support in emergency healthcare, and that a data-driven approach can support this, using a HM with three components: real-time data, predictive analytics, and a simulation model, triggered by forecasts to support system recovery. The review has also identified criteria for evaluating the HM, which is intended to be used recurrently. This requires an understanding of factors that contribute to its success or otherwise, including behavioural and organisational factors. Identifying these conditions supports future work in this area. From the literature review, the following criteria have been identified:

- As the HM is intended for recurrent use as a decision-tool integrated into ED operations, it should start with a system-level understanding into what matters in terms of patient experience, staff satisfaction, efficiency gains, or cost savings. Taking a QI approach supports the relevance of the study.
- The HM must support both task- and system-level understanding toward both adaptive short-term behaviours and escalation interventions, by supporting existing knowledge about what is happening, or is likely to happen. It is proposed that forecasts provide information which can support early adaptive behaviour to utilise spare capacity before queues build up. Where an escalation response is required, simulation can test a range of scenarios

according to the current need and allow the system to respond and mitigate the forecasted crowding, for example using demand management.

- As an integrated ED support tool, it is important that potential unexpected effects or uses are understood. Such tools need to be implemented with care, as unanticipated effects can result from the type or presentation of information, including technology-induced errors, or ambiguous information, which can actually reduce human decision quality and speed. A poorly designed tool can increase stress and workload, rather than reduce it.

- The application aims to support system resilience, by providing usable knowledge that supports anticipation about what to expect and information about how to respond. SA requires the perception of environmental information, the comprehension of its meaning, and a projection about the future based on this knowledge. This information must be comprehended by staff without interrupting workflow, hence usability, automation and integration of components within HM is required.

- Barriers to use of such an approach can exist at all levels of the organisation, and require understanding and managing. These include time and capacity, politics, resistance to change, and individual factors.

This literature review has identified that an integrated real-time HM approach in emergency healthcare can support SA and subsequent short-term decision-making. However in healthcare, this approach is in its infancy, and a range of barriers and challenges exist in practice.

The following gaps and contributions of this thesis have been identified:

- A framework for supporting the implementation of a short-term decision-support HM in sociotechnical systems is currently lacking in the literature. Positioning the framework in sociotechnical theory can support an understanding of the interacting roles of both social and technical elements in a complex system.

- An application of an integrated HM for short-term decision support has not been investigated in an applied setting, toward understanding its impact at the system level for both patients and staff. In healthcare, published studies are conceptual or prototyped. To address this gap, a HM with three components: real-time data, predictive analytics, and a simulation model, triggered by the predictions, will be developed and evaluated, supported by the framework.

- To understand the impact at the system level, the proposed benefits to both staff and patients needs to be understood. Real-time and forecasted information may be useful for patients, supporting attendance decisions using a demand management approach. However it is unknown whether this information changes health-seeking behaviour. Supporting patient decision-making is rarely considered, yet patients are stakeholders in the system under investigation. System-level evaluation can provide evaluation criteria to support future work applying similar interventions in similar domains.

The next chapter presents the philosophical assumptions and research methodology used to operationalise this research.

Chapter 3: Methodology

3.1 Introduction

This chapter outlines the philosophical approach: critical realism; the research strategy: Design Science; the research design: mixed-methods research; and the methods used to address the research questions. The mixed-methods design is a *partially mixed sequential equal status design* (Section 3.2), whereby the methods are not fully integrated. The approach is sequential, for example methods for defining the problem will precede the development and execution of the integrated hybrid model, while the final evaluation follows. In order to address the research questions, a range of methods are positioned within a Design Science research strategy (Section 3.4). The relationship between critical realism (Section 3.3) and Design Science research is outlined, and the steps involved in conducting design research are elucidated (Section 3.5). Research ethics are outlined in Section 3.7.

3.2 Research design: Mixed-methods

In the last twenty years, mixed-methods research has become a highly applied and debated topic of conversation (Given, 2017; Ghiara, 2020). Through collecting a stronger and richer array of evidence than single methods alone, mixed-methods approaches are considered particularly suited to addressing complex practical problems by exploiting the synergy between qualitative stakeholder engagement and quantitative outcomes to inform intervention planning, implementation, evaluation, and monitoring (Ivankova & Wingo, 2018). Mixed-methods research is "*the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g. use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration*." (Johnson, Onwuegbuzie & Turner, 2007, p123). The methods share the same research questions, collect complementary data, and conduct counterpart analysis (Yin, 2006). Intrinsic to this definition is the concept of triangulation, which is the convergence of findings through the use of multiple methods for validation, ensuring that explained variance is due to the underlying phenomena, rather than the individual method.

However the purpose of mixed-methods is not limited to triangulation, as data should be collected that will provide all of the information that is potentially relevant for the study (Johnson et al., 2007). Units of analysis may occur at more than one level (Yin, 2015): the system, the intermediate and the individual level, and multiple units of analysis may be involved at each level. In this thesis, the mixed-research methods are integrated within an overall Design Science research strategy (Section 3.4). Mixed-methods research has been identified as appropriate for Design Science research, which involves build-evaluate cycles, where evaluation, in particular, benefits from the strengths of a mixed-methods approach (Gregor & Jones, 2007; Ågerfalk, 2013).

The confinement of multiple methods to a single study forces the methods being used into an integrated design (Leech & Onwuegbuzie, 2009; Yin, 2017). Creswell and Plano-Clark (2018) are amongst a number of authors who have outlined a range of mixed-methods research designs. For example, consideration of the sequence of data collection, the relative priority to each paradigm, and the stage of the project in which the paradigms are implemented can all inform a design (Ågarfalk, 2013). However many of these typologies are complicated, or too simplistic, such that the important criteria for mixed-methods researchers are not captured. To maximise the benefits of a mixed-methods approach, the research design must reflect the conceptual, philosophical, and procedural congruence between the research question, the research design and the methods employed to make integration possible and justifiable. For this reason, Leech and Onwuegbuzie (2009) proposed a three-dimensional typology of mixed-methods designs, incorporating the level of mixing (partially mixed versus fully mixed); the time orientation (concurrent versus sequential), and the emphasis of approach (equal status versus dominant status). When undertaking a mixed-methods study, qualitative and quantitative methods are used at different stages. These may be conducted either concurrently or sequentially. The major difference between partially mixed-methods and fully mixed-methods is that whereas fully mixed-methods involve the mixing of quantitative and qualitative techniques within one or more stages of the research process or across these stages, with partially mixed-methods, both the quantitative and qualitative elements are conducted either concurrently or sequentially in their entirety before being mixed at the data interpretation stage (Leech & Onwuegbuzie, 2009).

According to this typology, the design of methods in this thesis is a *partially mixed sequential equal status design* (illustrated as P3 in Figure 4.1). There is a sequential component as the hybrid model (HM) will be informed by a patient questionnaire, and evaluated by staff interviews. Additionally, the qualitative and quantitative data sets are all analysed separately, and synthesis takes place at the data interpretation phase. This is necessary because the HM is quantitative, while defining the problem and evaluating the solution have qualitative components.

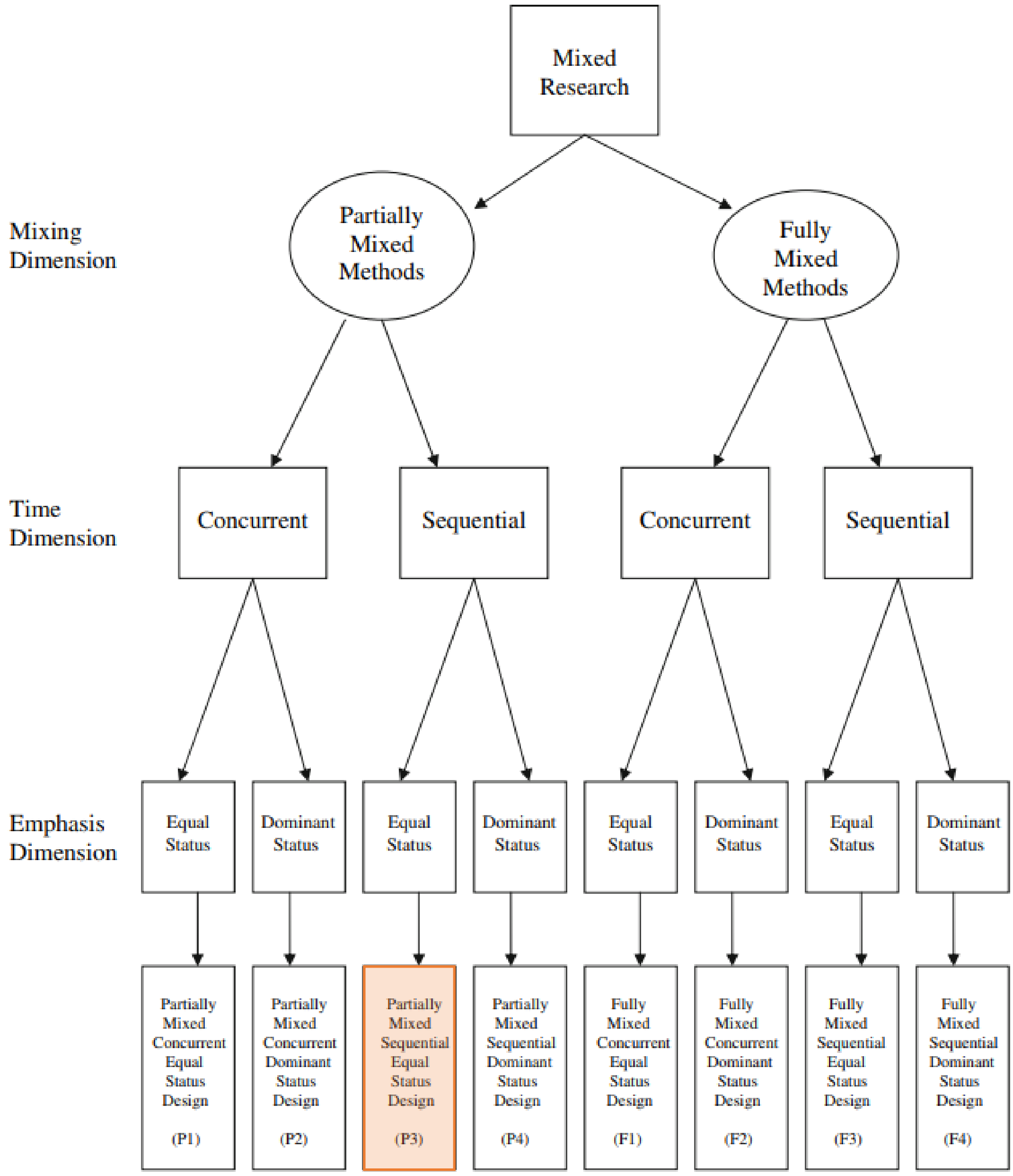


Figure 3-1 Typology of mixed methods research (Leech & Onwuegbuzie, 2009)

The philosophical approach, critical realism, is discussed in the next section, and the subsequent sections outline the Design Science justification and methodology, and the methods employed for data collection.

3.3 Philosophical Approach

3.3.1 *Introduction*

Discussions of *theory* in the subsequent sections are limited to the following definition, according to Abend (2008), who comprehensively defined multiple uses of the term in sociological literature. The main goal of a theory is to say something about empirical phenomena in the social world. This may shed new light on an empirical problem, help one understand some social process, or reveal what ‘really’ went on in a certain conjuncture. This is in contrast to a scientific theory, which has been tested and is widely accepted as valid, describes the causes of a particular natural phenomenon, and is used to explain and predict aspects of the phenomena.

The relationship between data and theory is much debated by philosophers of research, however a fundamental requirement of research is specifying the relationship between empirical material and theory, such that empirical data is used to test the strengths and weaknesses of a theoretical proposition or a conceptual framework (Sovacool et al., 2018). This is then revised to form new knowledge (Hancke, 2009). These theories or frameworks can be classified according to their underlying philosophical positions. For example, Easterby-Smith, Thorpe and Jackson (2015) outlined *positivism* and *social constructionism* as two opposing positions for conducting research. In their analysis, positivism holds an ontological assumption: that reality is external and objective; and an epistemological assumption: that knowledge becomes significant when based on observations of this reality. Theories in this paradigm are generally well suited to quantitative research methods such as experimental designs and analysis. Social constructionism, the opposing paradigm, developed as a reaction to positivism. It holds that reality is socially constructed and is given meaning by people. Social constructionism is an interpretive approach, which attempts to understand what people think and feel, why people have different experiences, and how these constructions and meanings drive action (Sovakool et al., 2018). Theories within this paradigm are suited to qualitative research methods, such as interviews and focus groups.

However few studies are pure examples, with many researchers choosing research designs and methods that blur the distinction between the two opposing approaches. While social constructionism is well supported by complementary qualitative methods, and positivism is associated with quantitative approaches, mixing methods can create philosophical difficulties when they represent very different ontologies (Easterby-Smith et al, 2015). However several philosophical positions provide coherent schools of thought, and partly reconcile these different perspectives to be consistent with both quantitative and qualitative research methods. These include Habermas' *Critical Theory* (1970, in: Brunkhorst et al., 2017), which views the motives and impacts of powerful groups and individuals on the behaviours and attitudes of the least powerful. This is relevant where knowledge may be determined by political processes. Giddens' (1984) *Structuration Theory* determines that structure and agency are not pre-ordained, but that structures are created through the agency and actions of individuals, and structure then guides and restrains individual agency. A continual interaction occurs between social structures and social action. This is relevant in management research for understanding the relationships between employees and their organisation, or between information systems which exist to facilitate action, and the resultant actions. A further school of thought is *Pragmatism*, which originated from the 19th Century work of John Dewey (1916, in: Dewey, 1998), in particular. Pragmatism claims that there are no pre-determined theories or frameworks shaping knowledge, but that meanings come from the lived experiences of individuals. This offers a synthesis between features often considered irreconcilable, such as positivism/anti-positivism. Finally, in the late 1970s Bhaskar's *Critical Realism* (Bhaskar, 2013; 2014), initiated a post-positivist perspective. He argued for a structuralist position, in that we can only understand the social world if we identify the structures at work that generate events, and a realist ontology, in common with the work of Habermas. This is in contrast to interpretative approaches, which can be less clear about the nature of reality, being relativistic in their ontological position. The implications of this for M&S research is that while the researcher is required to be sufficiently "epistemologically reflexive" and vigilant about investigations, models need to be answerable to empirical investigations.

3.3.2 *Research philosophy in OR*

In the field of OR, a positivist philosophy dominates (Holm et al., 2012), however the use of quantitative data need not imply the acceptance of a positivist epistemology. Mingers (2001) argued that quantitative data can - and should - be interpreted in the light of relevant social meanings. In social domains such as healthcare, isolating one part of the healthcare system from the rest can severely compromise the usefulness of the model in practice by shifting the problems elsewhere (Jahangirian et al., 2012). White (2009) argued that hybrid studies are a better approach for disorderly, complex processes. These approaches have been described as action research, reinforcing their contribution to practice as well as to theory (Howick & Ackerman, 2011). Similarly, Mingers (1997) asserted that to make the most effective contribution within rich social organisations, it is appropriate to combine methodologies and even paradigms. This has important implications for how knowledge is viewed. Combining paradigms presents challenges, and while Mingers & Rosenhead (2004) found through a survey that most researchers judged multi-methodological approaches to be more effective than single, they also found that relatively few combined hard and soft approaches. Howick and Ackerman (2011) reported in their review that the majority of papers avoided the multi-paradigm debate, which they conjectured may be to reduce effort, having decided that mixing methods is difficult enough without trying to reconcile potential incommensurabilty. Mingers and Brocklesby (1997) maintained that because of the uncertainties associated with any single paradigm, there is a need for conscious pluralism in OR research practice. 'Hard' methods assume that there is agreement on the nature of the problem-situation, while 'soft' methods assume that there are divergences of opinion. In addition, accounting for power and sociopolitical aspects is often necessary, and Brocklesby & Cummings (1996) argued that all of these perspectives can be complementary. While the issue of paradigm incommensurablity is still debated, the response within OR, where addressed, has tended to unite around Habermas' social theory (Jackson, 1985; Brocklesby & Cummings, 1996), or Bhaskar's critical realism (Mingers & Brocklesby, 1997; Mingers, 2015; Syed & Mingers, 2018). Both of these paradigms support a constructivist view of knowledge production with a realist ontology. These approaches make multi-methodology possible and sustainable, and allow for combining methods without having to constantly adjust the epistemological position, which can cause ‘stress and anxiety’ to the researcher (Kotiadis and Mingers, 2006).

3.3.3 *Critical Realism*

It is widely acknowledged that the most challenging current problems in management research are centred on people, rather than technical issues, hence Alvesson and Willmott (1996; 2012) stressed that the theory and practice of management requires a critical perspective to confront contemporary organisational challenges and problems. These approaches promote *epistemic reflexivity* to establish the necessary conditions for differentiating and establishing constructions of reality, and showing the possibility of alternative accounts. *Epistemic reflexivity* indicates reflection on the social conditions under which knowledge comes into being and gains credence (Bourdieu, 1990). Being epistemically reflexive requires seeking out new modes of engagement with research subjects that support the co-creation of knowledge, for example using participatory or qualitative approaches (Bryman & Bell, 2007). The critical approaches upheld by Giddens, Habermas and Bhaskar are in substantial agreement in many respects, for example each defends the possibility of objectively valid scientific knowledge, rejecting radical relativism, but also positivism, which is seen as depicting individuals as passive subjects of deterministic social systems (Pleasants, 2002). In other words, each embraces both epistemological relativism and ontological realism, and the latter necessitates the former. Similarly, M&S studies in social systems require a consensus of viewpoints toward understanding the structures, rules, processes, mechanisms etc. of social organisations. Any account of this ontology can be considered epistemologically contingent and fallible, as models are approximations of a real system relying on simplifications and abstractions, which is subjective because there can be neither a perfect nor accurate representation of a system (Padilla et al., 2013).

While Alvesson and Willmott (1996) emphasised that critical thinking does not exclude technical problem-solving approaches, they proposed Habermas’ social theory as a relevant approach to understanding and reflecting critically upon the dynamics of organisations and management practice. Habermas attempted to draw attention to the socio-cultural factors that influence experience, in particular communication. This approach makes much of the asymmetrical relations of power that inhibits the open formation and expression of views, and thus is well-suited to the application of soft OR/problem-structuring methods where it is prudent to be reflective upon the unequal intellectual and political positions of involved stakeholders. However social theory has been criticized for being too relative, despite ostensibly having a realist ontology, making it difficult to determine the implications of the research (Morrow & Brown, 1994; Pleasants, 2002). This is of particular relevance where, as in the case of this research, the implications are of a pragmatic and applied nature.

Gorski (2013) and Archer et al. (2013) advocated critical realism as a position that can supply a general schema for thinking about social behaviour. While there is not one unitary framework or set of beliefs that unite critical realism, Bhaskar's position described analyzing the world into discrete structures such as 'person' or 'network' and examining how interactions between these structures change their properties, the relationships, or lead to the emergence of new structures. As reality is complex, temporal and changing, these can be examined temporally, spatially or culturally, to determine how the actions of agents within the system have the power to change the system. The relevance of this is clear, as critical realism holds that structure both precedes and is an outcome of human agency. For this thesis, structural conditions, for example data visualisations and other information sources, aim to influence SA and human action, but may have unexpected side effects in practice. Additionally, while the interactions of patients and staff with real-time data applications can potentially change the behaviour of the system, it is clear that these changes may be difficult to determine statistically as many other factors are involved in these relationships. Ontology is central to critical realism, emphasising that many of the features of the world are not empirically verifiable or quantifiable, and may resist scrutiny. This approach allows combining methods, approaching causation critically, and using partial facts and events which account for the complexity and heterogeneity of the social world (Archer, 2016). As 'causal powers' or 'causal mechanisms' are dependent upon the nature of structures, it is contingent whether they are exercised at any particular time or place, for example, responses to forecasts of crowding will depend upon the context in which they operate: other conditions which constrain interventions, motivations to act, alternative knowledge sources or habit, perceived costs/benefits, for example.

Few OR researchers have directly addressed the philosophy of OR research in recent years, Mingers being a notable exception (2000; 2015; Syed & Mingers, 2018), although the relevance and importance of the epistemology of M&S is acknowledged (e.g. Tolk et al., 2013), Mingers (2000) argued that critical realism fits well with OR as an applied discipline, demonstrating examples of its fit with system dynamics as a modelling method, Soft-systems Methodology as an interpretivist method, and statistical modelling as an empirical method. For example, critical realism often rejects the possibility of prediction about social matters, as causality is held to have multiple causal mechanisms (Næss, 2015). However with application to time-series modelling, Mingers demonstrated the importance of contextual factors beyond the model itself when making interpretations of model outputs. While he noted that this is typical of practical, applied OR studies which do not become ‘trapped in the purely empirical domain of the data itself’, it is an example of the value of positioning OR research within a critical approach. Critical realism overtly combines explanatory theory with empiricism and confronts the radical anti-realism position of interpretivist/constructivist research philosophies, which allows research outputs to inform theory and real-world action. Additionally, within a Design Science approach critical realism enables a beneficial interaction between academics, domain experts, and end-users to support co-production across diverse stakeholder groups (Hodgkinson & Starkey, 2011).

3.4 Research Strategy: Design Science

3.4.1 *Introduction to Design Science*

Design Science research forms an appropriate research strategy for this thesis, which is interested in the practical relevance of real-time hybrid M&S to inform improvement in healthcare policy and practice, and to increase M&S research utilisation and relevance in healthcare.

Design Science studies progressive refinements of an intervention in its target settings, with the aim of developing practical knowledge. This works toward a more effective solution, with improved articulation of principles that underpin its impact, and aims to be applicable more widely (Van den Akker et al., 2006). While innovative M&S solutions to emerging problems continue to be published, the modeller's decisions in resulting designs are often implicit, yet more explicit learning can advance subsequent design efforts (Van den Akker et al., 2006; Richey & Klein, 2014). This is of particular importance in a tool such as real-time simulation, which is intended as a recurrent decision aid assimilated into routine practice, compared with single-use M&S applications.

3.4.2 *Justification for Design Science*

Yin (1994) defined *evaluation* as a particular type of research used to assess and explain the results of action projects or programs operated in a real-life setting (Yin, 1994; 2017). Case study research is a research strategy that supports mixed-methods approaches, and can be used as an evaluation tool. Case study research describes a research strategy which investigates human activity in the real world, where context is important such that precise boundaries are difficult to define between the context and the activity under investigation (Gerring, 2006). This allows the research to 'close-in' on a real-life situation (Flyvbjerg, 2004). Yin (2014) states that case study research is the preferred form of research where the main research questions are '*how*' or '*why*' questions; where the research has little or no control over behavioural events; and where the focus of the study is contemporary, i.e. here and now, rather than historical. Case studies are most commonly exploratory, and may generate hypotheses (Yin, 2017), however the case study method can serve evaluation needs directly by assessing outcomes and testing hypotheses (Yin, 1994). Although rigorous methodologies have been developed for implementing case study research (Stake, 1995; Eisenhardt & Graebner, 2007; Yin, 2014), the approach does not directly support the development of an artefact within its methodology, nor a focus on incrementally effective applicable problem solutions. This leaves problem generation and artefact/model development unspecified.

Similarly, Implementation Science provides a robust methodology which supports applied research to develop the critical evidence base informing the adoption of interventions by health systems (Allotey et al., 2008). Interventions have previously undergone sufficient scientific evaluation to be considered effective, while the implementation science methodology focuses on the design and evaluation of a set of activities to facilitate successful uptake of this evidence-based health intervention (Handley et al., 2016). This is an applied health science, focusing on the evaluation of the implementation of an artefact, but assumes an existing evidence base for the detailed intervention in its specific setting.

In contrast, Design Science explicitly integrates design as a major component of research. In line with the definitions provided by Collins, Joseph & Bielaczyc (2004), Nieveen and Folmer (2013) and Plomp & Nieveen (2013), the definition of Design Science used in this thesis is *the systematic analysis, design and evaluation of interventions with the dual aim of generating research-based solutions for complex problems in practice, and advancing our knowledge about the characteristics of these interventions and the processes of designing and developing them.* This supports applied research which also provides a set of design principles (Nieveen & Folmer, 2013):

- The purpose/function of the intervention
- The key characteristics of the intervention
- Guidelines for designing the intervention
- Its implementation conditions
- Theoretical and empirical arguments for the characteristics and procedural guidelines

These principles provide evidence for the potential impact of the intervention in its given domain, and how it might work in practice, in this case contributing to the knowledge base about the value of real-time hybrid modelling in healthcare for short-term decision support. In contrast, research strategies such as case studies or ethnography attempt to characterise events and relationships in real-life contexts, but there is no attempt to change practice (Yin, 2014). Meanwhile, experimental designs can analyse the effects of interventions, however controlling these research designs can distort real-life learning (Veal, 2006). In Design Science, the intervention is based on knowledge from the literature, such that evaluation of the approach in its applied setting contributes to expanding the knowledge base (Collins et al., 2004; Van den Akker et al., 2006). For future researchers, this provides information needed for applying similar interventions in similar domains. For policy makers, these principles assist in making research-based decisions for supporting complex problems in similar domains.

3.4.3 *Philosophical underpinnings of Design Science*

Donald Schön (1983), a philosopher of design, technological innovation and applied research, argued that practical, real-world problems rarely present themselves neatly enough for scientific generalisations to apply: the *rigour vs relevance* debate. In his argument, he considered that technical studies are often inapplicable to the 'swamps' of real-life practice. While Schön's work was founded on the work of Dewey (Waks, 2001), his epistemology of practice is centred on the practice of 'reflection-in-action' which is fundamental to critical realism. The focus is the critical evaluation of theories on the basis of empirical data. For Archer (2003; Archer et al., 2013), the focus on structure, context and causal mechanisms in critical realism is valuable in complex social systems, as it acknowledges that decisions and actions are contingent and uncertain, but that consensus toward an output is required in order to learn and progress in the real world. For M&S, it is acknowledging that stakeholders in applied research, a practical, problem-solving activity, are people. In work situations, people contend with varying workloads, abilities, stressors and distractors, and are rarely rational decision-makers who require simply a correct technical solution. How an intervention may or may not function in practice is of explicit interest, for example how people respond to specific features of the intervention. Design Science is not used to test theories, but to build interventions that are based on theories, and to test the effectiveness of the intervention in practice (Van den Akker et al., 2006). Faced with many ambiguities and unknowns, these results may be preliminary. However by incorporating diverse stakeholders, especially groups with conflicting agendas, the problem space, scope, and potential of new design solutions can be expanded (Hodgkinson & Starkey, 2011; 2012). The aim is to work toward improved designs by testing a use-case, where the findings can be generalised to similar cases. Carlsson (2006) argued that where design science is positioned in critical realism, the intention is to produce more detailed answers to the question of *why* and *how* an intervention works, for *whom*, and in *what* circumstances. Critical realism can be used to strengthen the theoretical foundations of a Design Science approach, by balancing the issues of structure and agency without over-prioritizing either (Hodgkinson & Starkey, 2012). This means attending to how and why an intervention has the potential to cause the desired change, which in this thesis will be investigated using a single use-case.

3.4.4 *Generalising from a use-case*

Many scholars have argued that it is not possible to generalise from a single case, and that case studies are only appropriate for the exploratory phase of an investigation. However these claims have been robustly countered, as all new knowledge enters into the collective process of knowledge accumulation (e.g. Flyvberg, 2006).

Case studies are a rigorous, in-depth methodology (Yin, 2014) which investigates one or more units of human activity in context, allowing a close-up investigation of a real-life situation, using a range of methods to triangulate findings. Use-cases differ from case studies by helping to understand how technology and related solutions can be applied to solve real-world problems. Technical descriptions can be too abstract to explain how the technology is useful in practice, and use-cases are considered a good basis for testing. Use-cases allow a description of the sequences of events that, taken together, lead to a system doing something useful (Bittner, 2002). By focusing on how an intervention is used in practice, the approach is significantly narrower than a case study, in which precise boundaries are difficult to define, and the focus is on understanding reality (Yin, 2014). Use-cases ensure the intervention is used in practice as intended.

Questions arise as to how this knowledge can be created in a cumulative way that can be generalised beyond individual solutions to individual problems. The problem arises from the requirement to create designs that are relevant to practice but at the same time contribute to the knowledge base. Offermann, Blom and Bub (2011) suggest that generalisability or transferability of findings occurs where settings are similar, especially when research involves social dimensions, and insights might be transferred from one to the other. They outline three types of design according to a range of scope. Short-range designs are aimed at a specific setting; mid-range designs are aimed at a specific *type* of setting; and long-range design are general insights about a type of design approach. This thesis focuses on mid-range design, by creating and validating (through evaluation) the utility of a short-range design, with the aim of increasing its generalisability. The mid-range design proposes that the application of the intervention within a certain scope of situations will yield a certain utility. This mode comes closest to cumulative knowledge creation, and the notion of "generalisation" in quantitative science (Offermann et al., 2011). To increase the robustness of the intervention, the more situations a design has been shown to work, the more likely it is considered to work for similar new problems. This presents a limitation of this formative research, as the intervention is tested in one use-case.

However this thesis proposes the development of a hybrid framework which is generic and testable in multiple, similar settings. Similarly, Zhang et al., (2013) proposed a generic Design Science framework to improve the integration of sustainable development between strategic, tactical and operational levels. Their framework aims to rank local and global environmental tools, supporting each activity, and stakeholder collaboration, along the design process. The framework proposed in this research supports the development of real-time HMs for short-term decision-making in sociotechnical systems.

3.5 Design Science Research Methodology

While the previous section characterised the focus of use-cases as narrower than that of a case study design, Collins et al. (2004) outlined a comprehensive methodological approach to Design Science research which employs use-cases. Design Science involves the building and evaluation of artefacts designed to meet identified business needs (Hevner et al., 2004). The methodology described by Collins et al. (2004) covered in-depth evaluation and analysis, and an iterative approach to implementing design research, requiring teams of cross-disciplinary researchers. While these features characterise Design Science in a large-scale application, several authors have provided definitions of stages and stakeholders involved in Design Science research. For example, a review by Ostrowski, Helfert & Hossain (2011) found that most researchers included some component in the initial stages of research to define a research problem, and common agreement on the outcome: an artefact or model. A process in the middle entails construction of the artefact, and this step requires relevant literature, existing theories/knowledge, and collaboration with partners. Additionally, Hevner et al. (2004) emphasised that such an approach requires rigorous evaluation of the utility, quality and efficacy of the proposed solution.

Peffers et al. (2007) synthesised the design, engineering, and Information Systems (IS) literature, and produced the following Design Science process model with six ‘Activities’: (i) Identify problem and motivation; (ii) Define objective of a solution; (iii) Design and Development; (iv) Demonstration; (v) Evaluation; (vi) Communication. ‘Activities’ are defined as the tools, methods, and/or actions taken by researchers to gain sufficient knowledge in order to create/produce/develop an artefact (Ostrowski et al., 2011). The process outlined by Peffers et al. (2007) forms a complete methodology for structuring and conducting Design Science research (Figure 3.2). While there is a process sequence, the research may start at a number of stages, depending upon the research objectives, and the process may be iterative.

A Design Science research methodology specified by Blessing and Chakrabarti (2009) can be mapped to these stages, and is detailed in the next chapter. In their comprehensive methodology, individual research projects may focus on one or two stages only, although iterations between stages will take place. For example, Salehi & McMahon (2009) focussed on the first stage of the framework outlined by Blessing (1994), and Blessing and Chakrabarti (2009) to explain the relevant criteria and results of the Descriptive Stage (Phase 1) in detail, while Zhang et al. (2013) developed a framework for stakeholder engagement which spans the duration of the Design Science process, but focuses only on how to support collaboration in tool integration across organisations.

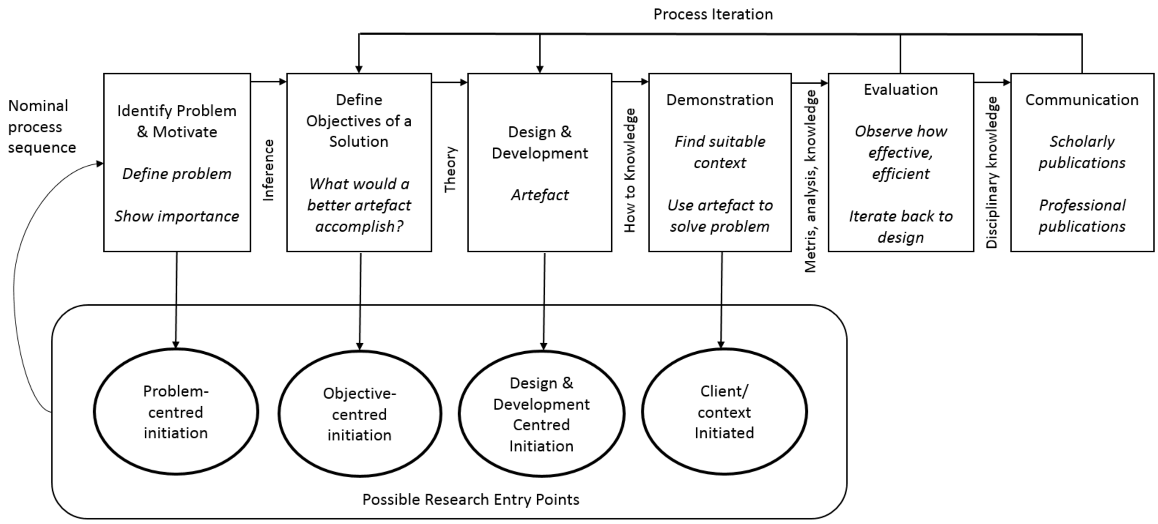


Figure 3-2 Design Science Research Methodology (reproduced from Peffers et al., 2007)

According to Blessing and Chakrabarti (2009), Design Science research must address the following issues:

- What are the criteria for evaluating the intervention? This involves identifying the key criterion at which the intervention is aimed, and aligns with the first and second stages of the methodology outlined by Peffers et al. (2007) (Figure 3.2), which defines the specific research problem and justifies the value of the solution [Activity (i)] and then infers the objectives of a solution from the problem definition and knowledge of what is possible and feasible [Activity (ii)].
- How is an intervention created? This involves identifying the influences on evaluation criteria, how these influences interact, and how they can be measured i.e. how to improve the design process. This aligns with Activity (iii) in Peffers et al. (2007) which creates the artefact, defined as *"any designed object in which a research contribution is embedded in the design",* and involves determining its functionality, its architecture, and developing the actual artefact based on knowledge of theory and other information sources.
- How do we improve the chances of developing a plausible intervention? This involves understanding how the knowledge gained from the design process can be used to develop guidelines, methods and tools, and how this design support can be evaluated. Evaluation is needed to determine whether the application can contribute to a plausible intervention as determined by the criteria. Peffers et al. (2007) (Figure 3.2) encompass this principle in Activities (iv), (v) and (vi), which involve demonstration of the use of the artefact toward solving the problem, and evaluation of how well the artefact supports a solution to the problem, through an appropriate method. Finally, Activity (vi) involves communication of the problem and its importance, the artefact and its utility, novelty, rigor of design, and effectiveness of its approach, to appropriate audiences in practice and academia. In academia, this might support progressing the design process, applying it to another research domain, or using it to solve a different problem.

These principles support a rigorous process for designing artefacts to solve observed problems, to make research contributions, to evaluate the designs and to communicate the results to appropriate audiences (Hevner et al., 2004; Peffers et al., 2007). A research methodology provides practice rules to implement these principles (Peffers et al., 2007), and in Design Science these include the development of an artefact, which may be a construct, model or method (Hevner et al., 2004). This should be a process that draws from existing theories and knowledge to develop a solution to a defined problem. The stages in the research methodologies outlined by Peffers et al. (2007) and Blessing and Chakrabarti (2009) will be examined in more detail in the next chapter toward the development of a framework to support the application of real-time hybrid modelling studies for short-term decision-support, which enables the use of a range of methods.

3.6 Methods

The term 'method' refers to a systematic procedure for carrying out an activity, specifically (i) How knowledge should be acquired; (ii) The form in which knowledge should be stated; and (iii) How the truth or falsity of knowledge should be evaluated (Polgar & Thomas, 1991). In this research, the mixed-methods approach is part of a Design Science research methodology (McKenney & Reeves, 2018), with each of the three research questions directing the analysis sequentially. The research will involve the development of a HM framework which will be tested in a single case, in the emergency department at NHS Trust in the South-West of England, and its surrounding urgent care network. The rationale for the need for short-term decision support in ED has been explained in Chapters 1 and 2, making ED a suitable case for this study. The methods and use-case are outlined in detail in Chapter 5.

To meet the objectives of the research questions, the methods chosen in this research to execute the framework are indicated in Table 3.1.

Table 3-1 Research questions, aims and objectives and methods

Research Questions	Aim	Objectives	Method
1. *How can simulation approaches support short-term operational decision-making in healthcare?*	To determine the need for short-term decision-support in healthcare, and to examine how simulation, real-time simulation, and hybrid modelling approaches have been used for short-term operational decision-support in the healthcare context, in particular emergency care.	**1.** To explore the need for short-term decision-support in healthcare, in particular emergency care	**1**. Literature Review (Chapter 2)
		2. To explore how analytics methods can be used for short-term decision-support.	**2**. Literature Review (Chapter 2)
		3. To critically evaluate simulation approaches used in healthcare for decision-support and to identify how simulation is used for short-term decision-support	**3**. Literature Review (Chapter 2)

		4. To determine criteria for evaluation of a hybrid simulation approach for short-term decision-support in healthcare.	**4.** Literature Review (Chapter 2)
2. *How can an integrated hybrid approach using real-time simulation and data analytics support short-term operational decision-making?*	To test and evaluate the potential of an integrated hybrid approach for short-term decision-support in healthcare combining real-time simulation with other analytics approaches.	**1.** To propose a generic framework supporting an integrated hybrid approach for short-term decision-making in healthcare	**1**. Literature Review and Design Science methodology (Chapters 2, 3 and 4)
		2. To apply the framework within a case study in a hospital ED	**2.**Direct observation, patient questionnaires, secondary data analysis, time-series forecasting, real-time simulation (Chapters 5 and 6)
		3. To evaluate the application of the framework.	**3.** Semi-structured staff interviews (Chapter 7)
3. What are the implications and the added value to the system of using real-time data applications for both patient and NHS decision-support?	To analyse the system level impact of the use of real-time data by both patients and staff decision support.	**1.** To critically evaluate the value that real-time applications provide at the system level.	**1.** Patient questionnaires and Semi-structured staff interviews (Chapters 5 and 7)
		2. To synthesise previous findings and evaluate the framework in light of the application.	**2.** Synthesis of findings (Chapters 7 and 8)

RQ1 has been addressed in the previous chapter. The next chapter develops the framework and the final objectives of RQ2 are addressed in the subsequent chapters.

3.7 Research Ethics

Specific ethical considerations arise when conducting design research in healthcare, in particular where human subjects are involved. In this research, this has been addressed in the following ways:

(i) The project was reviewed by the University of Exeter Business School Ethics Committee and given a favourable opinion (Appendix 1).

(ii) Hospital honorary contracts were obtained allowing staff-level access to departments, meetings and staff at the NHS Trust involved. A further honorary contract was required from a second large NHS Trust in the South-West of England to conduct patient questionnaires.

(iii) Honorary contracts allowed access to anonymised and pseudonymised secondary operational data.

(iv) All secondary hospital data used in this research is either fully anonymous (no identifiers), or has undergone pseudonymisation, which means that any identifying data has been replaced by one or more artificial identifiers.

(v) Patient questionnaires required ED department level approval. Signed, informed consent forms were gained from all patient participants.

(vi) Staff interviews required signed, informed consent prior to scheduling interviews.

(vii) Inclusion and exclusion criteria explicitly considered the potential risks to participants prior to selection, the avoidance of harm, and privacy and confidentiality.

(viii) All data used in this study is treated as confidential, and is stored in locked storage, and/or a password and full-volume encrypted computer.

3.8 Chapter Summary

This chapter has outlined the research strategy: Design Science research; the research design: mixed-methods research; and the research philosophy, critical realism, and why it is important and relevant for OR real-world research, in particular for implementing a real-time decision-support tool in a healthcare setting. It has drawn attention to the synergy between critical realism and design research, and outlined the steps involved in conducting design research. Research ethical concerns have been addressed. The next chapter reports the development of the proposed framework for implementation of the hybrid model. Its development is based on the literature review in Chapter 2, and the Design Science methodology which has been examined in this chapter.

Chapter 4: A generic framework supporting an integrated hybrid model for real-time decision making in healthcare

4.1 Introduction

This chapter proposes a framework for the development and application of real-time hybrid modelling (HM) studies in healthcare. The conclusion reached at the end of Chapter 2 is the recognition that a real-time decision-support system which combines real-time data, predictions, and simulation, has the potential to support short-term ED decision-making, however such an approach has to date not been evaluated in an applied setting. Chapter 3 identified Design Science as an appropriate methodology for investigating this. Additionally, a HM framework which supports the implementation of such an approach is currently lacking in the literature. Such a framework is motivated by the long-term focus on lack of evidence of real-world implementation and evaluation of simulation model results in the healthcare domain (Long et al., 2019), and the increasing need and opportunity to use real-time data to support quick and effective decision-making (Bumblauskas et al., 2017). Concepts derived from the Human Factors literature take account of sociotechnical system precursors of decision-making, including individual and team-level situation awareness (SA), and Quality Improvement (QI) theory is suggested as a means to bring together, in a generic framework, the concepts from data analytics, simulation and sociotechnical theory toward supporting short-term decision-making.

The development of the framework is done in two ways. The first development is through an examination of the stages of a Design Science methodology (Blessing and Charkrabari, 2009). The second is derived from insights from the literature review in Chapter 2, examining decision-making in dynamic, sociotechnical systems; data analytics and simulation for decision-support; and real-time simulation as a HM.

The framework is developed to be generic, as it is tested in practice, with transferable knowledge aimed at supporting similar future work in similar domains. It explicitly addresses the first objective of the second research question, which will be tested in subsequent chapters. RQ2 asks how an integrated hybrid approach using real-time simulation and data analytics (DA) can support short-term decision-making in healthcare, and its first objective proposes the development of the framework outlined in this chapter. The next section considers existing hybrid frameworks.

4.2 Existing hybrid frameworks

A conceptual framework should summarise the key factors and concepts of a subject matter, identify relationships between them and form definitions (Miles, Huberman, & Saldana, 2014). For multi-method approaches, Mingers (2001) argued that theoretical frameworks which can provide step-by-step guidelines will assist modellers in the development of hybrid models. Frameworks supporting hybrid simulation (HS) approaches have become popular, in particular SD and DES, due to the specific challenges and advantages of combining these different simulation approaches, namely the differing philosophies between the two modelling approaches, differing levels of decision-support, and the challenge of data exchange between continuous and discrete state changes. Some aspects of these are relevant to HM approaches.

Helal et al. (2007) introduced a methodology to integrate and synchronise HS based on a modular concept, where the modelled system is decomposed into several smaller modules, however their framework does not include conceptual modelling and addressed mainly technical integration. Similarly, the HS framework advanced by Alvanchi et al. (2011) was focused on technical interoperability. Motivated by the lack of a generic conceptual modelling framework, Chahal & Eldabi (2008) and Chalal, Eldabi & Young (2013) developed a HS framework for healthcare which focussed on the conceptual stage for SD/DES model hybridisation. The key elements of their framework are relevant for HM: problem identification, the mapping points between the models, and the mode of interaction, that is, how the models exchange information. Zulkepli and Eldabi (2015) argued that clearer guidelines on decomposition of the main objective into sub-objectives was required to understand the nature of communication between the hybridised models at the conceptual stage. They proposed a 3-phase framework that adds selection and communication elements as part of a series of guiding steps for developing the hybrid models. The three phases of their framework are the conceptual phase; the modelling phase; and the model communication phase, and it uses a modularisation approach similar to that of Helal et al. (2007). The authors argued that the conceptual modelling phase is the most important, as it supports modellers to think about some important issues before starting the hybridisation process. This includes addressing questions such as how both models could be linked using different packages, how to change the information, and how such information will affect the final result from both models. Each of these are relevant to HM.

Morgan, Howick & Belton (2017) took a broader stance, providing an overarching framework that examines the literature for 'all forms of mixing methods' (Howick & Ackerman, 2011). They used mixed-method designs to develop a conceptual framework for mixing OR methods. Although they applied it to HS using ABS, SD and DES, the principles are more broadly applicable. The features are: (i) The system modelling view, that is, whether a single view or multiple view is required; (ii) Method dominance – where information needs to be exchanged between methods, the direction of interaction, and the form of interaction, for example hard or soft methods; (iii) The mixed-method design including the number of methods, the number of points of information, the frequency and triggering of interaction, the separable roles of the methods, and the result of the mix in terms of the number of models and the modelling environments; and (iv) the technical justification of the mix. They analysed a range of mixed-methods designs from the literature (isolationism, parallel, sequential, enrichment, interaction, integration), and how they aligned with their identified features. The purpose of this mapping was to enable modellers to identify the design aligned with their perception of the problem and system. From their mapping, the features of the *sequential design* best captures the conceptualisation needed to support a HM approach using real-time data, prediction and simulation for short-term decision-support. Here, methods operate within their own paradigm, and one method follows the other, however they may be coded to support interaction. This is illustrated in Table 4.1, with the sequential design shown in the columns and the features which characterise the design indicated in the rows. The red text in Table 4.1 is an extension of the framework proposed by Morgan et al. (2017), who indicate that sequential designs do not require a trigger. Arguably, in the case of a recurrent use tool using real-time data, the use of a trigger maintains the sequential nature of the design. Moreover, their paper emphasises that sequential designs have a single interaction over a given time window. However, it is argued here that a sequential design used recurrently with an automated trigger may cycle through the sequence multiple times in a given time window. The number of methods, frequency of interaction, number of points of interaction, and type/frequency of triggers will be determined by the specific application. While this may share features in common with an *interaction design*, Morgan et al. (2017) note that interaction designs can interact in both directions; however where the interaction is one direction only, then the design is *sequential*, with interactive elements (the frequency of interaction and the type of interaction). How these elements will interact in the framework will be investigated in subsequent sections.

Table 4-1 Mixed methods approaches and categorisations, adapted from Morgan et al. (2017)

Features	**Mixed Method Design**	**Sequential Design**
What	View of the system	Need to capture different parts/behaviours of the same system
Why	Justification of mix	Allows for emergent insights as knowledge of system improves
	Number of methods under consideration	1 or more
	Number of separable roles of each method under consideration	Single problem theme with issues separable into more than model
How	Interaction likely	Yes
	Direction of interaction	One direction
	Form of interaction	Model insight and hard data
	Frequency of interaction over time window	Single pass of information or data from one model to the next **but may be cyclical**
	Number of points of interactions	Single to multiple
	Triggered or regular interactions	**Triggered by the state of the system, and/or regular triggers, every X timestamps**
	Number of models created	More than 1
	Modelling environment implications	May be built in single or two modelling environments

The toolkit developed by Morgan et al. (2017) supports the modeller to consider the input(s), the process and the output(s) of the project which all contribute to the selection of a mixed-method approach for hybrid simulation (HS). However HS studies can be distinguished from HM studies, which use simulation in conjunction with a range of other OR and cross-disciplinary methods. In the context of HM, the application of simulation with OR and cross-disciplinary methods is relevant not only in the model development/implementation stage of an M&S study (as is commonly the case with HS), but can be applied to other stages in the lifecycle, for example, conceptual modelling, input and output data analysis, model verification and validation, model formalisation, scenario development and experimentation, engaging with the stakeholders in the implementation of the results of a simulation study, and model documentation (Mustafee & Powell, 2018).

As problem definition is a key component of M&S study lifecycles (Robinson, 2004) and of the frameworks discussed above, it is appropriate to start the framework with a stage that aims to develop an understanding of the problem situation and determining the model objectives, with a view to using a real-time HM approach. As the purpose of the HM is short-term decision-support, an integral element of the approach will be to support SA to facilitate subsequent decision-making. Models of SA will be investigated in more detail in Section 1.4. However a closer look at Design Science methodology and how it might support the development of a real-time decision-support tool will be undertaken in the next section.

Key implications for a real-time Hybrid Modelling framework

A conceptual framework should summarise the key factors and concepts required by the HM, and identify relationships between them. This can be guided by a mixed-method research design which provides an overarching framework to inform the appropriate methodology. From the toolkit provided by Morgan et al. (2017), a sequential design with interactive elements provides a basis to consider the input(s), the process and the output(s) of the project. The interactive elements are necessary for a tool which is designed to be embedded into an organisational system. However it is important to distinguish HS studies from HM studies. HM studies use simulation in conjunction with a range of other OR and cross-disciplinary methods, at any stage of the M&S lifecycle, while HS combines methods at the model development/ coding stage. In HM, the simulation model forms a single component of the framework.

4.3 Stages of a Design Science research methodology

A set of conceptual principles defining what is meant by Design Science research were outlined in the previous chapter. These principles support a rigorous process for designing artefacts to solve observed problems, to make research contributions, to evaluate the designs and to communicate the results to appropriate audiences (Hevner et al., 2004; Peffers et al., 2007). A research methodology provides practice rules to implement these principles (Peffers et al., 2007), and in Design Science these include the development of an artefact, defined as "any designed object in which a research contribution is embedded in the design" (Peffers et al., 2007), which may be a construct, model or method (Hevner et al., 2004). This separates it from other research strategies, such as case study research and implementation science, which can effectively evaluate an intervention, but assume the pre-existence of an artefact. This should be a process that draws from existing theories and knowledge to develop a solution to a defined problem. The research methodology outlined by Peffers et al. (2007) is illustrated in Figure 4.1. These stages will be considered alongside the methodology advanced by Blessing and Chakrabarti (2009) in the subsequent sections with application to this study.

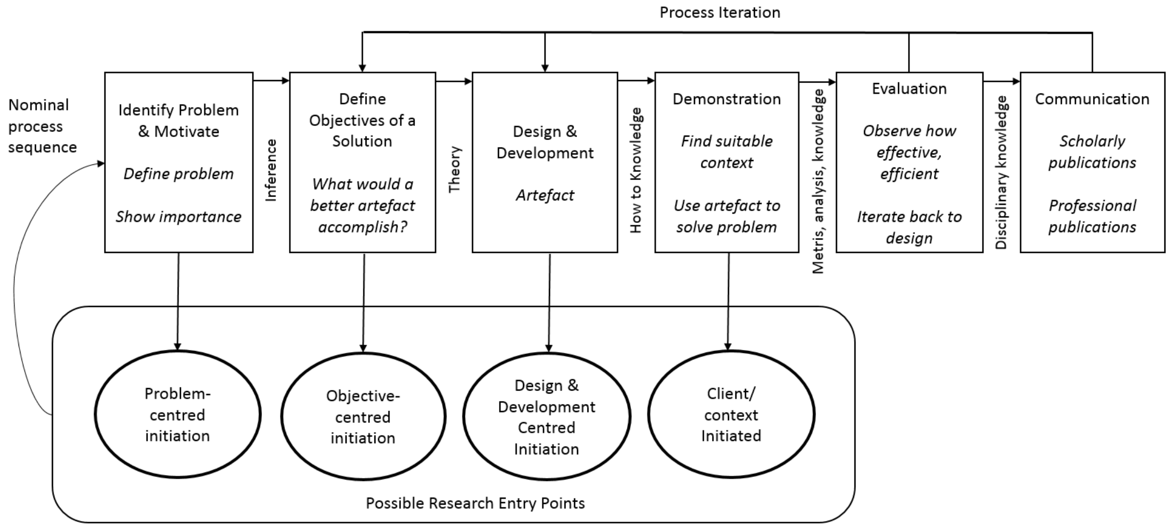


Figure 4-1 Design Science Research Methodology (reproduced from Peffers et al., 2007)

The research methodology developed by Blessing & Chakrabarti (2009) is illustrated in Figure 4.2.

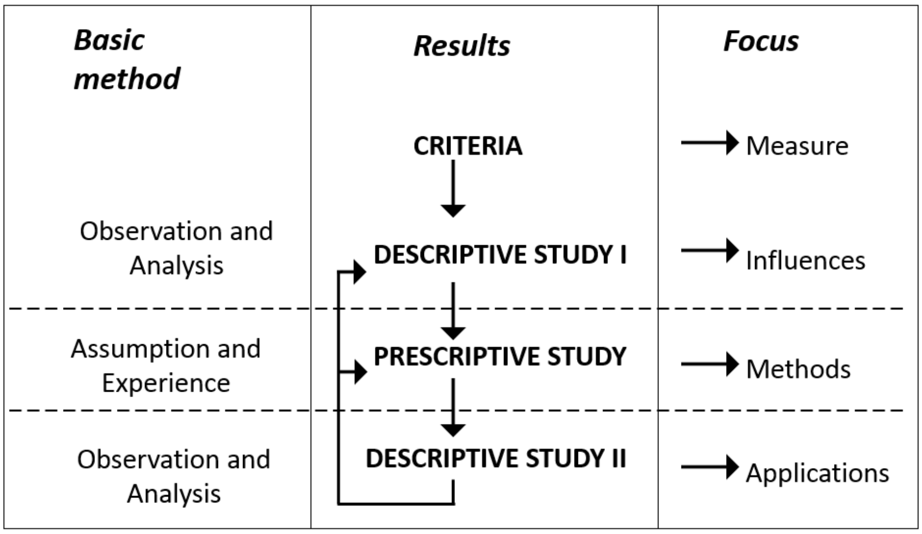


Figure 4-2 Design Research Methodology Framework, reproduced from Blessing and Chakrabarti (2009)

Similar to the methodology of Peffers et al. (2007) in Figure 4.1, the process in Figure 4.2 is not considered to be sequential, but iterative, and some stages may run in parallel. Additionally, entry may be at different points in the process depending upon the study aims, and a study may focus on one or two stages only. The purpose of the methodology is to achieve more rigour in Design Science, which the authors argue will improve the transfer of results into practice.

The individual components of the methodology will now be discussed [(a) to (d)]:

(a) <u>Criteria Definition</u>

A project should start with a clarification of the research by reviewing the literature, to determine the aim, focus and scope of the research project, and how the findings can be used to improve design (Blessing & Chakrabarti, 2009). As Design Science research aims ultimately at improving a situation, it is essential to determine criteria for evaluation. This involves Identifying the key criteria at which the intervention is aimed, and aligns with the first activity of the methodology outlined by Peffers et al. (2007), which defines the specific research problem and justifies the value of the solution. It is then possible to determine the factors that have a negative or positive influence on a plausible solution. The main tasks of these criteria are (i) the identification of the goals and purpose of the research; (ii) to focus the *Descriptive Stage I* on finding the factors that contribute to success; and (iii) to enable evaluation of the developed artefact in *Descriptive Stage II*. Criteria can be quantitative and qualitative. Quantitative criteria are essential, for example assuring data quality, and model validation.

However qualitative data is also important, as a technically ‘correct’ model may still fail to inform or be integrated into practice. Here, factors such as usability, confidence in the data and model, and perceived usefulness of the overall approach are important.

While the overall criterion for evaluation of a recurrent-use short-term decision-support tool is implementation through assimilation of the application in its real-world environment, this may not be possible, for example, where an initial design is proposed. Implementation assumes that the stakeholders have sufficient confidence in the application to successfully support short-term decision-making, and resultant action to improve system functioning. Many factors will influence confidence, and in the time-frame and limitations of the research, this may not be possible.

The criteria for evaluation are therefore factors that influence implementation, and are derived from the literature review (Chapter 2). These factors are:

(i) The usefulness, safety, efficacy, and cost-effectiveness of the application (Sections 2.3.1; 2.5.2; 2.6.1; 2.7.2).

(ii) Perceptions of the usability and functionality of the model (Sections 2.4.2; 2.6.3)

(iii) Confidence in the real-time data, the predictions and the simulation to provide short-term decision support, including its reliability and accuracy (Section 2.6.3)

(iv) The degree to which the model is considered to impact on SA in practice (Section 2.4.2)

(v) The degree to which the model fits into staff workflow (Section 2.4.2)

(vi) The capacity and technology-readiness of the organisation to innovate, the wider sociocultural context and how to sustain the application following assimilation (Sections 2.3.2; 2.4.2; 2.5.1)

(b) <u>Descriptive Phase I</u>

Having identified the criteria for success, an understanding of the various factors that influence, directly or indirectly, the above criteria is required. This focuses the modelling process and its evaluation on factors which contribute to success, and aligns with the second activity of the methodology outlined by Peffers et al. (2007), which infers the objectives of a solution from the problem definition, and knowledge of what is possible and feasible. These are derived from the literature as theoretical propositions (Carlsson, 2006), from site visits, direct observation (McKenney & Reeves, 2018), workshops (Blessing & Chakrabarti, 2009), and other methods such as interviews or questionnaires (Salehi & McMahon, 2009). User involvement takes place in the first and second phases [(a) and (b)], using methods appropriate for eliciting the required information to develop and evaluate the artefact.

Influencing factors are considered to be inter-related, creating a network of causes and effects connecting influencing factors with evaluation criteria. The literature review in Chapter 2 identified the following factors (Section 2.8):

- As the HM is intended for recurrent use as a decision-tool integrated into ED operations, it should start with a system-level understanding into what matters in terms of patient experience, staff satisfaction, efficiency gains, or cost savings. Taking a QI approach enhances the relevance of the study (Section 2.3.1).
- The HM must support both task- and system-level understanding towards both adaptive short-term behaviours and escalation interventions, by augmenting existing knowledge about what is happening, or is likely to happen. Predictions can provide information which can support early adaptive behaviour to utilise spare capacity before queues build up. Where an escalation response is required, simulation can test a range of scenarios according to the current need and allow the system to respond and mitigate the forecasted crowding, for example using demand management (Sections 2.3.4 and 2.3.5).
- As an integrated support tool, it is important that potential unexpected effects or uses are understood. Such tools need to be developed with care, as unanticipated effects can result from the type or presentation of information, including technology-induced errors, or ambiguous information, which can actually reduce human decision quality and speed. A poorly designed tool can increase stress and workload, rather than reduce it (Section 2.4.2).
- The application aims to enhance system resilience, by providing usable knowledge that supports anticipation about what to expect and information regarding what to do about it. SA requires the perception of environmental information, the comprehension of its meaning, and a projection about the

future based on this knowledge. This information must be comprehended by staff without interrupting workflow, hence usability, automation and integration of components within the HM is required (Section 2.4.2).

- Barriers to use of such an approach can exist at all levels of the organisation, and require understanding and managing. These include time and capacity, politics, resistance to change, and individual factors (Section 2.6.2).

Further data or information about the specific problem situation can be gained using appropriate methods for data collection, organisation, and analysis.

(c) <u>Prescriptive Phase</u>

The outcome of the descriptive study is used to develop the model toward the desired situation, and a conceptual framework can support this process. The *prescriptive phase* (Blessing & Chakrabarti, 2009) aligns with the third activity, 'Design and Development' in the methodology proposed by Peffers et al. (2007). This includes determining the functionality and architecture, as well as building the actual artefact, which may be a construct, model or method (Hevner et al., 2004). In order to assess the artefact, it needs to be developed, usually through prototyping (McKenney & Reeves, 2018). A single study may focus on one or more parts of this process, or one or more iterations (Blessing & Chakrabarti, 2009).

For building the model, a number of assumptions will be required, which must be made explicit so that the reasoning process can be traced (Robinson, 2006; 2013). The experience of the modeller contributes toward the resultant model (Keys, 2006), and must be considered at all stages. The validation of data, and each part of the model is performed during the model build process (Balci, 1989).

(d) <u>Descriptive Phase II</u>

The second descriptive phase is a formal evaluation, undertaken to determine whether the model has the expected effect on influencing factors identified in *Descriptive Stage I*, and whether these factors contribute to success. Evaluation provides feedback for further development. The success or otherwise of a modelling study may require sufficient descriptive narrative of the research process to enable conclusions to be made regarding the conditions under which the model was successful. Unexpected side effects may also occur. A reflective understanding of its limitations and how it is being used can ultimately increase the level of trust and confidence toward successful implementation. The fourth and fifth activities in the methodology described by Peffers et al. (2007) can be mapped to *Descriptive Stage II.* The fourth activity is ‘Demonstration’, where the artefact is demonstrated toward solving one or more instances of a problem. The authors suggest that this may be performed in situ, or may be through experimentation, simulation, case study, proof-of-concept etc. The fifth activity in their methodology is ‘Evaluation’, which Peffers et al. (2007) define as observing and measuring how well the artefact supports a solution to the problem, using quantitative and/or qualitative analysis as appropriate. At the end of this activity the researchers can decide whether to iterate back to the previous activity to try to improve the effectiveness of the artefact or to continue on to communication and leave further improvement to subsequent projects.

The purpose of *Descriptive Stage II* is to evaluate the functionality of the model from the user perspective (Blessing & Chakrabarti, 2009): is it useful in context? Does it address the need it was built to address? Are there any unexpected effects? Finally, it also assesses the criteria for evaluation, in this instance, what are its effects on SA? Is there confidence in each aspect of the model? Failures of evaluation (e.g. due to time or organisational constraints) can be reasoned, and can contribute to suggestions for improvement of the approach toward future applications (Blessing & Chakrabarti, 2009). As a proposed solution to a perceived need, based on a set of assumptions linking the solution to expected benefits, many factors may influence success or failure. For example politics, preferences, beliefs, motivations will all impact. As the HM intervention proposes to be assimilated into routine operational practice, use resources, and requires a certain technical infrastructure, these should be addressed in the evaluation.

The stages outlined above, advanced by Blessing and Chakrabarti (2007) and aligning with the work of Peffers et al. (2007), will inform the stages of the framework to support a HM application. A study using HM requires a conceptual framework to consider the constituent stages of a conventional M&S study and to explore complementary techniques (Mustafee & Powell, 2018). However, as the HM is developed for use as a recurrent decision-support tool, integrated into workflow, elements of the study process assume a different significance compared with single-use simulation models. While problem definition is an important stage in the process of all M&S studies, Design Science emphasises the importance of evaluation to support iterations, improvements and similar future work, to ensure that the modelling process starts with the assumption that the model will be useful in practice. For this reason, the problem definition stage must also consider the criteria and influencing factors for evaluation, for example, it should start with a system-level understanding of what matters in practice, to reduce the risk of unintended consequences in a complex system. The design of the model needs to be considered, such that it is not just useful, but useable in practice. It should support situation awareness (SA), which is an important component of short-term decision-making, and it should consider barriers to implementation of such an approach early in the design process. SA is common to all short-term decision processes, and will be examined in the next section to determine its conceptual role in a HM study framework.

Key implications for a real-time Hybrid Modelling framework

From the perspective of Design Science, the real-time hybrid model aims at improving a situation by being embedded into organisational workflow. A specific research problem must be determined, and the potential value of the solution justified by establishing criteria for evaluation. This requires problem definition, model development, and evaluation stages, which should be cohesive, and should consider more than the accuracy of the model. Evaluation should include the contribution of the decision-support tool to supporting SA and consideration of barriers and unintended consequences in practice.

4.4 Defining the problem and the objectives

The motivation for an M&S study is a real-world problem in an existing or proposed system (Robinson, 2004). For M&S studies in complex social systems, problem formulation requires a participative process (Jahangirian et al., 2015). A participative approach does not necessarily involve formal qualitative methods or problem-structuring methods, although these methods can support the elicitation of system requirements (Powell & Mustafee, 2016). Participative practice is an approach to research which incorporates local knowledge into research and planning, and collaborative activities in an iterative, flexible design (Cornwall & Jewkes, 1995). Balci and Nance (1985) and Balci (2012) outlined procedures for problem formulation, verification of the problem and a set of indicators for measuring errors of problem formulation. These emphasise both the importance of adequately formulating a problem, and of appropriate stakeholder engagement, and good methods for capturing outputs from this engagement. In Design Science studies, this stage has been supported by the literature as theoretical propositions (Carlsson, 2006), from site visits, direct observation (McKenney & Reeves, 2018), workshops (Blessing & Chakrabarti, 2009), and other methods such as interviews or questionnaires (Salehi & McMahon, 2009). In healthcare, studies argue for the necessity of involving patients and the public in quality improvement interventions (Ocloo & Matthews, 2016). The most appropriate methods for defining the problem and conceptualising the objectives of the modelling process depends on the particular situation, however some general principles from the literature can support a generic conceptual framework for real-time HM approaches. SA is integral to short-term decision-making, but can be enhanced or impeded by the provision of new information. For example, within a decision-support system, poorly presented information can increase stress and workload. For this reason, SA provides an overarching conceptualisation of the problem definition. The following sections explore models of SA, as introduced in Chapter 2 (Section 2.4.2) and how they might inform a framework for short-term decision-support, as a precursor to short-term decision-making.

4.4.1 *Existing literature*

4.4.1.1 *Individual situation awareness*

Endsley (1995) maintained that performance will be impeded where SA is incomplete or inaccurate, hence the purpose of a real-time decision-support tool is to project the development of a situation in an existing physical system over a short time period for short-term decision-support, thereby contributing to enhancing SA. SA is most relevant in a highly dynamic environment (Chiappe et al., 2012; 2015) and is a state of knowledge that provides the primary basis for subsequent decision-making (Endsley & Garland, 2000). As discussed in Chapter 2, it occurs at three levels: the perception of elements in the environment, comprehension of their meaning, and the projection of their status into the near future (Endsley, 1995; 2016).

An ‘agent-in-the-world’ model of knowledge creation (Boisot and Canals, 2004) describes how an individual receives stimuli from the physical world, perceives the stimuli as data (Level 1 of Endsley (1995), i.e. perception of elements in the current situation), conceptualises it in the context of their own expectations (Level 2 of Endsley (1995) i.e. comprehension of current situation), and alongside their own stored mental model and values, computes this as knowledge and acts upon it. This shares commonalities with Naturalistic Decision Making (Klein, 2008) where feedback loops seek additional information from the environment where the situation is unfamiliar or unexpected. Naturalistic Decision Making is defined as ‘the way people use their experience to make decisions in field settings’ (Zsambok & Klein, 2014), and is concerned with how people make decisions in complex, real world, uncertain contexts that can require real-time decisions in urgent situations. Real-time, data-driven information seeks to reinforce environmental cues to support faster, and more accurate, decision-making.

Figure 4.3 is adapted from Endsley’s (1995) three-level model of SA.

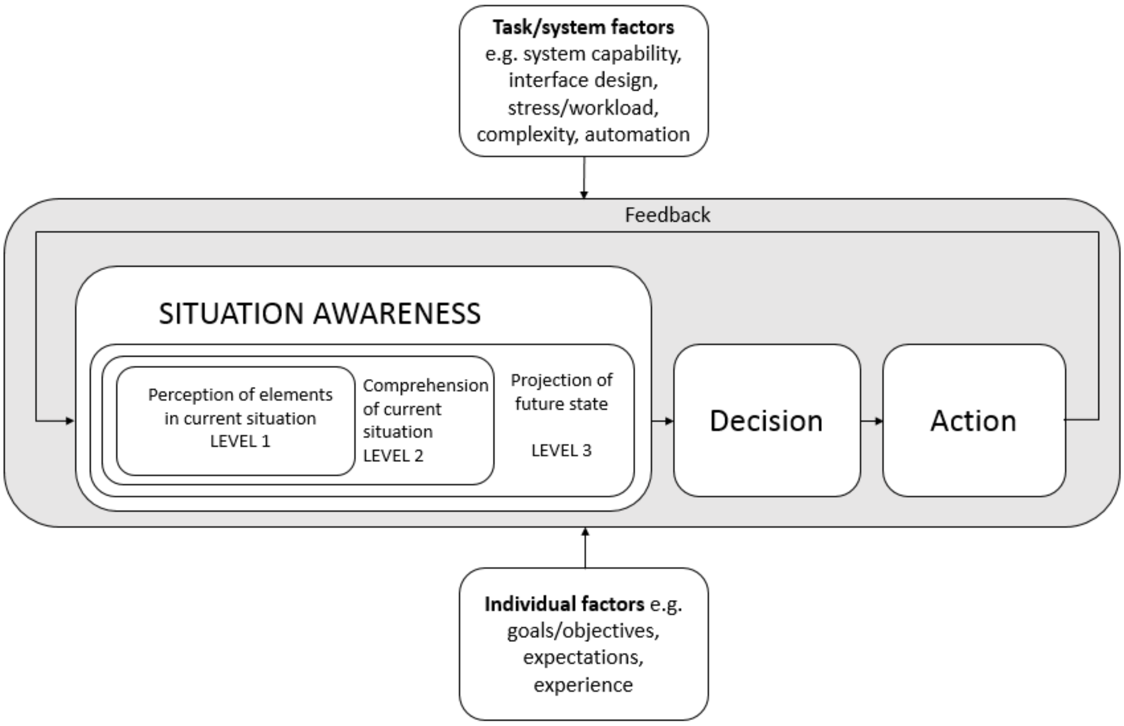


Figure 4-3 Three-level model of situation awareness in dynamic decision making, adapted from Endsley (1995).

In this model, knowledge creation (in the form of SA) involves perception of elements in the environment (Level 1), comprehension of their meaning (Level 2), and the projection of their status into the near future (Level 3). This final stage involves a decision-maker projecting how the situation will evolve in a future state, prior to taking a decision and acting upon it. The results of the action inform SA in a feedback loop, however the feedback may not be immediate. Real-time data can support this feedback loop by providing immediate information that is not readily visible in other ways, contributing to awareness of the current state of a situation by updating users' immediate knowledge and experience to make fast decisions that can inform adaptive action. This is conceptualised in Figure 4.4, adapted from Figure 4.3.

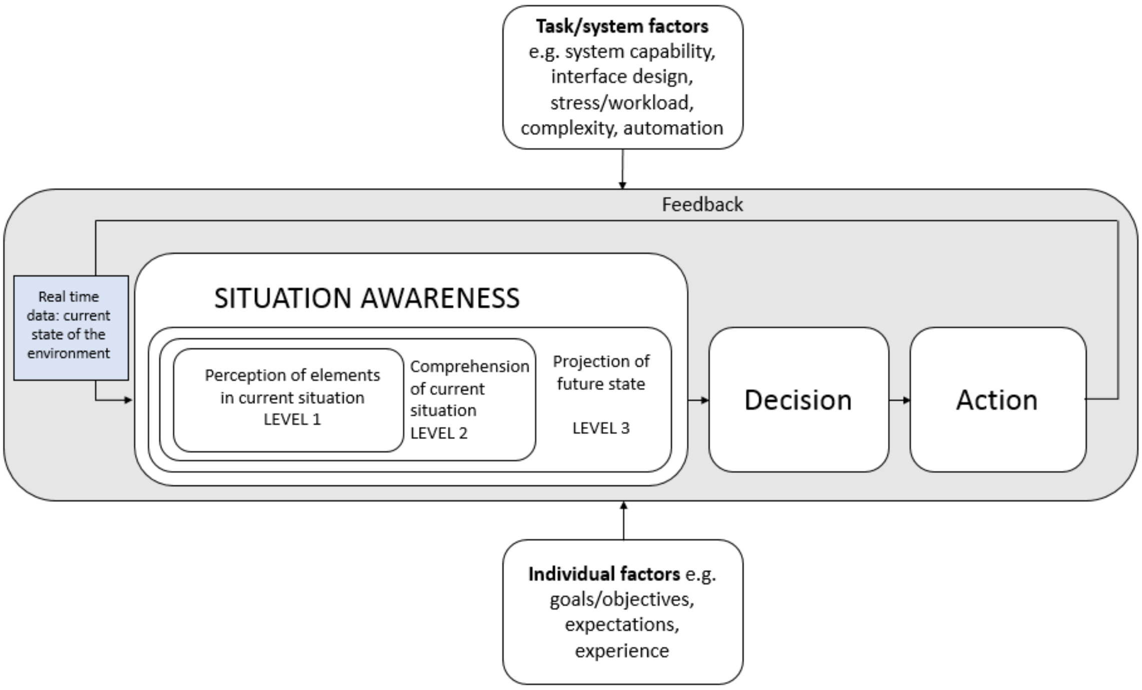


Figure 4-4 Three-level model of situation awareness in dynamic decision making, adapted from Endsley (1995), including a conceptualisation of how real-time data can support situation awareness and performance.

In Endsley's model, as explained in Chapter 2 (Section 1.4.2.1), a range of individual and environmental factors influence SA, decision-making and action, and subsequent performance. Environmental factors are relevant for the design element of the HM, as unanticipated effects can result from the type or presentation of information, for example, technology-induced errors (Peute et al., 2013; McGeorge et al., 2015). Similarly, IT systems that provide ambiguous information or with poor usability can actually reduce human decision quality and speed (Endsley, 2016). These have implications for evaluation, which is a necessary component of the framework, as concluded in Section 1.2.

The relevance of individual factors for a real-time decision-support tool depends on the problem situation. In social systems, with many uncertainties, the HM aims to reduce uncertainty rather than to provide definitive solutions: "*People are active participants in determining which elements of the environment will become a part of their (Level 1) SA by directing their attention based on goals and objectives and on the basis of long-term and working memory*" (Endsley, 1995, p .41). This protects the autonomy of workers to make informed decisions for which they are accountable, based on multiple sources of information, including their own experience, knowledge and instinct (Reddy et al., 2020).

However decisions are often made at the team level, where successful performance requires that team members maintain individual SA as well as shared SA. Specifically, shared SA requires team members to have an understanding of the type of information needed by others, knowledge of the devices used to distribute SA (e.g. dashboards), shared team processes to facilitate sharing of relevant information, and shared mechanisms such as a shared mental model. According to Stanton et al. (2006), each team member plays a role in the development and maintenance of other agents' SA. An agent with limited or degraded SA can enhance or update his/her SA through interaction with another agent, which may be human or non-human (e.g. documents, displays, etc.) (Stanton et al., 2006; Salmon et al., 2008).

4.4.1.2 *Team and distributed situation awareness*

Team SA comprises a team's collective awareness of a situation. Team members must possess SA related to their individual roles and goals within the team (some of which may be common or 'shared' with other team members), whilst also holding SA related to other team members, including an awareness of other team members' activities, roles and responsibilities, and also to the team overall, including goals and performance.

SA-related data and knowledge are distributed around the team through team processes such as communication, coordination and collaboration and serves to inform and modify team member SA, which is also informed and modified by the overall team's SA. This is represented in Figure 4.5, reproduced from Salmon et al. (2008), to illustrate how in a sociotechnical system, where members' workload has both individual task work (task-level) and teamwork (system-level) elements, SA is required at both levels. Information which contributes to both individual SA and shared mental models at the system level is therefore required to support system processes. This provides an overarching conceptualisation toward approaching and defining the problem. In Figure 4.5, a tripartite composition of team SA is apparent: individual team member SA (some of which may be common or 'shared' with other team members); SA of other team members; and SA of the overall team. Each of these forms of SA is impacted by team processes. In this model, the 'data' represents information coming in from the environment, which may be observations, technology, documentation etc.

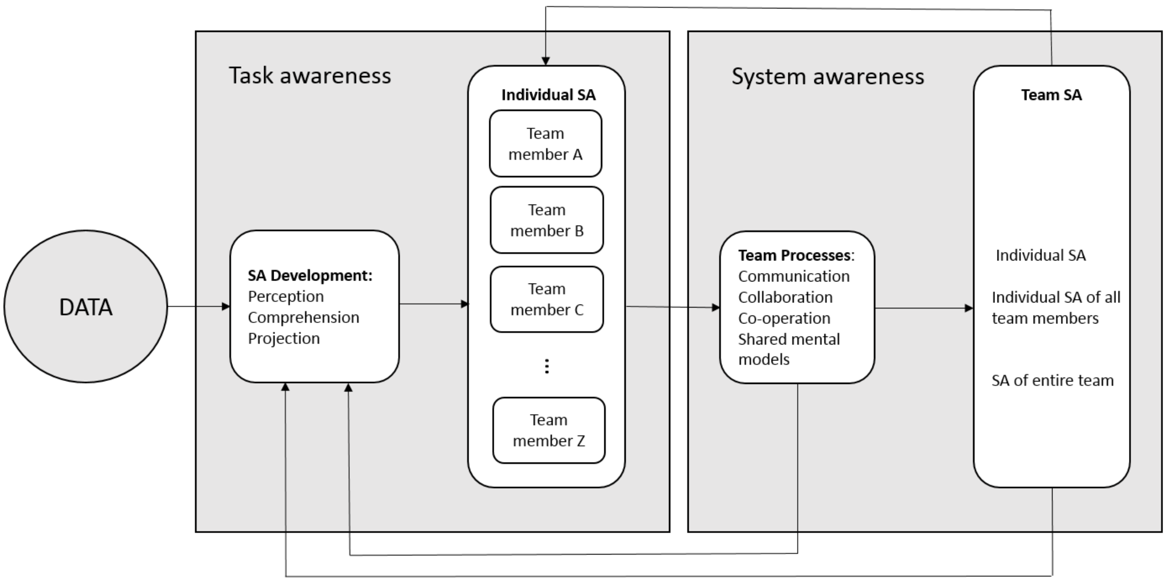


Figure 4-5 Model of team situation awareness, reproduced from Salmon et al. (2008)

While it is apparent that people can have both individual and shared SA, and that this can be influenced by both environmental information ('data') and team processes such as verbal and non-verbal communication, Salmon et al. (2008) conceptualised Distributed SA (DSA) as a network of humans and artefacts (e.g. technology) at the system level, where links exist at the artefact-artefact, human-human, and artefact-human levels. The human-human links represent team SA. DSA approaches therefore view team SA not as a shared understanding of the situation, but rather as an entity that is separate from team members' cognitive processes, and a characteristic of the system itself (Artman and Garbis 1998). Here, the SA of a team is distributed not only throughout the agents comprising the team, but also in the artefacts that they use in order to accomplish their goals (Stanton et al., 2006). In this conceptualisation, the information held by one agent modifies, and is modified by, other agents' information, creating an interacting network between people and artefacts. As decisions and actions (as per Figure 4.4) are taken based on SA which modify the environment, these changes will be reflected in real-time data applications as they occur, updating both task-level and system-level SA. This is conceptualised in Figure 4.6 (adapted from Figure 4.5), where real-time data informs SA at the task and system-level, and subsequently reflects changes in the system as a result of decisive action. A dotted line is used to indicate that these changes occur indirectly, through human decision-making and action.

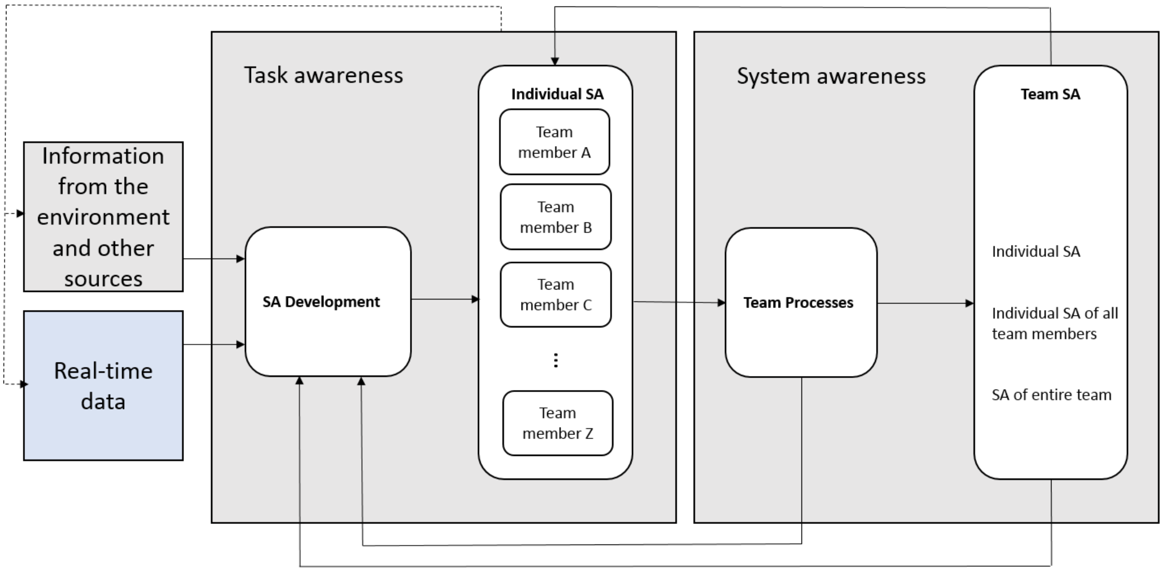


Figure 4-6 Model of distributed situation awareness, adapted from Salmon et al. (2008) conceptualising the contribution of relevant real-time data to support SA

Models of SA show how up-to-date environmental knowledge augments decision-making which informs action. Where real-time data, and applications which use real-time data, are a component of the environmental information, a feedback loop from actions can update environmental knowledge and SA for ongoing decision-support.

This closes the loop between environmental information→SA →decision-making → action by updating environmental information and SA as a result of any actions which make changes in the system. This information must align both individual (task awareness) and team (system awareness) knowledge. This forms the basis of an architecture for a framework using real-time data for short-term decision support by clarifying a generic element of the problem to be addressed (deficient

SA at task- and system-levels) and modelling objectives (enhancing SA for short-term decision-support in dynamic social systems).

While real-time data provides information, the value of the information it provides can potentially be enhanced by the use of analytics methods, such as prediction and simulation, which will be examined in Section 4.5. Existing real-time simulation frameworks will now be examined, and how they can be mapped to models of SA.

Key implications for a real-time Hybrid Modelling framework

In a sociotechnical system, members' workload has both individual task work (task-level) and teamwork (system-level) elements. SA is required at both levels for decisions which influence system functioning. This provides an overarching conceptualisation toward approaching and defining the problem. Real-time analytics can support SA by accurately reflecting the current state of the system for all team-members and augmenting decision-making for individuals and teams. However SA models emphasise that many individual and system factors, as well as team-processes, influence decision-making, so the autonomy to act remains with the decision-maker. Should action take place, changes in system behaviour as a result of action are reflected in the real time data, closing the loop between information, SA, decision and action.

4.4.2 *Existing real-time simulation frameworks*

One important component of the real-time HM is real-time simulation. This describes a simulation which can dynamically accept and respond to real-time data from the physical system to improve the accuracy of the model. However real-time simulation can also be mutually beneficial, in that the simulation system not only experiments with scenarios to change the physical system, it also accepts and responds to data from the physical system (Fujimoto et al., 2002). The physical system benefits from augmented decisions, and the simulation system benefits from the updated data.

Figure 4.7 illustrates the structure of a symbiotic simulation system, reproduced from Fujimoto et al. (2002). Via a data acquisition system, real-time or near real-time data is taken from the physical system. A control or decision-support function conducts "what if" experiments to investigate alternative scenarios. From an analysis of the output results, the physical system is optimised so that its performance is improved. In this representation of ‘symbiotic simulation’, which emphasises a mutually beneficial, continuous process, the results are also fed back to the control function for automatic validation and subsequent decision-making. The term symbiotic simulation reflects the close relationship between a simulation system and a physical system.

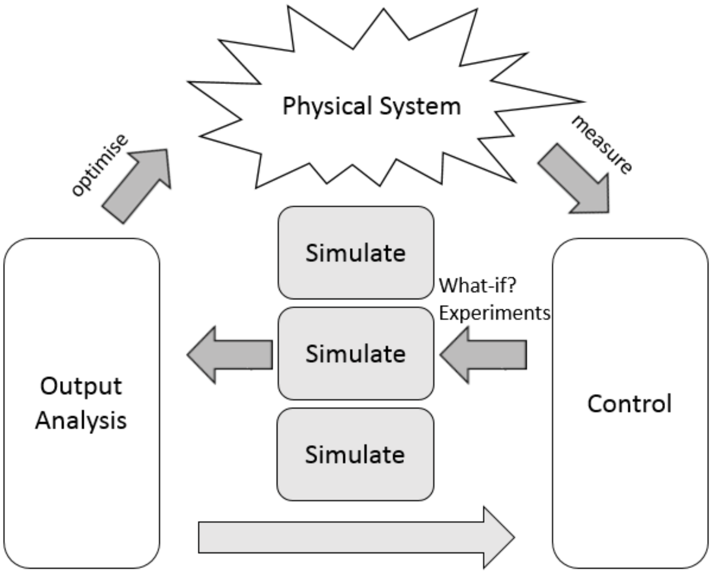


Figure 4-7 Real-time (symbiotic) simulation architecture, reproduced from Fujimoto et al. (2002)

The extended definition provided by Aydt et al. (2008a) differentiated between closed-loop, as per Fujiomoto et al. (2002) - where the decision is proposed to an external decision-maker, or directly implemented by means of actuators - and open-loop simulation systems, where no feedback is created to the physical system. Open-loop methods are concerned with decision-making, while closed-loop real-time simulations are used for forecasting, model validation, and anomaly detection (Aydt et al., 2009a).

In the closed-loop systems, either an actuator directly controls the system, or a decision-maker can control the physical system, rather than implementing the decision directly. In this case, control of the system is indirect as the decision belongs to the individual. Any changes in the physical system will be reflected in the real-time data, via the human-in-the-loop between the physical system and the simulation system. This is illustrated in Figure 4.8, reproduced from Aydt et al. (2008a). While this accurately reflects the conceptualisation of the real-time simulation as used here, in this research the term ‘real-time’ simulation is used, in preference to ‘symbiotic simulation’ to reflect the fact that other uses of the real-time data are proposed in this framework to support its intended purpose. The mapping between information, SA, decision and action conceptualised in models of SA can be seen with Figure 4.8, however the SA, decision and action components are all subsumed into 'External Decision-Maker", while the grey box in Figure 4.8 opens up the 'environmental information' components to illustrate the broad stages of a simulation.

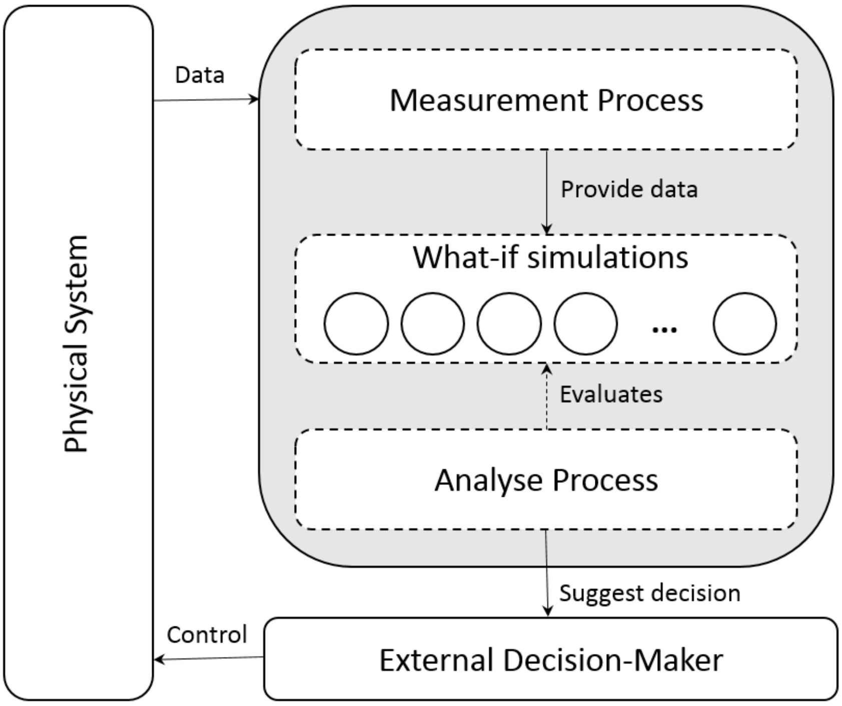


Figure 4-8 Human-in-the-loop closed-loop symbiotic simulation, reproduced from Aydt et al. (2008)

Other frameworks have been proposed in the literature. Tavakoli et al. (2008) proposed a framework that focussed on the data integration and processing layers, and their components, however they also conceptualised a 'data matching mechanism' which matched physical data generation with the simulation processes in an open-loop design. Similarly, Song et al. (2008) developed an open-loop framework for real-time simulation for heavy construction operations which conceptualised how real-time and historical data combined with process knowledge can update both the structure and input data of a simulation model. The open-loop framework proposed by Mousavi et al. (2011) illustrated how their model can reduce the time gap between measurement and improvement in the NHS, while Oakley et al. (2020) presented a symbiotic simulation as an early-warning system for hospital bed planning. Both used open-loop conceptualisations, however Hammad et al. (2012) conceptualised a two-way relationship between information sources, real-time simulation and site operations components for construction. The healthcare model proposed by Bahrani et al. (2013) also presented a closed-loop design explicitly incorporating decision-making using a human-in-the-loop, as the decision-maker is required to quickly and accurately evaluate alternatives and implement a decision to maintain the system in a healthy state. In their conceptualisation, the real-time monitoring engine enables process changes to be observed. All real-time simulations share enhancing SA as part of their common purpose. The human-in-the-loop architecture can be readily mapped with SA models to emphasise the purpose of the real-time simulation in supporting SA at both task- and system-levels to augment subsequent decision-making and action. This is illustrated in Figure 4.9, adapted from the model by Aydt et al. (2008a) in Figure 4.8.

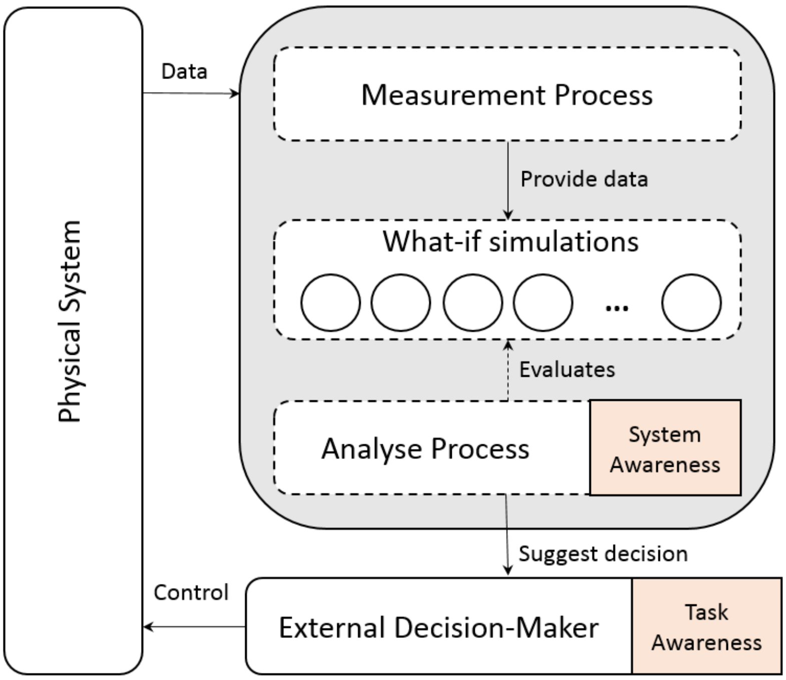


Figure 4-9 Human-in-the-loop closed-loop symbiotic simulation, mapped with SA (system-level and task-level), adapted from Aydt et al. (2008)

Any simulation method may be appropriate depending upon the problem. DES has been commonly used in real-time simulation as per the above discussion; ABS is also used (e.g.Low et al., 2005; Seekhao et al., 2016; Rivas & Cahmoso, 2017; Rasheed et al., 2019), and SD has been proposed (Zhang et al., 2017).

This section has provided an examination of real-time simulation frameworks, concluding that the human-in-the-loop design for supporting short-term decision-making augments SA at system and individual levels. This can be mapped to models of SA examined in Sections 4.4.1. As discussed in Chapter 2 in Sections 2.5 and 2.7, a range of analytics techniques can add further value to data for decision-support, in addition to simulation. These are now examined in the following sections.

Key implications for a real-time Hybrid Modelling framework

The human-in-the-loop conceptualisation of real-time simulation implicitly illustrates how real-time simulation models enhance SA at both system and individual levels to augment decision-making. Real-time simulation can support all levels of individual SA: perception of the current situation, comprehension of the current situation, and projection of the future state. In common with models of SA, actions as a result of these decisions will change the behaviour of the physical system. These changes will be reflected back into the real-time data for ongoing visualisation and analysis. This human-in-the-loop cycle forms the outline of the framework, to emphasise both the relationship between the simulation model and the physical system, and that human decision-making can be enhanced by a real-time decision-support tool.

4.5 The use of Data Analytics in Hybrid Modelling

As defined in Chapter 2, Data Analytics (DA) is the use of data, statistical and quantitative analysis, and fact-based management to drive decisions and actions. DA methods have been applied with simulation to support different stages of an M&S project, to combine the benefits from specific methods, to overcome the weaknesses of specific methods, or to consider the wider system in a modelling problem (Marshall et al., 2016; Greasley & Edwards, 2019). Marshall et al. (2016) argued that simulation and other DA methods offer distinct but complementary value in healthcare.

Hospitals are increasingly able to accumulate large amounts of operational data due to advances in data collection, storage and use of standards (Chen, Lin & Wu, 2020). While hospital data is not ‘big data’, much of it is generated at high velocity and in a variety of formats, including multiple hospital input databases, mobile devices, Internet of Things sensors, and patient sensors. It therefore exhibits the characteristics of ‘big data’ in that it is growing exponentially in volume, velocity, variety and veracity (Galetsi & Katsaliaki, 2019 b,c). However, while the veracity of the data is improving in terms of its accuracy (e.g. Mbizvo et al., 2020), much of it is inputted by clinical staff who are prioritising patient care, and may suffer from inaccuracies. A further significant challenge to data analytics in healthcare is the need to standardise and secure the process of extracting healthcare datasets (Galetsi et al., 2019). Nonetheless, the potential value in using data to support decisions is undisputed. The challenge has shifted from collecting and storing sufficient data, to using the data to add value to the healthcare organisation.

Figure 4.10 illustrates a widely used conceptualisation of DA categories, functions and examples (Shao, Shin & Jain, 2014; Khalifa & Zabani, 2016). This hindsight-insight-foresight hierarchical framework is adapted from Davenport and Harris (2007) who emphasised that the degree of intelligence rises as the methods progress from access and reporting, to predictive and prescriptive analytics. Descriptive analytics involve observing historical data, diagnostic methods involve exploratory analysis, predictive analytics involve prediction of future observations, while prescriptive analytics enable the best course of action to be determined, under certain circumstances, supporting the ability to influence the system towards its goal performance.

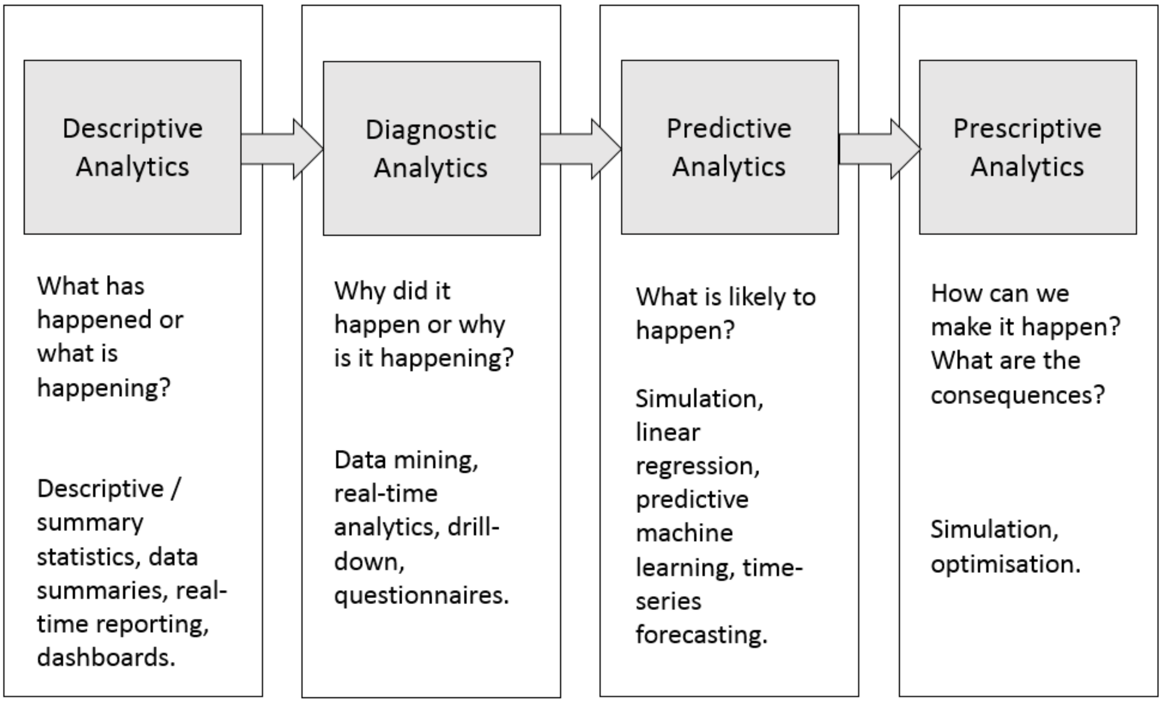


Figure 4-10 A functional categorisation of data analytics, adapted from Shao, Shin & Jain (2014)

These are explained in more detail in the following subsections.

4.5.1 *Descriptive analytics*

Descriptive analytics analyses and presents data using techniques such as descriptive or summary statistics, real-time reporting, graphs, charts and dashboards, that is, traditional business intelligence and visualisation techniques (Chen & Storey, 2012; Saxena & Srinivasan, 2013; Delen & Zolbanin, 2018). Data is collected, maintained, and processed to allow decision-makers to quickly assess performance against Key Performance Indicators (KPIs) by comparing current performance against targets for business objectives (Peral et al., 2017). Descriptive analytics converts raw data into meaningful information, and information into insights for decision-support (Evelson, 2010). It is also used to prepare the data for further analysis (Delen & Zolbanin, 2018). While Haas et al. (2011) asserted that historical data alone, no matter how it is presented, remains simply a record of history which provides limited insights or solutions, Mustafee et al. (2018) argued that the combined use of historical and real-time data can alleviate some of these criticisms. Historical analytics give organisations insights into past events, but real-time descriptive analytics allow information to be used as the situation is unfolding, changing the operational environment in the present.

Despite the plethora of data available, gaining access to real-time data in healthcare can be challenging. In the UK, NHS Information Governance is complex, with a legal framework which includes the NHS Act 2006, the Health and Social Care Act 2012, the Data Protection Act, and the Human Rights Act, and a raft of NHS standards and guidelines for information security. Due to the enormous complexity of Information Governance and patient confidentiality codes of practice, specific arrangements for NHS research ethics, and internal policies protecting patient-level information access (Department of Health, 2007), few NHS staff members will be fully confident in the detail of their hospital's Governance Framework. While the secondary use of data is recognised as essential for improving the quality of health services (NHS England, 2019), it can be challenging making a case for accessing NHS data, especially for the continuous feeds required for a real-time application. Additionally, the data must be validated by NHS staff for the purpose for which it is required, which will mean agreeing an information exchange standard for sending data from the hospital IT management system, and the data-exchange format. The quality and accuracy of the data will also need to be considered.

For this reason, if a real-time data research application is planned in the NHS, it is recommended that the starting point is to determine what data is available, or could be made available, and how often it is required. Secondly, consideration should be given to whether the data will be used internally only, or if it needs to be made available external to the organisation. A third consideration is maximising the value that can be gained from the real-time data that is made available. Finally, historical data and other data sources may also be required to meet the study objectives.

4.5.2 *Diagnostic analytics*

Diagnostic analytics is considered to take descriptive data a step further to understand why an event or performance happened. It uses exploratory data analysis including correlations, data mining, root cause analysis, and drill-down and drill-through processes, focusing on processes and causes. This may include understanding the impact of input factors and operational policies on performance measures (Shao et al., 2014). Diagnostic analytics is considered to require domain knowledge, and uses existing data, or may need additional data to be collected (Khalifa & Zabani, 2016; Delen & Zolbanin, 2018; El Morr & Ali-Hassan, 2019).

Diagnostic analytics may be useful as part of a HM study using real-time data, depending upon the objectives. However one key use of diagnostic analytics in this framework is determining the conditions for triggering interactions between methods, that is, how the methods interact when the components of the HM are run. This was introduced in Section 1.2. Events in one method are implicitly triggered by threshold levels in another, therefore there is a variable time gap between the different methods in the integrated model, where events are triggered by the state of the system.

Aydt et al. (2008b) highlighted a wider application, where triggers do not necessarily occur in line with changes in the physical system. He categorised triggers as reactive, preventative, or proactive. According to these definitions, a reactive trigger can be observed in the physical system, and are events that require immediate action. A preventative trigger is observed in forecasts, and therefore is limited to conditions that can be forecasted. Finally, a proactive trigger occurs at fixed, regular intervals for continuous improvement, and does not rely on the notion of a triggering condition.

In Chapter 2 (Section 2.7.5) attention was drawn to the potential advantages of using simulation to plan for recovery based on forecasts of a critical event. For staff rostering, this combined approach was used by Park et al. (2008) and Lin and Chia (2017), while Augusto et al. (2018) proposed to implement preventative triggers in their M&S framework for supporting emergency units in times of crisis. The addition of a preventative trigger using predictive analytics, where possible, can add significant advantage to the HM by augmenting decisions that allow the system to recover before the critical situation has actually occurred.

4.5.3 *Predictive Analytics*

Predictive Analytics, in its most general sense, refers to any method which can predict future observations, including machine learning, data mining, forecasting, and mathematical approaches such as time-series approaches (Delen & Demirkan, 2013; Waller & Fawcett, 2013; Shao et al., 2014). These can also include cross-sectional data leading to a categorical prediction through classification, judgemental approaches (Fildes et al., 2008), and simulation (Shao et al., 2014; Adra, 2016). More recently, predictive analytics often takes the form of data-driven machine learning methods for making predictions by specifying the values of new observations based on the structure of the relationship between inputs and outputs (Abbott, 2014; Mortenson et al., 2015; Delen & Zolbanin, 2018) and is often considered to be a subset of, or synonymous with, 'big data' applications (e.g., Koh & Tan 2005; Janke et al., 2016; Vidgen et al., 2017).

Taking the broader perspective, predictive analytics describes a set of methods which have been used extensively in healthcare to predict events based on prior foreknowledge from historical data and other sources of information (Soyiri & Reidpath, 2013). The purpose of predictive analytics in this framework is to predict a critical event, such that subsequent decisions are 'preventative' rather than 'reactive'. A range of methods could be used as appropriate to the available data.

Using reactive real-time simulation, the purpose of the simulation is to find a solution that recovers and mitigates the effects of the detected critical situation. This is possible because the simulation runs much faster than real-time. In this case, the simulation and its multiple runs must be completed in a short time-frame in order to be useful for subsequent decision-making. Aydt et al. (2008b) contrasted this approach with a preventative trigger, termed 'preventative what-if analysis', where a critical condition is predicted, triggering the simulation. In this case, the purpose of the simulation is to prevent the critical situation from arising in the first place. This is illustrated in Figure 4.11, reproduced from Aydt et al. (2008b). In this figure it can be seen that recovery to a normal operational state will be faster the earlier the critical condition can be detected.

The accuracy of the forecasts with regard to both false negatives and false positives is an important consideration. For example, in a safety-critical application, failing to detect a critical situation may be a greater error than over-detecting and compensating unnecessarily. In healthcare situations, this may often be the case, and both types of error may need to be handled. A further consideration is that not all critical events can be predicted, for example a major incident or event such as large traffic or rail incident.

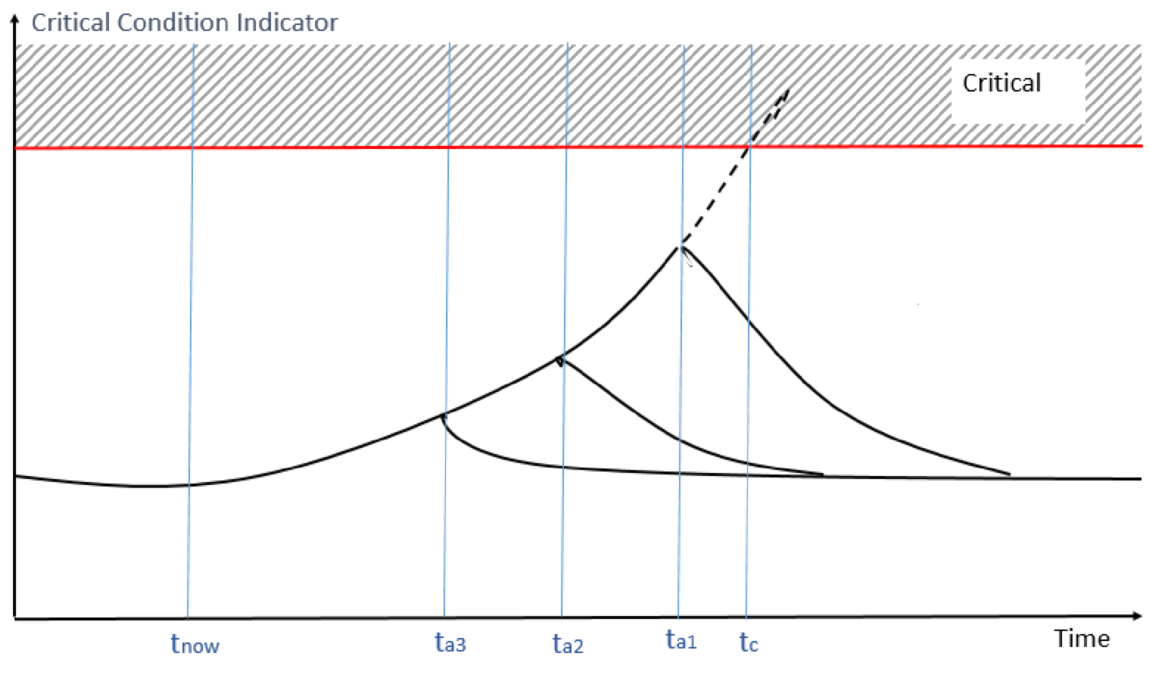


Figure 4-11 The timing of detection of a critical condition, and its relationship with recovery. Reproduced from Aydt et al. (2008b)

Figure 4.11 illustrates where triggers might occur based on forecasts to prevent a critical situation from escalating. This enables the real-time simulation to run, potentially providing solutions to recover from the forecasted or escalating situation. The next section discusses prescriptive analytics in a general sense, tying back to Section 4.4.2, which examined frameworks for real-time simulation.

4.5.4 *Prescriptive analytics*

Prescriptive analytics informs decision-making by suggesting a solution path, for example, simulation can anticipate the consequences of unforeseen interactions and prescribe interventions on the basis of tested scenarios (Marshall et al., 2015), while optimisation is a prescriptive method as it suggests the 'best available' values for a given function (Hoad et al., 2015). Haas et al. (2011) argued that prescriptive models and what-if analysis should be on an equal footing with other analytics methods to make sense of real-world complexity and support real-world decision-making. In practice, prescriptive analytics can continually and automatically process new data to improve recommendations and provide better decision actions (Delen & Zolbanin, 2018). A key challenge is to facilitate integration of datasets, along with simulation, analytical, statistical, and optimisation models, and Haas et al. (2011) suggested that research is needed to understand whether these integrated tools are feasible, practical, flexible, cost-effective, and usable.

Figure 4.12 illustrates how analytics approaches are proposed to interact in the real-time hybrid modelling framework for short-term decision-support, adapted from Figure 4.10.

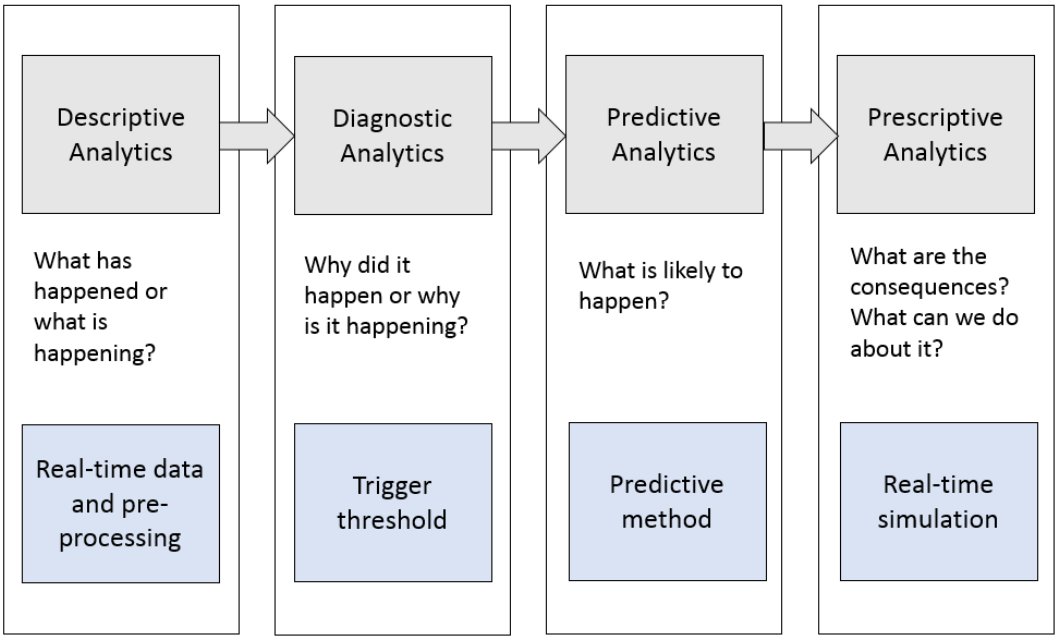


Figure 4-12 Proposed use of the functional categories of data analytics in framework (Adapted from Shao, Shin & Jain, 2014)

Key implications for a real-time Hybrid Modelling framework

A range of descriptive, diagnostic and predictive methods can be used to support real-time simulation in a HM. The real-time data will need to be accessed, validated and processed. The conditions for triggering interactions between methods must be determined. Where a critical situation can be forecasted, a predictive component can be used to trigger the simulation, aiming to prevent the onset of a critical situation. Finally, the simulation model can support decisions toward either preventing the critical situation or mitigating its effects. These components form the backbone of a real-time HM framework, supporting a range of applicable methods.

4.6 Integration of analytics methods

As described in the previous section, the proposed approach can be used in a HM, whereby the problem situation is addressed using multiple methods, supporting synergies between the methods. While the research design is sequential, as each element follows the next sequentially, the HM is designed to be a recurrent use decision-support tool, during busy periods, so minimal manual interaction is important. This information must be comprehended by staff without interrupting workflow, hence automation and integration of components within the HM is required. As discussed in Section 4.2, the number of methods, the frequency of interaction, the number of points of interaction, and type/frequency of triggers will be determined by the specific application.

The right infrastructure is important, and the integration component represents a process within which the individual methods interact to form a single, complete model. Figure 4.13 conceptualises the integration of the HM. The real-time data is accessed and pre-processed as required. This will occur at predefined intervals, proactively triggering the preventative component. For a preventative trigger, a set of values are sent to the predictive component which returns a forecasted value. Should a reactive or proactive simulation trigger be required, the predictive component is not executed. Given a predefined threshold, the real-time simulation is triggered, and returns decision support. Resultant action will be reflected in the real-time data. This forms the HM component of the framework, maximising the value that can be gained from the real-time data for short-term decision-support.

In Figure 4.13, the cycle identified in Section 4.4 can be mapped by following the arrows. The real-time data is integrated with a prediction model (an optional component) and a simulation model. Either the real-time data or the predicted data can trigger the simulation. The simulation provides information which can support a decision. Actions as a result of the decision change the physical system behaviour and are therefore reflected back in the real-time data. In Figure 4.13, the dashed line between the real-time data and the integration component reflects the predefined intervals for updating the data. The arrow between the decision of a manager (or other stakeholder) and the real-time data is dotted, as the 'human-in-the-loop' retains the autonomy to take a different decision.

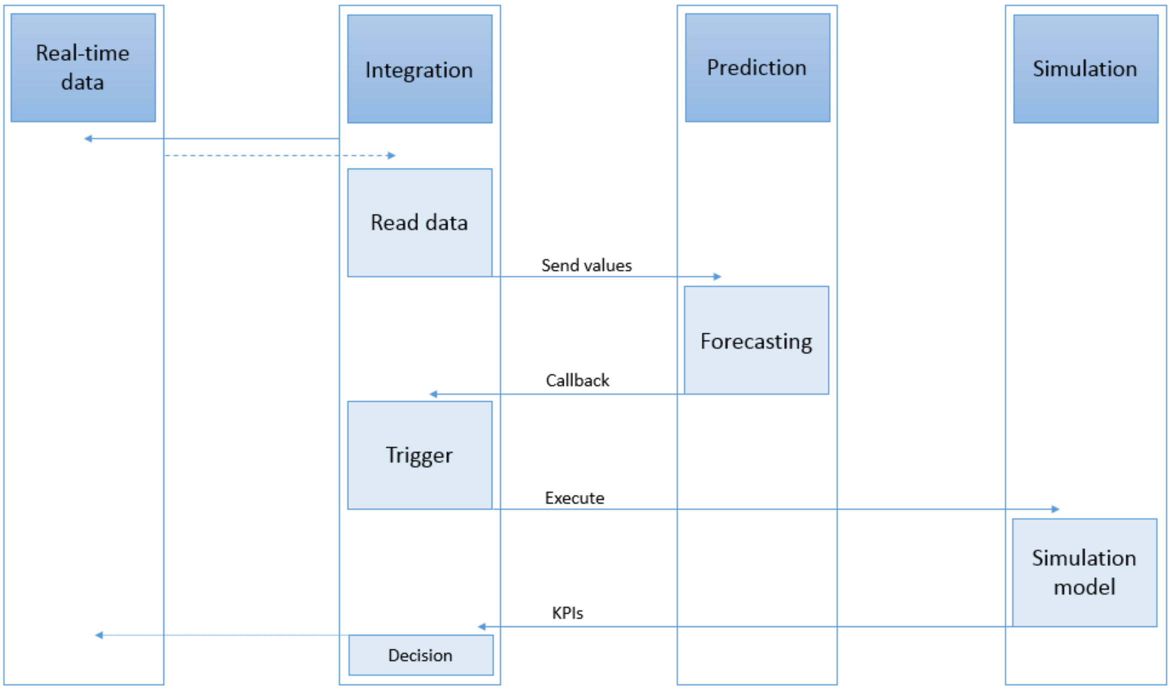


Figure 4-13 A conceptualisation of integration of components

As discussed in Section 4.3, evaluation of the HM is necessary to provide feedback for improvements and iterations. Evaluation can take many forms, and Venable et al. (2016) highlighted the possibility of reducing risk by evaluating early, before undergoing the cost and effort of building an artefact. This may be integrated into the problem definition phase as a 'formative evaluation', defined as "*the systematic assessment of the worth or merit of some object*" (Nieveen & Folmer, 2013). This identifies both its inherent, intrinsic value, and its contextually determined value. This is discussed in the next section.

Key implications for a real-time Hybrid Modelling framework

For real-world recurrent use, all sequential components of the HM should be integrated and automated in a way that supports comprehension by staff without interrupting workflow. This creates a single HM, with a number of elements, and with a single purpose.

4.7 Evaluation of the model in context

In healthcare, real-time simulation is a relatively new approach, although a number of studies have progressed the methodology (Tavakoli et al., 2011; Espinoza et al., 2013; Tan et al., 2013; Oakley et al., 2020). The majority of studies are proof-of-concept or prototyped applications, hence none of these applications have been implemented in the real-world. This means that to date, it is difficult to determine whether such an approach would be considered useful or applicable in the healthcare domain, and the circumstances under which it would be useful.

An evaluation should be undertaken to determine whether the model has the expected effect on decision-support suggested from the literature and existing studies, to provide feedback for further development, and to explore any unexpected side effects. Within a Design Science methodology, the criteria for evaluation are an important component of the problem definition phase. The success or otherwise of a modelling study may require sufficient analysis of the research process to enable conclusions to be made regarding the conditions under which the model was successful. A reflective understanding of its limitations and how it is being used can ultimately increase the level of trust and confidence toward successful implementation. Peffers et al. (2007) suggested that within Design Science, the model (artefact) should be demonstrated in situ, or through experimentation, simulation, case study, proof-of-concept etc., to observe and measure how well it supports a solution to the problem. There are clear advantages to evaluating the model in use by real users solving real problems where possible. This may require quantitative or qualitative analysis as appropriate. Following evaluation, researchers can decide whether to iterate back to the previous activity to try to improve the effectiveness of the artefact or to leave further improvement to subsequent projects.

Issues such as reliability, accuracy, validation and verification of the model should be integrated into the development of each stage of the HM build. The purpose of this evaluation stage is to evaluate the functionality of the model from the user perspective, however it should also address elements such as the usability of the model, any barriers to its implementation, its contribution to supporting SA, and consideration of unintended consequences in practice. As the HM intervention proposes to be assimilated into routine operational practice, use resources, and requires a certain technical infrastructure, these should be addressed in the evaluation.

A framework for Design Science research evaluation was proposed by Venable et al. (2016) which considers why, when, how, and what to evaluate, across the dimensions of formative/summative, and artificial/naturalistic. The formative perspective captures the possibility of reducing risk by evaluating early. This offers the possibility of incorporating early evaluation in the problem definition phase, alongside identifying criteria for final evaluation. In Venable et al. (2016), the summative perspective offers the possibility of evaluating the artefact in reality, not just in theory. Naturalistic evaluation methods evaluate the artefact in use by real users solving real problems, while artificial evaluation methods offer the possibility of controlling potential confounding variables more carefully. The criteria identified should guide the appropriate choice of strategy. The decision to implement the results of an M&S study, or in this case to embed the tool into operations functions for recurrent use, belongs to the organisational stakeholders. A variety of factors affect the outcome of such a decision, many outside of the control of the modeller. However for future work, it is important to understand which of these can be controlled, as IT interventions are unlikely to be adopted by healthcare professionals until they are ‘fit for purpose’, and stakeholders have confidence in the intervention (Ross et al., 2016; Liberati et al., 2017). As iterations progress, continuous, systematic evaluations are likely to be needed, however the focus of evaluation will change as the intervention becomes closer to an ideal solution, and the risk of implementation failure lowers.

Key implications for a real-time Hybrid Modelling framework

Many factors influence the potential implementation of a real-time decision-support tool in practice. The chosen methodology supports iterative evaluation to determine factors which contribute to the usefulness of the approach in its applied setting, for improving the design or for informing future work. These factors go beyond the technical proficiency of the approach, to consider, for example, usability of the model, any barriers to its implementation, its contribution to supporting SA, and consideration of unintended consequences in practice.

4.8 Integrated Hybrid Analytics Framework (IHAF)

Through an examination of Design Science methodology, consideration of existing HS frameworks, and a review of the literature focusing the purpose of the study, a generic integrated hybrid analytics model (IHAF) is illustrated (Figure 4.14).

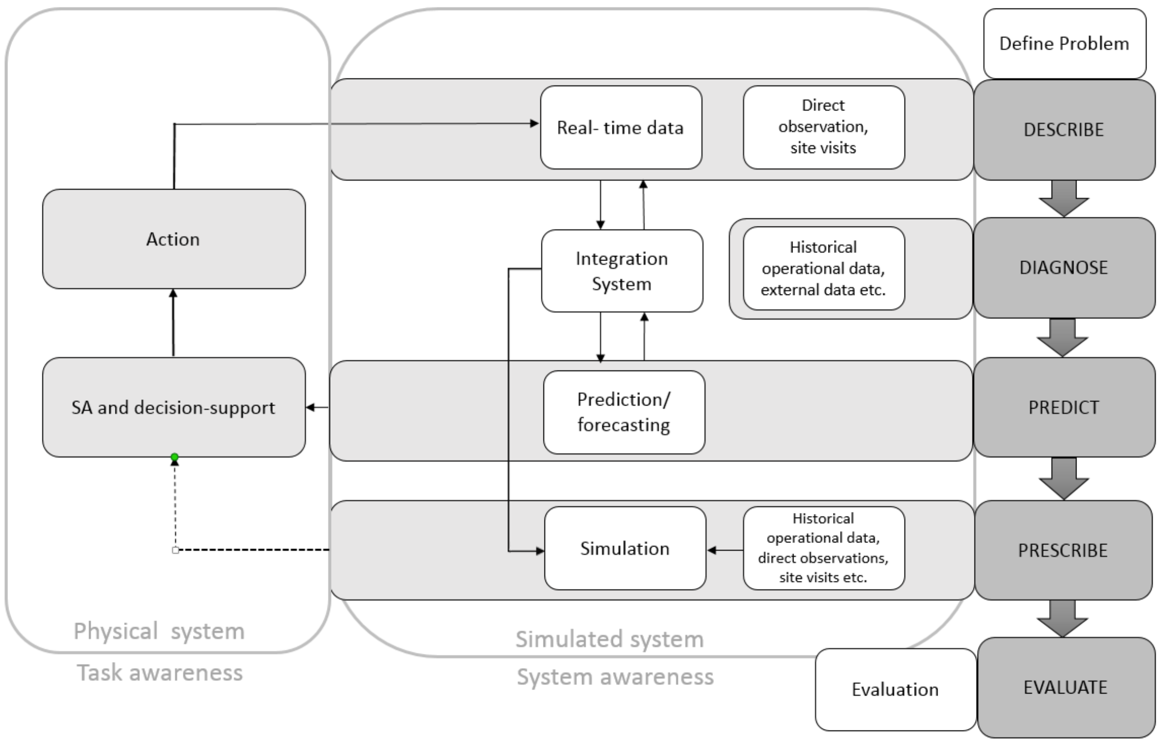


Figure 4-14 Integrated Hybrid Analytics Framework (IHAF)

The first stage, as supported by multiple M&S guidelines, is problem definition (Balci & Nance, 1985; Shannon, 1998; Robinson, 2004; Law, 2009; Martin et al., 2018). Positioned in Quality Improvement, problem definition requires participative practice, defining stakeholder groups, and consideration of methods such as site visits, direct observation, workshops, interviews or questionnaires.

From the perspective of Design Science, the real-time decision-support tool aims at improving a situation. This means determining a specific research problem, but also justifying the value of the solution by determining criteria for evaluation, which should consider more than the accuracy of the model. It should start with a system-level understanding of what matters in practice, to reduce the risk of unintended consequences in a complex system.

Through the literature review in Chapter 2, and in Section 4.4, the purpose of the model - short-term decision support in sociotechnical systems - has emphasised the role of SA as a precursor to decision-making, which focuses the purpose of the HM. As team-work is a defining feature of sociotechnical systems, SA is required at both the task-level and system-level. Additionally, the information→ SA→decision→action process is closed by a feedback loop, where system changes implemented as a result of decisive action informs ongoing SA via environmental cues and other information sources.

It is proposed that real-time information can help to close this feedback loop by updating immediate feedback. Models of SA at the team and individual levels conceptually map with human-in-the-loop models of real-time simulation, whereby the simulation model is initialised using real-time data, scenarios are investigated, and decisions are suggested to an external decision-maker. This updates system-level and task-level awareness to augment decision-making, however the autonomy of the decision-maker is retained. For this reason, control of the system is indirect. In the framework, this is represented as a dotted line between simulation processes and decision-making. Changes to system behaviour as a result of action are reflected in the real-time data, subsequently updating the simulation model at its next initialisation.

Data analytics methods require consideration for adding further value to a model for supporting short-term decision-making. The literature review concluded that a HM approach which combines real-time data, predictions, and simulation has the potential to support short-term ED decision-making. Descriptive analytics using real-time data offers value by allowing information to be used as a situation is unfolding. However it also presents significant challenges, including accessing the data in real-time, and consideration should be given to maximising the value that can be gained from any real-time data that is made available. Furthermore, consideration should be given to whether the data will be used internally, or if it needs to be made available external to the organisation. Historical data and other data sources may also be required to meet the study objectives.

Diagnostic analytics may also be important, and as outlined in Section 4.5.2, a key use of diagnostic analytics in this framework is determining the conditions for triggering interactions between methods. This may require additional operational or external data, which may need to be collected. Integration of methods are required. Events in one method are implicitly triggered by threshold levels in another, therefore there is a variable time gap between the different methods in the integrated model, where events are triggered by the state of the system. The right infrastructure is important, and the integration component represents a process within which the individual methods interact to form a single, complete model. Where the trigger is preventative, that is, triggered by forecasts, the forecasted threshold is used to trigger the real-time simulation model. In this case, the purpose of the simulation is to prevent the critical situation from arising in the first place. The combined use of descriptive, diagnostic, predictive and prescriptive analytics supports system-level SA and task-level SA for augmented decision-support which informs action, and completes the loop. Changes in the system as a result of decisive action are reflected in the real-time data, which updates both SA and the HM.

As the HM is intended for recurrent use, evaluation is necessary to provide feedback for improvements and iterations. Evaluation can take many forms, including evaluating before undergoing the cost and effort of building the model. However it is valuable to demonstrate and evaluate any iteration of the model in context toward determining the ‘usefulness’ of a real-time decision-support tool, using quantitative or qualitative analysis as appropriate. The purpose of the evaluation stage of IHAF is to evaluate the functionality of the model from the user perspective, however it should also address elements such as the usability of the HM, unexpected effects, and barriers to its implementation, which may be at different organisational levels, and between different stakeholder groups. Following evaluation, researchers can decide whether to iterate back to a previous activity to try to improve the effectiveness of the artefact or to leave further improvement to subsequent research. Each of these stages and activities has enabled the development of IHAF for supporting short-term decision-making in sociotechnical systems.

The next section indicates how this chapter contributes to addressing the research questions (RQ) in this thesis, and how it will be used in the subsequent chapters.

4.9 IHAF framework application to address research questions

This chapter has developed and proposed a HM framework, the Integrated Hybrid Analytics Framework – IHAF - for supporting the development of a HM for short-term decision-making in sociotechnical systems. This addresses the first objective of RQ2.

To test this framework, phases (b) *(Descriptive phase I),* (c) *(Prescriptive phase),* and (d) *(Descriptive phase II)* outlined in Sections 4.3 will address the second and third objectives of the second research question (Table 4.2), by applying IHAF in a use-case, and evaluating the application.

Table 4-2 Research question 2, and its aim and objectives

Research Question	**Aim**	**Objectives**
2. *How can an integrated hybrid approach using real-time simulation and data analytics support short-term operational decision-making?*	To test and evaluate the potential of an integrated hybrid approach for short-term decision-support in healthcare combining real-time simulation with analytics approaches.	**4.** To propose a generic framework supporting an integrated hybrid approach for short-term decision-making in healthcare. **5.** To apply the framework within a case study in a hospital ED. **6.** To evaluate the application of the framework.

Stage (d) (*Descriptive phase II*) intends to evaluate the potential system-level benefits of using real-time data applications for both patients and staff. This is important because:

- In the context of a sociotechnical system, patients as well as staff are considered to be integral components of the urgent-care system, and decisions of both sets of stakeholders will impact on system behaviour.

- In the context of quality improvement, healthcare service interventions should be designed to support the needs of end-users, the patients and their families, as well as those of service providers.

For this reason, a formative evaluation as well as a summative evaluation will be undertaken (Venable et al., 2016). The formative evaluation will inform both the model development (as part of the problem definition stage) and the summative evaluation, forming the evaluation stage. Additionally, both will be address RQ3:

Table 4-3 Research question 3, and its aim and objectives

Research Question	Aim	Objectives
3. What are the implications and the added value to the system of using real-time data applications for both patient and NHS decision-support?	To analyse the system level impact of the use of real-time data by both patient and staff decision support	**3.** To critically evaluate the perceptions that patients and NHS staff have regarding the value that real-time applications provide at the system level. **4.** To synthesise previous findings and to evaluate the framework in light of the application.

Figure 4.15 illustrates the mapping of the stages of the Design Science research methodology and IHAF, with each of the three RQs.

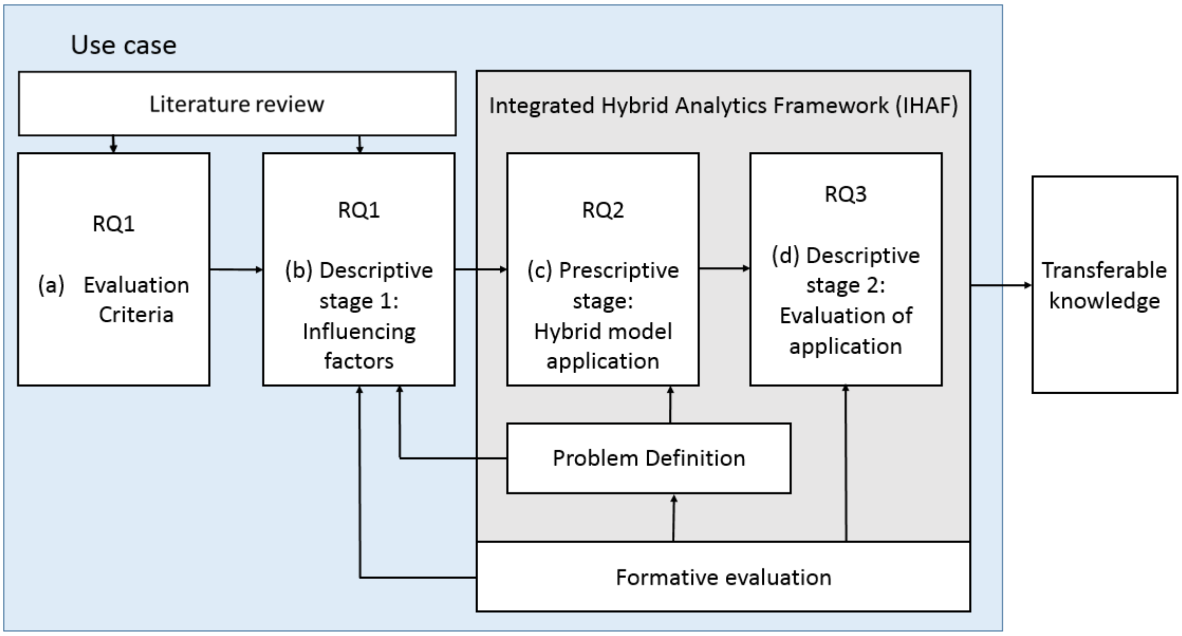


Figure 4-15 Alignment of Design Science, the proposed framework, and the methods

Phases (a) and (b) align with RQ1, and are addressed in the literature review in Chapter 2. Alongside the literature review, site visits, direct observations, and a formative evaluation will inform the development of the interview schedule for the summative evaluation, which forms part of IHAF.

Stage (c) is addressed with application of the IHAF framework (proposed in Chapter 4 and implemented in Chapters 5, 6 and 7), which includes the summative evaluation component, staff interviews (Chapter 7). Stage (d) takes this evaluation further, synthesising it with the formative evaluation, and the literature, to address RQ3 (Chapter 7).

4.10 Chapter Summary

This chapter advanced a generic HM framework, IHAF, for the development and execution of real-time HMs for short-term decision-support in sociotechnical systems, motivated by a healthcare application. This is accomplished through an examination of Design Science methodology, consideration of existing hybrid frameworks, and a review of the literature focusing the purpose of the study. This explicitly addresses the first objective of the Research Question 2, to propose a framework supporting an integrated hybrid approach for short-term decision-making in healthcare. The framework is developed to be generic, with transferable knowledge which aims to support similar future work in similar domains.

The model progresses through a sequence of stages. The problem definition stage is partly addressed with reference to the literature on situation awareness (SA) at individual and team levels. As the framework is positioned within the principles of QI, the problem definition stage is considered to require participative practice. SA and real-time simulation provide an overarching conceptualisation. Models of SA and closed-loop real-time simulation, though arising from different disciplines, conceptually align in their component parts. In SA models, SA is informed by environmental cues, with subsequent decision-making supporting action. This can change system behaviour, and via a feedback loop can inform ongoing SA. However feedback may be delayed. In human-in-the-loop real-time simulation models, real-time data initialises a simulation model, providing decision-support. Subsequent action changes the system, reflected in real-time data, providing faster and more accurate feedback than system observations.

The use of analytics to add value to real-time data forms the backbone of the framework. A HM decision-support tool which combines real-time data, predictions, and simulation has the potential to effectively support short-term decision-making. Hence, an architecture integrating descriptive, diagnostic, predictive and prescriptive methods has been proposed. The diagnostic component determines the conditions for triggering interactions between methods. The trigger may be reactive, or predictive, depending whether a predictive component is required, where the purpose of the simulation is prevention.

As the HM is intended for recurrent use, evaluation is necessary to provide knowledge toward improvements and iterations. This is an essential component of a Design Science methodology, and is necessary in real-world applications as a variety of factors affect a decision to implement a recurrent-use model, many outside of the control of the modeller.

The next chapter describes the application of IHAF, describing the use-case, and starting with the problem definition stage.

Chapter 5: Application of the Integrated Hybrid Analytics Framework in the use case NHS Trust

A. Define Problem

5.1 Introduction

This, and the subsequent two chapters, apply and test the Integrated Hybrid Analytics Framework (IHAF) proposed in the previous chapter in a use-case. This framework was developed and proposed in Chapter 4 as a real-time Hybrid Modelling (HM) framework for short-term decision-support in healthcare, with a particular focus on Emergency Departments (ED) (Figure 5.1; *Define Problem* component highlighted). This chapter introduces the use-case and the methods used to test the framework, and details the problem definition phase. The following two chapters (Chapter 6 and Chapter 7) apply the HM stages, and the evaluation stage respectively. Together, these three chapters address Research Question 2, to test and evaluate the potential of an integrated HM for short-term decision-support in healthcare combining real-time simulation with data analytics.

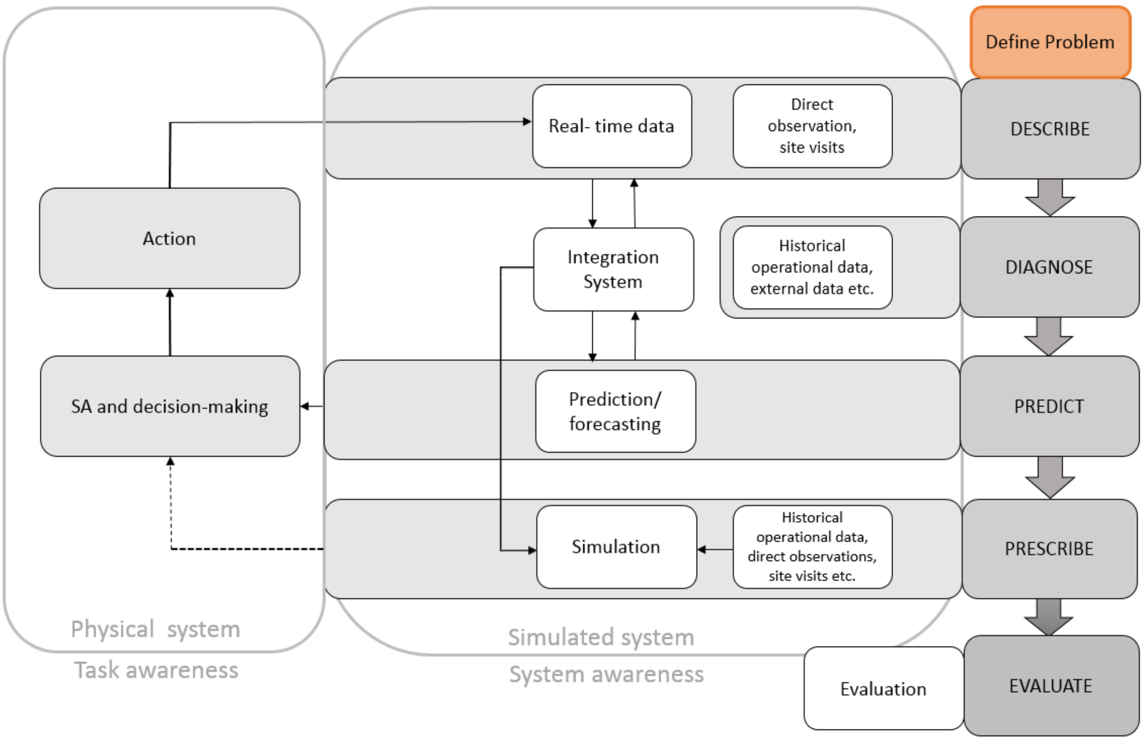


Figure 5-1 Integrated Hybrid Analytics Framework (IHAF) with problem definition stage highlighted

The HM is tested in a single case at a NHS Trust in the South-West of England, and its surrounding urgent care network. The rationale for the need for short-term decision support in ED has been explained in Chapters 1 and 2, making ED a suitable case for this study.

The aim of IHAF is to support the development and application of a HM which projects the progression of a situation over a short time period for short-term decision-support in an applied setting. Real-time predictions forecast the onset of a critical situation, and real-time simulation provides knowledge about recommended escalation actions to recover from the situation, supporting task- and system-level situation awareness (SA). Providing information derived from real-time or near real-time data seeks to reinforce environmental cues to support faster, and more accurate, decision-making. SA is influenced by both system-wide factors, such as workload and quality/availability of information, and by individual factors, such as experience. Salmon et al. (2008) illustrated how in a sociotechnical system, members' workload has both individual task work (task-level) and teamwork (system-level) elements. Information which contributes to both individual SA and to shared mental models at the system level is therefore required to support system processes. This provides an overarching conceptualisation toward approaching and defining the problem.

The use-case and real-time data will be described in Section 5.2. Following this, an overview of methods used to address the Research Questions, mapped to the Design Science methodology and IHAF, is clarified in Section 5.3. Section 5.4 addresses the first stage of IHAF, problem definition, which is achieved through direct observation and patient questionnaires. The rest of this chapter then presents the development, implementation and analysis of the patient questionnaire which contributes to the problem definition, and provides a formative evaluation (Nieveen & Folmer, 2013; Venable et al., 2016). This will be subsequently synthesised with the final evaluation in Chapter 7. The next section justifies the choice of the use-case.

5.2 The use-case: An NHS ED in the South-West of England

For this study, the use-case is a 'typical case' (Ridder, 2017), meaning it is representative of a broader set of cases. This is justified at two levels:

(i) ED as a case of healthcare: In contrast to many areas of healthcare, EDs function within a rapidly changing operational state requiring constant short-term decision-making. While other areas of healthcare may benefit from short-term operational decision support - such as GP services, or ward discharge planning - emergency services exemplify the requirement, as arrivals are unplanned and largely unpredictable in the short-term. Further, in common with many areas of healthcare, EDs are struggling operationally, and are well-placed to potentially benefit from additional information to support operational activities.

(ii) The use-case as a case of ED: The ED involved as a use-case operates with low physical visibility due to the fragmented layout, thus may benefit from additional information about crowding. It is located in a geographical area where staff and patients have access to a subset of real-time operational data through *NHSquicker* (Mustafee et al., 2017b) which is able to provide the real-time data for this study. Further, it is part of an urgent care network (UCN) consisting of one ED and three minor injury units (MIU), which are all spaced roughly equidistant geographically from ED. The NHS Trust involved has a prior history of innovation (Thistlethwaite, 2011) and thus has been open to improvement; nonetheless, a prolonged period of austerity inevitably inhibits innovation, as capacity is constrained, morale is reduced and funds are limited (Kelly & Young, 2017).

5.2.1 *NHSquicker: the use-case real-time data*

As discussed in Chapter 3, gaining access to healthcare data can present significant challenges for conducting applied studies in the healthcare domain, in particular where data is required in the form of live feeds. For this reason real-time data needs to be considered early in the study alongside problem definition if the application is to proceed.

For this study, data has been made available by *NHSquicker* (Mustafee et al., 2017b). *NHSquicker* (https://www.nhsquicker.co.uk/) was developed by the Health and Care IMPACT (Information, Modelling, Prediction and Evaluation to inform ACTion) Network (https://www.health-impact-network.info/) as a digital platform with the aim of shaping demand across urgent care networks. The IMPACT network, a collaboration between University of Exeter Business School and NHS organisations in the south-west of England, aims to facilitate collaborative working between universities and health and care organisations, and one of its stated missions is to create a thriving academic community for PhD students undertaking research in health and care. As a member of the IMPACT network, this research has benefitted from collaborating on projects such as *NHSquicker*, for example by gaining access to the data, and has also contributed to *NHSquicker* development, for example supporting elements of its evaluation.

One component of the *NHSquicker* platform is a mobile phone application which provides real-time wait-time data for patients across the southwest of England, with the aim of supporting attendance decisions for low-acuity patients. The NHS Trust use-case is one hospital which contributes real-time data to *NHSquicker* from its ED and three MIUs. These four facilities form an UCN. The data is available both historically and in near real-time, with three variables: (i) The total number of patients in each department (ii) The number of patients waiting to be assessed by a clinician (iii) The maximum wait time to be assessed by a clinician.

NHSquicker data has been validated by a number of NHS Trusts for its purpose (supporting attendance decisions for low-acuity patients), with jointly agreed data exchange standards. It provides data for most treatment centres in the UCN which can be utilised for decision support. This provided an opportunity to leverage *NHSquicker* data for real-time hybrid modelling.

Currently, patients and staff have access to real-time wait-time data via *NHSquicker*. It is proposed that both patients and staff will have access to real-time predictions, and that staff will have access to the real-time simulation. However at the time of conducting this research, no patients (as participants) were familiar with or had used *NHSquicker* for supporting attendance decisions.

5.2.2 *Research focus*

Inappropriate ED attendance for problems that are better suited to MIU, walk-in centres (WIC), general practice (GP), pharmacy or self-treatment can contribute to crowding. The resultant demand-capacity mismatch has wide-ranging impacts, associated with poor patient outcomes, longer hospital stays, poor patient experience, and a reduced staff morale (Bond et al., 2007; Morris et al., 2012; Sun et al., 2013; Boyle & Higginson, 2018; Morley et al., 2018; Higginson & Boyle, 2018; Abir et al., 2019). The definition and proportion of 'inappropriate' or 'non-urgent' attendance varies widely in the literature, for example a systematic review undertaken by Durand et al. (2011) found considerable variability in the proportions of visits deemed non-urgent, from 4.8% to 90%, with an overall median of 32%. The causes are multifactorial, with input, throughput and output factors all contributing, such that addressing the problem is a complex endeavour. Kluth et al. (2014) classified approaches to complex decision-making, with one approach categorised as 'focusing on individual factors'. This involves dividing the main complex problem into single smaller problems. We cannot expect to understand complex systems completely, and selecting system boundaries and scope are of vital importance in defining the problem area, and in the design of decision-support approaches (Daellenbach et al. 2012).

The focus of the study is low-acuity patients visiting ED. Low-acuity presentations were defined by Dinh et al. (2016) as those who self-presented (were not transported by ambulance), were assigned a triage category of 4 or 5 (low-urgency or non-urgent) and were discharged to their usual residence from ED. In this study, the first two criteria are used, as no information is available about the discharge destination of patients.

5.3 Methods overview

In Chapter 3, Table 3.1 illustrated how methods are used in this research to address the individual research questions and objectives. In the previous chapter, Figure 4.15 illustrated the alignment of the Research Questions (RQ), the Design Science methodology, and IHAF. This is updated in Figure 5.2 illustrating the alignment of the RQs, the phases of the Design Science methodology described in Chapters 3 and 4, the IHAF framework developed in Chapter 4, and the individual methods used to test the framework in the use-case.

Phases (a) and (b) address RQ1. Phase (a) has been defined using the literature, as discussed in Chapters 2 and 4, and determine the aim, focus and scope of the research project, which has informed the development of the framework. Phase (b) focuses the modelling process and its evaluation of factors which contribute to a plausible design, and are derived from the literature as theoretical propositions (Carlsson, 2006), and from other information sources, in this case site visits, direct observation (McKenney & Reeves, 2018), and questionnaires (Salehi & McMahon, 2009), as described in subsequent sections of this chapter. While the primary aim of this phase in the Design Science methodology is to determine the criteria for evaluation of the model, it is also used within IHAF to define the problem, and forms its first stage.

The prescriptive stage (c) involves building the model. This addresses RQ2 and applies IHAF. In the use-case, the real-time data comes from *NHSquicker*, the integrated real-time predictions use time-series forecasting, and the real-time simulation is a discrete-event simulation model (Chapter 6). Finally, the evaluation stage (d) uses staff interviews, and is the final component of IHAF (Chapter 7).

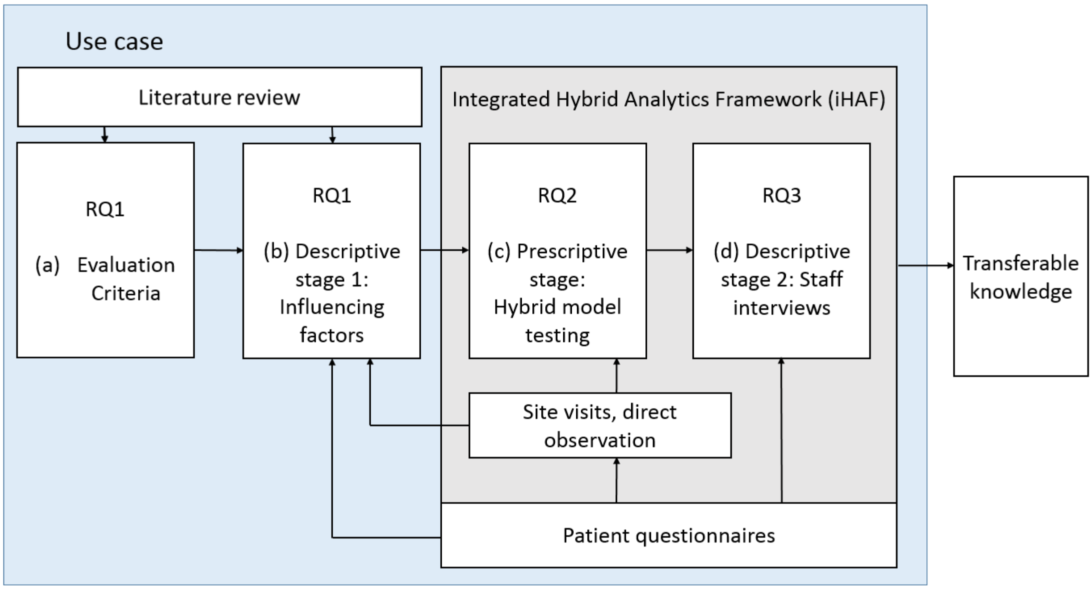


Figure 5-2 Relationship between research questions, DS methodology, IHAF, and methods in use-case

For RQ3, the staff interviews and patient questionnaires are synthesised with the literature to analyse the system level impact of the use of real-time data for short-term decision-support, and to evaluate the potential value and barriers to implementation of the HM. The following sections of this chapter focus on the problem definition phase of the IHAF application in its use-case.

5.4 IHAF: Problem definition stage

5.4.1 *Site visits and direct observations*

Defining the problem requires a triangulation of approaches, in this case the literature review, site visits and observational data. While the available real-time data does not define the problem, it is worth considering early, as access or restrictions may place limits on the application. Reviews of the literature have already been undertaken in Chapters 2 and 4 toward defining the problem using an overarching approach (to develop the framework) and with a more focused approach (for application to ED). Having selected a use-case, the specific problem to be addressed, and the objectives of the HM approach need to be clarified. A real-time HM starts with the assumption that it will be useful in practice. For this case study, field notes and observations contributed to understanding the problem, and the desire for solutions. Observation is a method of data collection which looks at people and places in their natural settings. Participant observation provides direct experiential and observational access to the social reality of people, involving not just observation but also listening. Observation is less disruptive and more unobtrusive than interviewing (Holloway & Wheeler, 1996).

Observations may be of four types: (i) The ‘complete participant’ takes an insider role and uses covert observation; (ii) The ‘participant as observer’ requires access permission and explains their observer roles to participants, but may also have a contributory role to play; (iii) The ‘observers as participant’ are only marginally involved in the situation, with no role to play in the setting apart from gathering data; (iv) The complete observer takes no part in the setting, and is a ‘fly on the wall’. Whichever approach is used, observation generally progresses from unfocussed and unstructured, to more focussed observations of specific actions and events (Creswell & Creswell, 2017).

Field issues in any type of field research can present challenges. When using qualitative methods, this can include maintaining a balance between objectivity and sensitivity, a process of reflectivity that requires openness, a willingness to listen and ‘give voice’, and representing multiple views as accurately as possible. Some of these issues are discussed at length by Corbin and Strauss (2014).

For this case study at an NHS ED, two types of participant observation were used:

(i) ‘Observer as participant’: for ED observation of processes, behaviours, data collection for developing the simulation model (Chapter 6)

(ii) ‘Participant as observer’: Workshops and events specifically aimed at understanding how real-time data is useful to patients and the NHS, as part of the co-development of *NHSquicker* (Appendix 2).

The contribution of the observational data (ii) to defining the problem is as follows:

1. 3rd IMPACT network event 21 June 2016, University of Exeter Business School. Thirty-nine NHS management, IT, clinical and communications staff, patient representatives and academics attended the one-day event. The purpose

was to focus aspects of the design of *NHSquicker* toward its subsequent development.

- There was general concern by managers and clinicians that more information is needed about how people with low-acuity conditions use ED and that this required engaging with patients about what mattered to them, for example, do they have misconceptions that ED can offer 'better' care? What do patients base attendance decisions on? Why do MIU attenders choose MIU?
- Academics and communications staff agreed that understanding how patients use services requires engaging with patients to find out what 'value' means to them.
- Clinicians were keen that assumptions were not made about who 'needed' to be seen in ED, and that providing information does not necessarily mean facilitating a decision, which may be multifactorial. Managers and communications staff emphasised that decisions about where to attend should involve patients.
- There was agreement that providing patient access to a subset of real-time data would: support joint working between providers across the urgent care network; empower, educate, and inform patients; improve resource utilisation across the network by spreading demand; reduce anxiety in patients; reduce waits for patients.
- Several participants mentioned that the NHS is often fearful of change or action, with various examples of concern, for example patients might make suboptimal decisions. An unintended consequence may be empty EDs, and MIUs underperforming against the 4-hour target.

2. 4th IMPACT network event Qualitative Systems Dynamics workshops (Powell & Bradford, 2000) 18 July 2018 and 27/28 June 2019 (NHS Hospital). These involved a manager, a clinician, several lay attenders, two academics (one an experienced facilitator). The purpose was to gather intelligence to confirm assumptions relating to demand and supply side factors contributing to crowding and the effects of real-time data. This was a data-gathering exercise for academic research.

- Patient anxiety was a central theme; real-time data was perceived to reduce patient anxiety by providing additional information and increasing confidence, dependent upon the actual and perceived accuracy of the real-time data.
- Previous experience of a service is perceived to influence risk aversion and anxiety, either positively or negatively.
- ED wait times are counterbalanced by more effective demand-capacity management across the urgent care network, including redirecting patients from ED to MIU.
- Patient anxiety and perceived urgency influence inappropriate attendance to ED.
- As inappropriate attendances increase, more patients will be redirected from ED to MIU. This occurs when ED is overcrowded.

All field notes are documented in Appendix 2.

The *problem statement* in the context of the use-case application is therefore that a proportion of low-acuity (non-urgent) patients who could safely be seen elsewhere in the Urgent Care Network (UCN) are attending ED, contributing to crowding. The focus of the use-case is investigating how real-time data analytic approaches might support short-term decision making toward the safe distribution of low-acuity patients across the UCN.

From the literature review and the observations, it was determined that more information was needed about what was important to patients when using services in the UCN, how they made attendance decisions, and how they might make use of real-time data to support health-seeking behaviour in the wider system. This was done using patient questionnaires. The overall process is outlined in the following subsection. Following this, the remainder of this chapter details the development, data collection, analysis and results of the patient questionnaires.

5.4.2 *Patient Questionnaires*

To support the problem definition stage of IHAF, it was determined necessary to understand how the end-users of the health system, the patients, use the real-time data currently available to them. It also determined the perceived usefulness of predicted data for both themselves, and for the NHS. This stage has a dual purpose:

- It supports a technology- or market-pull approach. By involving end-users, who are part of the system under investigation, an understanding of the current requirements and perceived value for patients, as end-users, can be considered in the design and function of a decision-support tool. This also supports the evaluation stage of IHAF by informing the interview schedule.
- It explores the implications and the added value to the system of using real-time data applications for patient decision-support, for later synthesis with the staff interviews (Chapter 7).

A significant body of research has uncovered a range of factors influencing the attendance of low-acuity patients to ED. The focus for this formative evaluation was therefore which of those factors could be influenced by real-time knowledge of wait-times, and the characteristics of patients who considered wait-time predictions to be useful for attendance decisions. It was determined that for this stage of the evaluation, responses needed to be representative and aggregated. This required a larger sample size than exploratory approaches such as semi-structured interviews or focus groups could provide. An on-site user questionnaire was chosen, as it has no interviewer effects, and is quicker and more convenient than conducting structured interviews. However due to the length of the questionnaire, a facilitated approach was required to alleviate some of the disadvantages of a self-complete questionnaire. For example, it was possible for patients to ask for clarification if required, it reduces the risk of missing data, and it increases the response rate (Bryman & Bell, 2007). Additionally, several open questions provided additional information.

5.4.2.1 *Questionnaire development*

Based on a purposeful review of the literature (Section 5.6) and direct observations (Section 5.4.1), the questionnaire was developed in three parts. The first section contained demographic data, the second section captured reasons for attendance, and the third section captured the perceptions of usefulness of the real-time data for attendance decision-making both now and in the future. The questionnaire consists of a series of brief, standardised response questions using a mixture of response-types, including open-ended questions. Multiple response questions were chosen for the main data collection rather than forced choice 'yes-no' questions. While the implicit assumption is that checked items correspond to 'yes' and unchecked items correspond to 'no', higher endorsement rates are observed for ‘yes-no’ questions, which were considered could confound results hypothesising a relationship between ‘reasons for attending ED’ and the outcome measures (Meyners & Castura, 2014; Callegaro et al., 2015). In order to capture the classifications derived from the literature (Section 5.6), a range of related ‘check-all’ questions were used. This is necessary because in multiple response items, interpretation of the unchecked box could be ‘no’, a missed entry, uncertainty, or not wanting to answer the question (Callegaro et al., 2015). A final section contained open questions about the perceived usefulness of real-time DA to patients and the NHS.

5.4.2.2 *Questionnaire setting and selection of participants*

The study was conducted across two NHS Trusts in Devon, with approximately 150,000 ED visits each year across the two sites. The purpose of using multi-sites in this stage was to increase the generalisation of the results. Participants aimed to be a representative sample of low-acuity patients seeking urgent health care. *Low acuity* was defined by the triage nurses’ assignment of Triage Category 4 or 5, and that patients walked-in, that is, were not transported by ambulance. No assumptions were made as to whether the visit was ‘appropriate’ based on this categorisation. All questionnaires for participants under 16 years were completed by a parent or carer. Questionnaires were administered under NHS honorary contracts, with university ethical approval and signed consent from all participants, parents or carers prior to participation. Questionnaires and consent forms were given in paper format, and were facilitated by a researcher who remained with the participant.

5.4.2.3 *Questionnaire design and validation*

The questionnaire was validated for face validity by NHS staff and piloted on 50 patients (Figure 5.3). Piloting ensured that no vague, complex or ambiguous questions, double-barrelled questions, technical jargon or formatting errors were included (Choi & Pak, 2005). This was a cross-sectional facilitated questionnaire-based study on patients from a defined catchment area. A convenience sample of low-acuity patients were recruited whilst waiting for care in ED, MIU and WIC waiting areas. Inclusion criteria were patients seeking urgent care classified as Triage Category 4 or 5, with either the ability to consent or accompanied by an adult who could consent on their behalf. Patients who declined to participate were excluded. The sample was recruited across four shifts in proportion with historical arrival patterns. The sample was cross-checked with attendance data to ensure it was representative for age, gender and time of day. Questionnaires were facilitated to ensure a high return: 152 questionnaires were returned, a response rate of 94%.

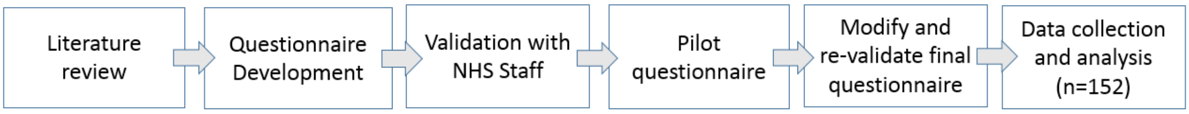


Figure 5-3 Questionnaire development process

5.4.2.4 *Questionnaire analysis*

Descriptive statistics were used to examine the demographic characteristics and survey results of the sample. Statistical analysis was performed using chi-square for categorical data for comparison of proportions, and Mann-Whitney U test for interval data for comparison of means. Statistical significance was set at $p<0.05$ and all analyses were performed using SPSS 24.0. Open questions were coded and analysed thematically using NVivo Pro 12 against the closed question responses.

The questionnaire development and results are presented in the following sections. The results provide an understanding of the current requirements and perceived value of real-time data and predictions for patients, as end-users, and can be considered in the design and function of the HM, and support the final evaluation of the model.

5.5 Patient Questionnaire Development: Closed questions

5.5.1 *Introduction*

The focus of the study is low-acuity patients visiting the ED, defined in section 5.4.2.2. This is a problem that can be viewed both from the patient (demand) side and the NHS (supply) side.

As outlined in Chapter 4, defining the problem can require a triangulation of approaches. Through field notes and observations, it was found that clinicians, managers and other NHS staff expressed concern that there is little understanding of why people with low-acuity conditions use ED and on what they base their attendance decisions. Staff considered that real-time data analytic (DA) applications might be useful to patients by reducing anxiety, and changing attendance decisions, and that the effects of this would benefit both patients and the wider system. However patient views were unknown, therefore the questionnaires aimed to determine the characteristics of patients attending ED, and the expected value to patients in using real-time DA applications. At the system level, the DA applications are expected to support ED crowding in two ways:

(i) At the user-end by providing patients with new information that can support attendance decisions (closed questions, Section 5.8).
(ii) At the provider-end through decision-support, for example about redirecting appropriate patients to alternative services before queues become unmanageable (open questions, Section 5.9). Patients were asked how the NHS might benefit from this information in planning and delivery of services.

5.5.2 *Focus of questionnaire*

To support urgent and emergency care attendance decisions, it is important to understand why low-acuity patients attend ED, and their perceptions regarding the value of real-time DA for supporting attendance choices. The definition and proportion of 'inappropriate' or 'non-urgent' attendance varies widely in the literature, for example Weinick et al. (2010) estimated between 14 and 27% non-urgent attendance; Mason et al. (2017) found 23% of adults and 31% of children were non-urgent attendees; meanwhile Hsia & Niedzwiecki (2017), using a conservative definition, found 3.3% avoidable ED attendances. A systematic review undertaken by Durand et al (2011) found 51 different methods to categorise visits to ED into non-urgent and urgent cases. Many focussed on the main complaint, the duration of the complaint, vital signs and the need for diagnostic tests or treatments performed in ED. The most common categorisation focussed on the delay until seeking treatment. The authors also found considerable variability in the proportions of visits deemed non-urgent in the literature, from 4.8% to 90%, with an overall median of 32% (Durand et al, 2011). A House of Commons Library briefing paper (Baker, 2017) reported that almost 38% of ED attendances resulted in guidance or advice, and a further 11% resulted in no treatment - totalling almost half of recorded attendees. NHS Digital (2019a) defines unnecessary attendance as the "First attendance with some recorded treatments or investigations all of which may have been reasonably provided by a GP, followed by discharge home or to GP care", with a national average of 10.8%; however this doesn't account for those patients who may have more been appropriately seen at a MIU or other facilities.

This low level of agreement highlights the lack of reliability of these categorisation methods, and reflects the complexity of the issue. For example the classification of 'inappropriate' is done in retrospect, and many of these patients have been referred to ED by an alternative healthcare provider based on described symptoms, without examination or investigation. Moreover, the majority of patients who self-refer to ED consider their attendance appropriate, based on their own assessment of their condition, and knowledge about existing resources. In any case, not all low-acuity patients are 'inappropriate'.

A significant amount of research has explored the decision factors that contribute to patients with low-acuity medical problems using the ED rather than an alternative care facility appropriate for treating minor conditions. Many studies have suggested that patient self-triage and decision-making regarding ED attendance is reasonable and appropriate based on the information available with which to make a decision (e.g. Nelson, 2011; Land & Meredith, 2013; Breen & McCann; 2013; Chapman & Turnbull, 2016; Cheek et al., 2016; Krebs et al., 2017; Weber, Hirst & Marsh, 2017). How real-time DA, as additional information, can affect attendance decisions is unknown, and this is the focus of the first part of the questionnaire, which will subsequently inform the evaluation of the HM. The next section proposes the use of a conceptual model for supporting analysis of the questionnaire data, Andersen's Behavioural Model, which is used to both explain and predict healthcare utilisation, and health-related outcomes such as patient satisfaction, through a set of determinants of health-seeking behaviour. Following this, a set of known factors which drive ED attendance decisions in patients with low-urgency conditions is derived from the literature (Section 5.6).

5.5.3 *Conceptual framework for analysis: Andersen's Behavioural Model of Health Service Use*

Healthcare utilisation is the point in health systems where the needs of patients and the service provided by the system meet. Utilisation is strongly dependent upon both the structure of the system and need-related factors (Babitsch et al., 2012). The Behavioural Model of Health (Andersen, 1995) describes three categories of individual and contextual determinants of health-seeking behaviour:

(i) *predisposing characteristics* such as demographic and social factors, psychological factors including health beliefs and individual resilience and trust in/familiarity with the medical organisation; (ii) *enabling resources* such as family/social influences, availability of medical resources, access to care, social support, and convenience of services including organisational factors; and (iii) the *perceived need* to the individual, such as evaluated health status, self-reported health, psychological distress and anxiety (Andersen, 1995; Babitsch et al., 2012).

Andersen's model provides an appropriate conceptual basis for understanding ED utilisation because it considers both human attitudes, such as beliefs about the healthcare system, and health-seeking behaviour. It demonstrates the complex interaction between factors which enable or impede utilisation of a health service, an individual's predisposition to use a service, and their need for care.

The model was originally developed by Andersen (1968) and has been expanded through numerous iterations. Andersen (1995) extended the model past the use of services to end at health outcomes, and included feedback loops to illustrate that health outcomes may affect aspects such as health beliefs and behaviour. The sixth revision (Andersen, Rice & Kominski, 2011) expanded further to include quality of life as an outcome and emphasises contextual as well as individual determinants of access to medical care (Figure 5.4). Dimensions of utilisation are defined according to components of the framework. Contextual characteristics are the circumstances and environment of health care access, including health organization and provider-related factors as well as community and family characteristics, while individual characteristics belong to individuals.

The arrows leading from the contextual characteristics indicate how they can influence health behaviours and outcomes in multiple ways. They can work through individual characteristics, for example low distribution of MIUs may lead to increased use of ED by low-income persons who don't own a car, while contextual characteristics can also influence health behaviours and outcomes directly, for example poor GP accessibility and continuity can result in increased use of ED, independent of individual characteristics (van den Berg et al., 2016).

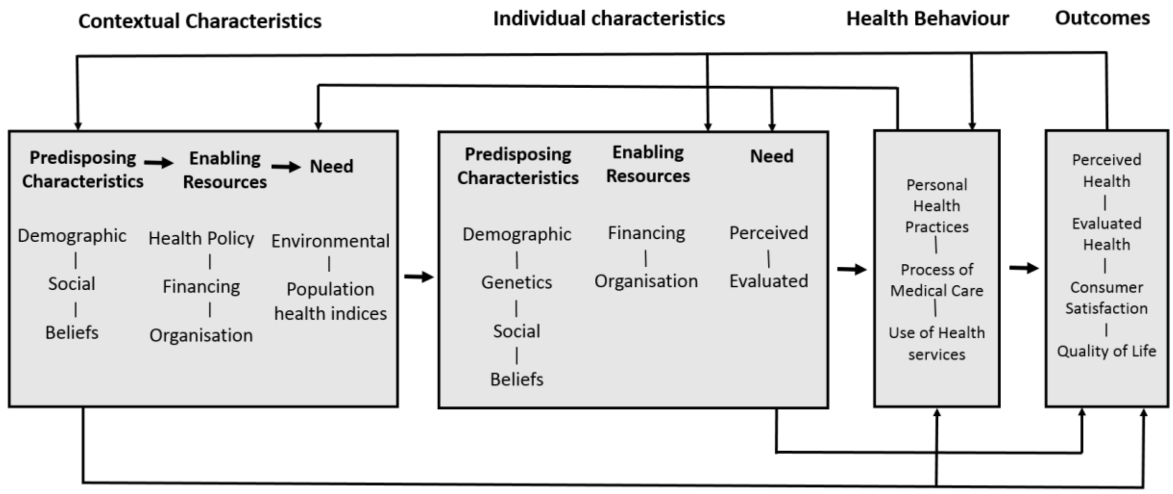


Figure 5-4 The Behavioural Model of Health, reproduced from Andersen (2013)

The model suggests that major components of contextual and individual characteristics determine utilisation. These are:

(i) Contextual predisposing characteristics, which are existing conditions that predispose people to use or not use services: demographic factors, (e.g. age, gender, marital status); social factors (e.g. educational level, ethnic and racial composition, measures of spatial segregation, employment level) and beliefs (e.g. community or organisational values, cultural norms, political viewpoints regarding how health services should be organized, financed, and made accessible to the population).

(ii) Contextual enabling conditions that facilitate or impede use of services, including public policies at all levels from local to national, financing characteristics including socioeconomic levels, the method of compensating providers; organisational factors include the number and distribution of services, staffing, structure in the community, resources, opening hours, and facilities (Andersen, 2008).

(iii) Contextual need or conditions that lay-people or health care providers recognize as requiring medical treatment (Andersen, 1995; Andersen, Davidson & Baumeister, 2014), quality of housing, rate of disease and injury. Population health indices are more general indicators of community health such as mortality and morbidity rates.

Individual characteristics which determine utilisation include:

(i) Individual predisposing characteristics, such as age, gender, and genetic factors can determine individual need for care. Social factors determine the status of a person in the community and their ability to cope with presenting problems, including education, occupation, and ethnicity. Social networks, such as presence of family and friends can facilitate or impede utilisation. Health beliefs are attitudes, values, and knowledge people have about health and health services that can influence their perception of need for health services (Andersen, 1968; Bradley et al., 2002).

(ii) Individual enabling characteristics enable access, for instance income is relevant even where healthcare is nationalised, for example the cost of parking or public transport, or taking time off work. Social support, for example emotional or practical help can be an enabler. Organisation of health services describes, for instance, whether the patient is registered with a GP (Blackwell et al., 2009). It also includes means of transportation, reported travel time, and waiting time for care (Andersen et al., 2014).

(iii) Individual need characteristics: Perceived need is how people view their own general health, perceptions of the severity of their presenting complaint, and how they respond to symptoms of illness, pain and anxiety (Afilalo et al., 2004; Hoot & Aronsky, 2008). Evaluated need represents professional judgment and objective measurement about a patient's physical status and need for medical care, which can also have a social component (Andersen, 1995; Blackwell et al., 2010; Andersen et al., 2014). Quality of life has been recently added to the model to reflect an increased focus on patient-centred care.

Health behaviours are personal behaviours that influence health status, such as diet or exercise. The process of medical care is the behaviour of providers interacting with patients. Personal health behaviours interact with use of health services to influence health outcomes. The patient's perceived health status is influenced by health behaviour, personal health services use, and individual characteristics, as well as the contextual environment (Hoot & Aronsky, 2008). Evaluated health status is dependent on the judgment of the professional. Consumer satisfaction is how individuals *perceive* the health care they receive, and can be judged by patient ratings of travel time, waiting time, communication with providers, and technical care received. Central to the model is feedback, depicted by the arrows from outcomes to health behaviours, individual characteristics, and contextual characteristics. Feedback allows insights about how access might be improved. Feedback can occur at the national level, or at the regional or Trust level, resulting in contextual changes in the organisation and processes of care for patients (Babitsch et al., 2012, Andersen et al., 2014).

The model emphasises the dynamic and iterative nature of health service use, such that the outcome of using a service in turn affects subsequent health behaviour (Andersen, 1995). By using the framework's relationships it is possible to determine the directionality of the effect following a change in an individual's characteristics or environment. For example, if one experiences an increase in need as a result of a minor injury, Andersen's model predicts this will lead to an increased use of services, where all else remains equal. It is a useful way of conceptualising a range of interacting factors known to contribute to health seeking behaviour.

An important concept within the model is the concept of *mutability.* This is a determination of the degree to which each factor can be altered, and therefore can potentially influence behavioural change. Policies are implied first by determining what variables explain utilisation, and this is achieved using a structured literature review (Section 5.6). To be useful for promoting or controlling access, a variable must also be considered mutable, or point to policy changes that might bring about behavioural change. If a factor has a high degree of mutability (i.e. can be easily changed) a policy would be justified in using its resources to address this intervention, rather than a factor with low mutability. For example social structure is judged as being of low mutability, as ethnicity is not changeable, and altering educational or occupational structures is not a viable short-term policy to influence utilisation. Health beliefs are judged as having medium mutability since they can be altered and can affect behavioural change. Characteristics that fall under demographics are difficult to change, however many *enabling resources* are assigned a high degree of mutability as individual, community, or national policy can take steps to alter the level of enabling resources for an individual. For example, the UK government has made sustained attempts to clarify the difference between 'emergency' and 'urgent' care to ensure patients seek 'the right care, in the right place, whenever they need it.' With the aim of reducing pressure on ED, NHS111 provides access to urgent care advice, while the roll-out of urgent care centres for convenient, local treatment of urgent conditions aims to improve access (NHS England, 2020a). At the local level, Trusts communicate availability and appropriateness of services, facilities within those services, and opening hours. At the individual level, patients can educate themselves about services in order to make urgent-care attendance decisions. Adequate communication by Trusts and other services to collate this information makes this a straightforward task for individuals, and should reduce uncertainty about where best to attend for treatment, directly influencing perceived need.

Andersen (1995) explained that 'need' encompasses not just actual (evaluated) need, but also perceived need for care, which may be increased or decreased through health education, incentives to use particular services, and other factors such as anxiety. The degree to which real-time DA impacts on the perceived need for care in ED is the purpose of the questionnaire, and will ultimately inform RQ3, the evaluation of the system level impact of the use of real-time DA for both patient and staff decision-support. This is an exploratory questionnaire, as no participants actually used real-time information to inform their attendance decisions, hence how 'perceived need' translates into health behaviours and outcomes needs to be addressed through future work. The next section reviews the health-related literature to understand reasons for low-acuity ED attendance. These will inform the development of the set of closed questions, and subsequent analysis according to Andersen's conceptual model.

5.6 Literature Review: closed questions

A literature review was undertaken to collate a set of known factors which drive ED attendance decisions in patients with low-acuity conditions. CINAHL (Cumulative Index to Nursing and Allied Health Literature) is an index of English-language and selected other-language journal articles about nursing, allied health, biomedicine and healthcare. Preliminary searches for papers aiming to categorise reasons for low-acuity (non-urgent) patients attending ED indicated that they were mainly published in medical (particularly emergency care) or medical policy journals. MEDLINE, a comprehensive database, focuses on biomedical literature and returned thousands of clinical and biomedical papers. In Web of Science, which includes MEDLINE, the same search criteria limited to 'Health Policy Services' found relatively few relevant papers. CINAHL was chosen as a small database which indexes a high number of medical journals, and returned a large proportion of relevant papers. Other databases are likely to return further relevant papers, however it was judged that the selection retrieved had provided sufficient information to inform the questionnaire.

The search was undertaken in the following way (Table 5.1):

Table 5-1 Search strategy for questionnaire development

Database(s)	CINAHL (EBSCOhost)
Content Search	Abstract
Search criteria	("Emergency department" OR "emergency room" OR "accident and emergency" OR "accident & emergency" OR a&e OR "a & e") AND (non-urgent OR "low acuity" OR inappropriate)
Limit(s)	Full text results, Academic Journals, 2008-2018 (full years), English language
Results	730
Filtered by title/abstract	31

The search returned 730 papers. They were filtered by title, then abstract, for papers that specifically sought to understand and explore factors, reasons and characteristics associated with ED attendance by patients with low-acuity conditions. Papers were included where a subset of characteristics were examined, for example reasons for using ED rather than primary care services (Nelson, 2011; Shaw et al., 2013), or where participants have specific inclusion criteria, such as age (Rowe et al., 2015).

A full range of factors were sought, so papers were not restricted to the UK, although the search period was limited to the last decade. The reason for this is that in the UK, changes in the structure of health services have occurred over time, for example GP-led WICs were introduced between 2007-2010 in the UK to lower the barriers to accessing primary care, and a 48-hour target for GP appointments was introduced in 2008 (Monitor, 2014).

The majority of studies used patients as participants to determine decision factors, however a small number of papers whose respondents were staff were also included (e.g. Breen & McCann, 2013; Chapman & Turnbull, 2016). Most papers were published in medical journals (e.g. American Journal of Emergency Medicine, BMJ, Internal and Emergency Medicine, International Emergency Nursing), health policy journals (e.g. Health Policy, International Journal for Quality in Health Care), while two were commissioned reports (Rowe et al. 2015; Mason et al. 2017). The majority of the studies used cross-sectional questionnaires, while eight used qualitative methods (e.g. Backman et al., 2008; Shaw et al., 2013; Chapman & Turnbull, 2016; Beache & Guell, 2016) and one was a longitudinal design (Mason et al., 2017).

During analysis of the papers starting from the most current, saturation of new decision factors occurred early in the process. This means that starting from 2018 and working backwards by date of publication, all decision factors were identified by 2015. This supported the assumption that the resultant decision factors identified from the search constituted a complete or near-complete set. Figure 5.2 illustrates the results of this process. Decision-making reasons were identified (columns) by reading each of the papers, and a red dot is placed against each factor where identified (lines).

Totals are included in the table to provide an indication of the frequency of each of the factors, however it is important to note that the use of one database only (CINAHL), and the wide inclusion criteria mean that these totals are indicative only. Data about decision-factors are gathered using a variety of methods within the studies, and at times are subject to interpretation, for example the interviews by Backman et al. (2008) found that anxiety, feeling ‘disturbed about symptoms’, and a history of hospitalisation are the main factors discriminating between patients seeking healthcare at an ED and those attending a primary care centre. In Figure 5.2, these findings are classified as ‘anxiety’, ‘my condition is urgent’, and ‘I may need to go to hospital’.

The questionnaire (closed-questions) aimed to understand patients’ reasons for attending ED, and to identify whether real-time DA presents a possible solution to the challenges associated with uneven demand. The specific objectives of this part of the questionnaire were to understand patient awareness of, access to, and use of urgent care services locally, how referral and advice impacted on patient decision-making, and levels of patient satisfaction with using urgent care services. By categorising the mutability of these decision-making reasons, it is possible to determine which of these are potentially mutable to real-time DA, and may subsequently result in behavioural change.

The individual questions are first categorised by similarity according to how they were used in the literature. This is done in three ways:

(i) The validated questionnaires developed by Coleman et al. (2001), Penson et al. (2012) and Mason et al. (2017) informed the questions, categorisations, and mutability allocations. These authors grouped their questions by similarity and categorised them according to the 'strength of attendance reason' to describe the degree to which categories of attendance reasons might be amenable to change.

(ii) The objective is to explore how real-time DA might influence individual characteristics, according to Andersen's concept of *mutability* of determinants of health seeking behaviour. Descriptions of the characteristics (Andersen, 1995; Andersen et al., 2014) were examined to determine where each of the categories of questions could be mapped to Andersen's model, and his descriptions of mutability, with a view to drawing some general conclusions about the potential impact of the real-time DA for patients. These validated the categorisations from (i).

(iii) The categories were further validated in this context in consultation with two NHS clinicians (one surgical consultant, one senior ED nurse) and one NHS manager (quality improvement) to support content validity, i.e. whether the factors adequately measure the categories (Tsang et al., 2017).

To summarise the above process: A search was undertaken in the CINAHL database for papers investigating low-acuity ED attendance. From each paper, a list of factors influencing attendance was identified from the empirical findings. These factors were then grouped by similarity, and allocated a mutability ranking, as per the work of Coleman et al. (2001), Penson et al. (2012) and Mason et al. (2017). These were validated in consultation with NHS staff. Figure 5.2 tabulates the papers, the individual decision-factors, and the categories of decision-factors identified via the above process. The following sections describers the categories and mutability in more detail, which subsequently inform the questionnaire development.

Table 5-2 Reasons for ED attendance by patients with low-acuity conditions, and categorisations

Categories		Ease of access				Uncertainty re alternatives					Wait times					Past experience		Anxiety					Perceived Seriousness								Advise to attend			Other	
Authors	Date	Proximity	Travel/transport	More convenient	Easier to get here	Unaware of Alternatives	Unaware find alternatives	Uncertain if alternatives approp	Couldn't find info online re althernatives	Unsure if alternatives open	Will be seen more quickly here	Open 24 hours	Don't want to wait for GP	Don't have a GP	Don't want to see GP	Past experience	Past satisfaction	Anxiety	Raised patient expectation	Reassurance re severity	Condition is urgent	Want to see doctor asap	May need to go to hospital	Confidence in this service	Another service would refer me here anyway	condition is emergency	Want to see a specialist	Might need an Xray	Most appropriate place	Percieved seriousness	Advised by HCP	Advised to attend by family/friend	My child's condition is urgent	Instinct/knowledge	Searched symptoms online
Lega & Mengoni	2008												●		●			●			●	●							●	●					
Hau et al.	2008				●								●					●																	
Backman et al.	2008																	●			●		●												
Brim	2008	●	●	●	●	●							●		●																				
Williams et al.	2008														●	●													●	●	●		●		
McGuigan & Watson	2010	●				●								●							●				●			●	●	●	●	●			
Backman et al.	2010															●		●			●														
Tsai et al.	2010	●		●	●																														
Nelson	2011			●								●	●								●							●	●		●	●			
Penson et al.	2012	●			●						●		●	●	●		●			●		●	●	●	●	●	●	●			●				
Knowles et al.	2012					●	●	●																							●				
Agarwal et al.	2012	●			●							●	●		●		●	●						●			●		●	●					
Shaw et al.	2013	●	●	●	●	●							●		●	●	●				●					●				●	●				
Breen & McCann	2013					●							●		●					●	●		●					●	●	●					
Nagree et al.	2013			●																															
Land & Meredith	2013	●			●	●					●		●	●		●	●	●		●	●	●			●	●		●		●					
Alyasin & Douglas	2014			●								●	●	●		●	●				●	●						●							
Beache & Guell	2015			●																											●	●			
Atenstaedt & Evans	2015				●		●	●	●	●		●																	●				●		●
Rowe et al.	2015					●	●	●			●		●		●			●		●	●	●	●	●		●	●		●	●		●	●	●	●
Chapman & Turnbull	2016			●		●	●	●		●					●			●	●	●	●	●		●			●		●				●		
Cheek et al.	2016	●		●	●						●		●	●	●	●	●		●					●	●				●		●				
Unwin et al.	2016			●	●			●								●	●									●		●			●				
Liscott	2016	●		●		●	●	●				●	●			●	●	●			●											●			
Krebs et al.	2017																												●		●				
Mason et al.	2017					●	●	●		●		●	●	●	●		●	●	●		●	●		●		●					●	●			
Idil et al.	2018	●		●		●					●	●					●				●														
Pearson et al.	2018		●															●			●					●				●					
Sancton et al.	2018																												●	●					
Dixe et al.	2018										●																		●	●					
Andrews & Kass	2018				●	●															●									●	●	●			●
TOTAL		10	3	12	11	12	6	7	1	3	6	7	14	6	11	8	10	11	3	5	16	7	4	6	4	7	4	7	13	12	12	7	4	1	3

The categories are described in the next section, alongside the individual decision-factors that contribute to each, and the allocated mutability. Each of these can be mapped to Andersen's conceptual model as 'Individual Characteristics'. Contextual characteristics are considered to be fixed. In Andersen's model, a combination of contextual and individual characteristics result in health-seeking behaviour and health outcomes.

5.7 Variable classification description

1. *Ease of access: Medium mutability*

Several studies found that convenience, transportation barriers and location of urgent care facilities contributed to attendance decisions in 35-63% of patients (Penson et al., 2012; Shaw et al., 2013; Land & Meredith, 2013; Alaysin & Douglas, 2014; Cheek et al., 2016; Idil et al., 2018; Andrews & Kass, 2018). Nelson (2011) found that most of their participants - low-acuity patients visiting ED - lived within a 15-mile radius. Non-urgent patients are reported to attend for convenience of location, despite being aware that their condition is not urgent or serious, and could wait (Chapman & Turnbull 2016; Mason et al., 2017). In Andersen's model (2013), this is considered to be an *Individual Enabling Resource (Organisation)*, as convenience, proximity, and travel options enable access and utilisation. It's clear that proximity and convenience is a factor in decision-making, and is rated as 'medium' strength by Mason et al. (2017), Penson et al. (2012) and Coleman et al. (2001) suggesting it may be amenable to change.

2. *Uncertainty about alternative facilities: High mutability*

While most people are aware of the location of EDs and its 24 hour access, many are unsure of opening hours, location and facilities of alternative services (Penson et al., 2012; Atenstaedt & Evans, 2015; Rowe et al., 2015). Even patients who do not consider their condition to be serious will attend ED if they do not know where else to go, although these are relatively low rates of 5-20% (Penson et al., 2012; Land & Meredith, 2013; Unwin et al., 2016; Mason et al., 2017). Correspondingly Chapman and Turnbull (2016) reported that there is also a lack of awareness among healthcare professionals about which conditions are most appropriate for which services, where services are located, the facilities they provide, and the times they are available. Nonetheless, educating patients about alternative services is considered to be a priority by staff (Worthington et al., 2005; Brim, 2008 Breen & McCann, 2013; Chapman & Turnbull, 2016). In Andersen's model (1995), this is considered to be an *Individual Enabling Resource (Organisation)*, as knowledge of alternatives enables access and utilisation. These are considered to be weak attendance reasons, highly mutable to change (Coleman et al., 2001; Penson et al., 2012; Mason et al., 2017).

3. *Wait times: Medium mutability*

A proportion of people believe that they will be seen more quickly at ED than other services (Penson et al., 2012; Land & Meredith, 2013). In some cases, patients prefer the convenience of a same day 'quick-fix', even if they believe it involves a wait, than to wait a few days for a GP appointment (Tsai et al., 2010; Nelson, 2011; Penson et al., 2012; Agarwal et al., 2012; Alyasin & Douglas, 2014; Chapman & Turnbull, 2016; Mason et al., 2017; Andrews & Kass, 2018). Convenience of one-stop facilities, the belief by patients that they'll be seen more quickly in ED, and not wanting to wait for a GP appointment are all classified as 'medium' strength reasons (Coleman et al., 2001; Penson et al., 2012). As patients currently have no knowledge of actual wait-times, in Andersen's model, wait times can be classified as an *Individual Predisposing Characteristic (Beliefs)*, representing a patient's attitudes or knowledge about health services. The provision of actual wait times will change this to an *Individual Enabling Resource (Organisation)*, as knowledge of wait times enables a more informed choice.

4. *Past experience: Low mutability*

Treatment-seeking behaviour is repetitive and reinforcing, such that past experience will influence attendance behaviour (Liscott, 2016). Many studies focus on previous satisfaction with the efficient delivery of care, availability of tests such as X-ray, and perceived expertise and specialty care available in ED (Land & Meredith, 2013; Unwin et al., 2016). Agarwal et al. (2012) found that patients or carers who were anxious about presenting conditions sought reassurance by turning to services with which they were familiar. Cheek et al. (2016) found that 16% of low-acuity patients attended ED out of habit or convention. Past experience is a *Predisposing Characteristic (Beliefs),* representing knowledge, attitudes or values people have about health services that can influence their perception or need for care (Andersen et al., 2014). Previous experience is rated as a 'strong' attendance reason by Mason et al. (2017), Penson et al. (2012), and Colemen et al.(2001), suggesting it is resistant to change.

5. *Anxiety: Medium mutability*

Anxiety is a characteristic of patients in ED which can lead to the belief that their condition is more serious and/or urgent that it is in reality (Shaw et al., 2013; Land & Meredith, 2013). Higher levels of anxiety, and stronger beliefs that conditions are more serious or urgent than that attributed by a triage nurse are found in non-urgent ED patients compared to GP patients (Lega & Mengoni, 2008; Backman et al., 2008; Agarwal et al., 2012). Anxiety appears to strongly influence decision-making (Pearson et al., 2018). Mason et al. (2017) found that people, particularly those of younger generations, have become more demanding of the healthcare system, and that social media and the internet can exacerbate anxiety with self-diagnoses. Anxiety is associated with a need for rapid reassurance, with 30-50% of low-acuity patients indicating that they needed to see a doctor as soon as possible (Penson et al., 2012; Land & Meredith, 2013; Idil et al., 2018), an issue which is on the rise (Mason et al., 2017). Most patients report being in the most appropriate place, but this is often due to misperceptions that they would be seeing a clinician who is more qualified than their GPs (McGuigan & Watson, 2010; Atenstaedt & Evans, 2015). In Andersen's model, anxiety can be viewed as multi-faceted. *Individual Predisposing Characteristics (Social)* include normative health beliefs associated with, for example, education, occupation and ethnicity, and *(Beliefs)* are attitudes, values and knowledge that can lead to perception of need for health services. Both of these predisposing characteristics can affect anxiety, and influence the perceived need for services, that is, how patients respond to symptoms of illness, pain and anxiety (Andersen et al., 2014). This factor is considered to be of medium mutability (Coleman et al., 2001; Penson et al., 2012).

6. *Perceived severity: Low mutability*

Shaw et al. (2013) found that low-acuity patients with pain or perceived need for investigations, even where aware of alternative providers, will choose to attend ED. Penson et al. (2012) estimated that two-thirds of low-acuity patients could have been managed outside the ED, despite strong beliefs that they were in the most appropriate place for treatment. They concluded that there is a low overall likelihood of behaviour change in patients for whom attendance reasons include wanting to see a specialist, the perception that their condition is an emergency, a concern that they may need to go to hospital, or having been referred or recommended to attend by a health care professional. People seek help from the place that has the facilities they perceive they need (Land & Meredith, 2013).The frustration of wait times is offset by the benefits of perceived staff expertise (Atenstaedt & Evans, 2015), access to X-rays or other tests (Lega & Mengoni, 2008; McGuigan & Watson, 2010; Penson et al., 2012; Cheek et al., 2016; Unwin et al., 2016), or other facilities (Land & Meredith, 2013; Unwin et al., 2016; Andrews & Kass, 2018). Parents of under 5's are more likely to be confident in ED as they perceive it to be more consistent, specialist, up-to-date and thorough than other services (Rowe et al., 2015). Perceived severity is an *Individual Need Characteristics (Perceived)* which represents how people view their own general health, perceptions of the severity of their presenting complaint, and how they respond to symptoms (Andersen et al., 2014). Mason et al. (2017), Penson et al. (2012) and Coleman et al. (2001) consider these to be 'strong' attendance reasons, therefore of low mutability.

7. *Advised to attend by Health Care Professional: Low mutability*

A common theme in the literature is the issue of alternative services referring non-urgent patients to ED due to risk averse behaviour (Mason et al., 2017) or lack of capacity within their own service (Chapman & Turnbull, 2016). Patients who have been referred to ED by a healthcare professional (HCP) are very likely to consider themselves to be in the most appropriate place (Penson et al., 2012), making this category of low mutability. Referred patients range from 30 to 52% in the literature, and are increasing over time, particularly from GPs who are under-capacity (Penson et al., 2012; Unwin et al., 2016; Mason et al., 2017; Krebs et al. 2017). In the UK, this is a growing problem, frequently presented as a 'crisis' (Marchand et al., 2017). This aligns with *Individual Need Characteristic (Evaluated)* in Andersen's model. Although there may be a range of reasons patients are referred to an alternative service, including social reasons, patients are likely to view this as representing professional judgment and an objective measurement about their physical status and need for medical care (Andersen, 1995). It therefore may have a perceptual element. Mason et al. (2017), Penson et al. (2012) and Coleman et al. (2001) consider these to be ‘strong’ attendance reasons, therefore of low mutability.

8. *Advised to attend by family/friend: Medium mutability*

Interpersonal factors are a driver of attendance decisions, with carers and relatives influencing a reduced tolerance for risk both for ED attendance and for self-management (Liscott, 2016). Social networks seem to sanction risk averse behaviour (McGuigan & Watson, 2010), with roughly 15% of patients attending ED when advised to attend by friends or family (Nelson, 2011; Penson et al., 2012), while McGuigan and Watson (2010) found that the second most common reason for non-urgent attendance at ED was ‘advised to attend by someone else’. Parents of young children are considered risk averse and will attend ED to be cautious (Rowe et al., 2015; Chapman & Taylor, 2016). Atenstaedt and Evans (2015) found that parental anxiety increased if they have to wait to see a doctor, even until later in the day. In Andersen’s model, this is considered to be an *Individual Predisposing Characteristic (Social),* as social networks such as the presence of family and friends can facilitate or impede utilisation (Andersen et al., 2014). Penson et al. (2012) and Coleman et al. (2001) classified this reason as ‘medium’ strength.

5.8 Results: closed questions

Based on the literature review, the questionnaire was developed in three parts: (i) The first section contained demographic data; (ii) The second section captured reasons for attendance, as described in the previous section. By categorising the mutability of these decision-making reasons, and investigating which are influenced by real-time data, it is possible to determine which may subsequently result in behavioural change; (iii) The third section, open questions are in Section 5.9. The consent form and full questionnaire are in Appendix 2a. Samples of the raw data and analysis are in Appendix 2b.

5.8.1 *Study setting*

The study was conducted in two EDs and two MIUs across two NHS Trusts. Two researchers collected the data over a six-week period. Participants aimed to be a representative sample of low-acuity patients seeking urgent health care. All questionnaires for participants under 16 years were completed by a parent or carer. No participants had previously been exposed to *NHSquicker*, although it had launched to the public in December 2017, and questionnaires were completed in August 2018. This was partly due to low early adoption and partly due to the relatively small sample size, which was collected over approximately 30 hours, one hour per visit, with approximately 5 questionnaires per hour completed.

5.8.2 *Results*

5.8.2.1 *Summary statistics*

This section presents the results of the closed questions in the questionnaire. Questionnaires were facilitated on-site by a researcher. A total of 152 completed questionnaires were analysed. One hundred and sixty-two patients were approached, and seven declined to participate; a further three questionnaires were incomplete as patients were called for treatment. The most common reasons for refusal were for eye or hand injuries. 55% were completed by the patient and 45% by a respondent answering on behalf of the patient. 53% of the sample were female and 47% were male. In accord with the findings in the literature, 90% of the sample checked ‘this is the most appropriate place for me today’ (e.g. McGuigan & Watson, 2010; Beache & Guell, 2015; Sancton et al., 2018), evenly distributed over those who would and would not have found the real-time DA useful for their current visit. Very few patients indicated that they were motivated by parking concerns, transport options available to them, the weather, 24-hour access, not being registered with a GP, or information obtained over the internet about available services or their condition.

The following is a summary of the perceived usefulness of the real-time information:

Perceived usefulness of the real-time information	Yes	No	Unsure
It would have been useful for me today	38%	31%	31%
It will be useful for me in the future	68%	24%	8%
I will recommend it to a friend	76%	22%	2%

There is a strong age relationship between those who would or would not find the data useful. For example:

- In those aged over 50 years, only 17% would have found the data useful today, though 48% thought it might be useful in the future and 71% would recommend it to a friend.
- For those aged over 70 years, no patients would have used it today, though 39% thought they may in the future.
- In the 18-35 year age groups, 56% would have found it useful today, and 87% in the future.

5.8.2.2 *Statistical analysis*

The sample was classified into those who would, and would not, have found the real-time descriptive information useful for supporting their attendance decision today. The 'unsure' category was removed to ensure two definitive categorisations. The purpose is to determine the decision variable characteristics of those for whom the provision of real-time descriptive information forms a factor in their decision-making, and therefore which characteristics are potentially mutable to real-time DA. The null hypotheses (H0) is that there are no differences between the groups. The alternative hypotheses (H1) is that different factors influence attendance decisions for low-acuity patients between those who do, and do not, find real-time information useful (for todays' visit).

Chi square statistical analysis is used for testing relationships between categorical variables. Plots and tables below show the observed frequencies, observed percentage of overall sample, and the p-value. Assumptions for a chi-squared independence test are: (i) Independent observations: each case is a unique person so this assumption has been met; (ii) All expected frequencies must be >5: where small numbers are involved, these relationships were excluded from the analysis (marked as n/a).

The Mann-Whitney U test is used to compare the differences between the two independent groups where the dependent variable (DV) is ordinal, i.e. those measured using a Likert scale. It is a non-parametric test that can be used in place of an unpaired t-test. It is used to test the null hypothesis that two samples come from the same population (i.e. have the same median) or, alternatively, whether observations in one sample tend to be larger than observations in the other. The assumptions for Mann-Whitney U test are: (i) The DV should be ordinal or continuous. For Likert-scale questions, this assumption is met; (ii) The independent variable (IV) should consist of two categorical, independent groups: those who believe the real-time descriptive data would or would not have supported their decision-making for today's visit, meeting this assumption; (iii) Independent observations: each case is a unique person so this assumption has been met; (iv) The variables do not need to be normally distributed, but the shape of the distributions of the two groups must be determined. In this case, most of the distributions are not identical, but have a 'similar' shape, and the Mann-Whitney U test is used to compare the medians of the independent variable of the two groups. In those variables that do not meet the assumption, the output uses mean ranks only.

As patients were given the options 'yes', 'no' and 'unsure' for a number of questions, where 'unsure' is included in the analysis, the independent-samples Kruskal-Wallis test is used. This is a non-parametric test with the same assumptions as Mann-Whitney: samples are random and mutually independent, the DV measurement scale is ordinal and IV is categorical. These assumptions are all met. The analysis is presented below. The test used for analysis is indicated for each. All analyses are 2-tailed, as no assumptions are made about the direction of difference between groups.

The difference between age groups is significant, using Mann-Whitney U mean ranks, with younger people significantly more likely to consider the real-time data useful (Figure 5.5).

Age	Useful		Not Useful		
	n	%	n	%	p-value (2-tailed)
<5 years	3	5.2	4	8.5	*0.002
6-18 years	4	6.9	3	6.4	
18-35 years	29	50.0	10	21.3	
35-50years	12	20.7	5	10.6	
50-70years	10	17.2	15	31.9	
<70years	0	0.0	10	21.3	

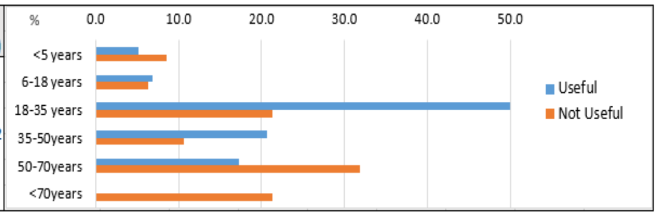


Figure 5-5 Mann-Whitney U mean ranks for age group (DV) and perceived usefulness of real-time data (IV)

No significant difference was found using Mann-Whitney U tests between usual health for those who would, and would not, find real-time descriptive data useful, however there is a general trend toward better general health for those who would find it useful. This may reflect the relatively younger ages (Figure 5.6).

Usual Health	Useful		Not Useful		p-value (2-tailed)
	n	%	n	%	
Poor	4	6.9	1	2.1	
Fairly Poor	2	3.4	1	2.1	
Moderate	6	10.3	8	17.0	0.489
Good	17	29.3	17	36.2	
Very Good	29	50.0	20	42.6	

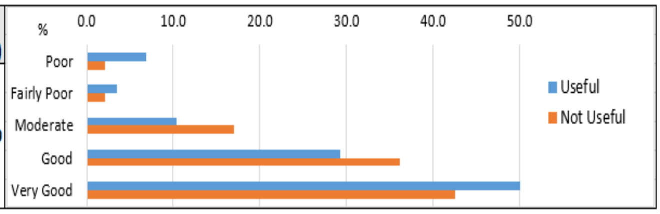


Figure 5-6 Mann-Whitney U test for usual health (DV) and perceived usefulness of real-time data (IV)

No significant difference was found using Mann-Whitney U tests between urgency and seriousness for each group, however there is a general trend toward lower perceived urgency and lower perceived seriousness in those who considered that real-time descriptive information would have been useful today (Figure 5.7).

Urgency	Useful		Not Useful		p-value (2-tailed)
	n	%	n	%	
Low	3	5.2	4	8.5	
Fairly Low	12	20.7	6	12.8	
Moderate	26	44.8	16	34.0	0.217
High	12	20.7	15	31.9	
Very High	5	8.6	6	12.8	
Seriousness	**Useful**		**Not Useful**		**p-value (2-tailed)**
	n	%	n	%	
Low	6	10.3	2	4.3	
Fairly Low	14	24.1	10	21.3	
Moderate	23	39.7	18	38.3	0.174
High	11	19.0	12	25.5	
Very High	4	6.9	5	10.6	
Certainty	**Useful**		**Not Useful**		**p-value (2-tailed)**
	n	%	n	%	
Low	2	3.4	2	4.3	
Fairly Low	3	5.2	1	2.1	
Moderate	9	15.5	4	8.5	*0.038
High	16	27.6	7	14.9	
Very High	28	48.3	33	70.2	

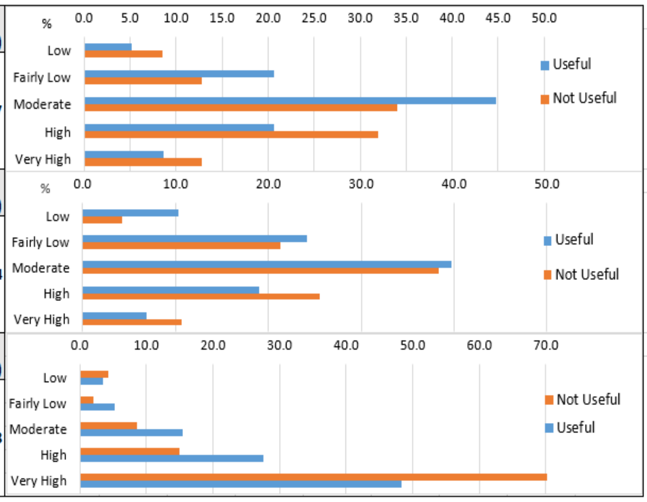


Figure 5-7 Mann-Whitney U tests for perceived urgency, seriousness, and certainty (DVs) and perceived usefulness of real-time data (IV)

There is a significant difference between those who find real-time data useful, and not useful, and the certainty of patients that they are in the most appropriate place for today's attendance. This suggests that where there is uncertainty about where best to attend, the additional information provided by the real-time data may be useful for supporting attendance decisions (Figure 5.7). The null hypotheses is that there is no difference between certainty ratings between the two groups.

Using chi-square analyses, there was no significant difference in the sample between those who are aware of each alternative facility (Figure 5.8). The majority of patients claimed to be aware of alternative facilities, apart from urgent care centres (UCC), which in the South-West of England are an emerging service. Note that questions that are 'tick all that apply' are treated as independent analyses.

However there is a significant difference in perceptions of the appropriateness of alternative facilities, in particular UCCs, WICs ($p<0.01$) and GPs, with more patients who would have found real-time descriptive information useful for their current visit considering these services appropriate today (Figure 5.9). The null hypotheses is no significant difference between the two groups.

Awareness of alternatives	Useful		Not Useful		p-value(2-sided)
	n	%	n	%	
ED	58.0	100.0	47.0	100.0	
UCC	21.0	36.2	13.0	27.7	0.352
MIU	49.0	84.5	35.0	74.5	0.202
WIC	52.0	89.7	38.0	80.9	0.200
GP	52.0	89.7	42.0	89.4	0.961
NHS111	52.0	89.7	40.0	85.1	0.482
Pharmacy	51.0	87.9	45.0	95.7	0.155
Other	37.0	63.8	39.0	83.0	*0.029

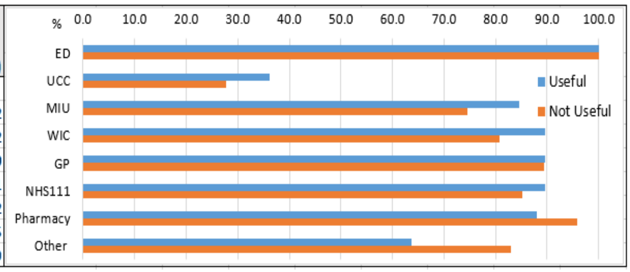


Figure 5-8 Chi square analyses for awareness of alternative facilities

Perceived appropriateness	Useful		Not Useful		p-value(2-sided)
	n	%	n	%	
ED	49.0	84.5	42.0	89.4	0.465
UCC	13.0	22.4	4.0	8.5	0.054
MIU	22.0	37.9	15.0	31.9	0.521
WIC	28.0	48.3	11.0	23.4	*0.009
GP	19.0	32.8	8.0	17.0	0.067
NHS111	5.0	8.6	3.0	6.4	n/a
Pharmacy	2.0	3.4	3.0	6.4	n/a
Other	4.0	6.9	3.0	6.4	n/a

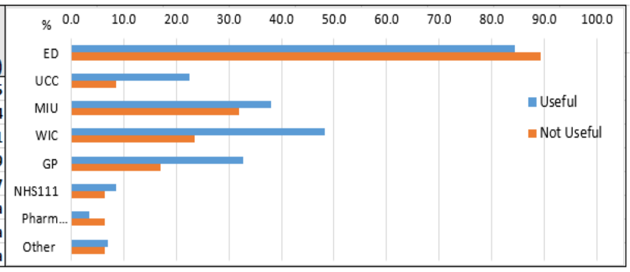


Figure 5-9 Chi square analyses for perceived appropriateness of alternative facilities

Each of the questionnaire items are individually analysed categorically using chi square analyses prior to grouping into the mutability categories derived in Sections 5.6 and 5.7. (Figures 5.10 and Table 5.3). The internal consistency was checked by ensuring the direction of responses within the categories (Tsang et al., 2017). As with Figures 5.8 and 5.9, Chi square tests are carried out on the *actual numbers* of occurrences, not on percentages, proportions, means of observations, or other derived statistics. However the accompanying plots illustrate the percentages.

Interestingly, there is no difference in the findings between those who were referred in to the facility and those who were not, and whether they would or would not have found the real-time descriptive data useful for today's decision. However those who were 'advised to attend' are statistically significantly different for each group, and these may represent patients who could not get a same-day appointment and were advised to go elsewhere if they considered their condition to be urgent. In other words, the patient made the attendance decision, rather than a healthcare professional.

Many of these patients may not have been aware of alternative facilities, and this is reflected in the statistically significant difference between those who were 'not sure where else they could have gone today' who considered the real-time data useful and those who did not.

Table 5-3 Questionnaire items grouped by validated classification criteria, with individual chi square analyses

Classification of attendance factors	abbreviation	Useful		Not Useful		
		n	%	n	%	p-value (2-sided)
Advised to attend by Health Care Professional						
Referred by another service	ReferHCP	25	43.1	25	53.2	0.303
Advice from Health Care Professional	AdviceHCP	20	34.5	28	59.6	*0.01
Ease of access						
This is the closest service for me today	ClosestServ	31	53.4	16	34.0	*0.047
The service is the nearest ot me	NearestServ	16	27.6	10	21.3	0.456
It's easier to get here than another service	EasierHere	9	15.5	2	4.3	n/a
Nearness to my location	Nearness	19	32.8	12	25.5	0.42
Uncertainty about alternative facilities						
I'm not sure where else I could have gone	UnsureAlt	14	24.1	4	8.5	*0.035
I don't know where to find other services	UnsureFindAlt	3	5.2	2	4.3	n/a
Im not sure other services are right for me today	UnsureApprop	9	15.5	6	12.8	0.689
Personal knowledge about this service	Knowledge	30	51.7	16	34.0	0.069
I don't know what other services are open	UnsureOpen	6	10.3	3	6.4	n/a
Wait times						
I'm more likely to have to wait at other services	WaitAlt	3	5.2	1	2.1	n/a
I think I'll be seen more quickly here	QuickerHere	5	8.6	0	0.0	*0.039
My GP wasn't available	GPUnavail	14	24.1	9	19.1	0.539
I don't know if my GP was available	GPUnavailUnsure	3	5.2	3	6.4	n/a
I didn't want to see my GP	Don'tWantGP	4	6.9	2	4.3	n/a
I would have to wait for my GP	WaitGP	9	15.5	5	10.6	0.465
Perceived severity						
I wanted to see a specialist	SeeSpecialist	8	13.8	13	27.7	0.077
My condition is an emergency	Emergency	8	13.8	10	21.3	0.312
I might need to go to hospital	MaybeHospital	7	12.1	9	19.1	0.315
Another service would have referred me here anyway	AltReferHere	13	22.4	15	31.9	0.274
Anxiety						
I need reassuring that its not serious	Reassurance	21	36.2	11	23.4	0.156
Personal instinct about this service	Instinct	15	25.9	7	14.9	0.17
I'm confident in this service	Confidenthere	12	20.7	10	21.3	0.941
I need to see a doctor as soon as possible	SeeDrASAP	17	29.3	5	10.6	*0.019
Advised to attend by family/friend						
Somebody with you when decided to seek treatment	Accompanied	44	75.9	35	74.5	0.869
Advice from family/friend/other	AdviceFamily/friend	10	17.2	9	19.1	0.801
Past Experience						
Previous experience about this service	PastExperience	12	20.7	11	23.4	0.738
I've used this service before and was satisfied	SatisfiedHere	10	17.2	12	25.5	0.299
Real-time data as NHSquicker						
Do you think you would have found it useful today?	NHSqUsefulToday	58	100.0	0	0.0	
Do you think it would be useful at a later time?	NHSqUsefulLater	57	98.3	23	48.9	*0
Would you recommend it to your friends/family	RecommendFriend	57	98.3	32	68.1	*0

For those who chose the 'closest service', there is also a statistically significant difference, suggesting that some patients may be prepared to travel further to be seen more quickly. Similarly, there is a significant difference between those who chose their current facility because they thought they'd 'be seen more quickly', suggesting that being seen promptly is a priority and that this can be supported by knowledge of wait times. However very few of the participants indicated that wait times were a priority in their decision-making processes. This may be an artefact of self-report bias, or it may be that other factors took priority in making the decision. This is further reflected in the anxiety measure 'I need to see a doctor as soon as possible', which is significant at $p<0.02$ in a group of non-urgent patients for those who would, compared with those who would not, have found the real-time information useful today. Those for whom it would have been useful today are also statistically more likely to find it useful in the future and to recommend real-time descriptive information to a friend.

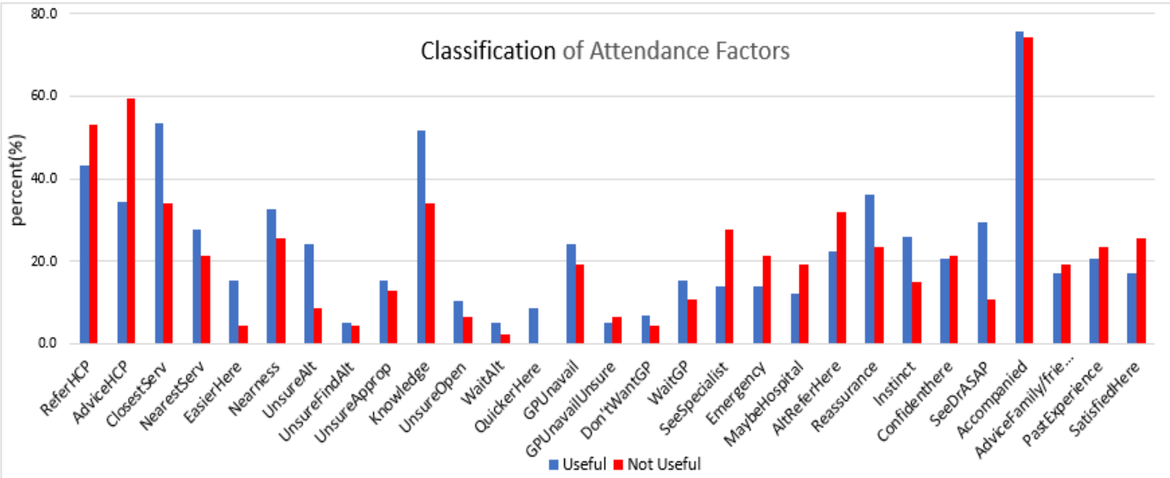


Figure 5-10 Attendance factors by perceived usefulness of real-time descriptive data (refer to abbreviations in Table 5.3 for full descriptions)

5.8.3 *Categorisation of variables*

5.8.3.1 *Validity and reliability*

The questionnaire was developed according to the categories identified in the literature (Sections 5.6 and 5.7). Items in the questionnaire aim to be representative of these categories. The face validity, how well the questionnaire measures what it intends to measure, and the content validity, how well the questionnaire items cover all aspects of these constructs, was validated by two NHS clinicians and one NHS manager.

Construct validity is the extent to which the survey measures a construct that is not directly observable, such as attitudes or beliefs. To increase the construct validity of the questionnaire, on several occasions the same question was asked twice but reworded slightly. Additionally, groups of questions relate to aspects of the underlying theoretical construct as used in the literature. Construct validity is often evaluated using confirmatory factor analysis. It shows how the correlations between the questionnaire items can explain the underlying latent construct. A weakness of this questionnaire is that many of the questions are categorical, which means that correlations between variables cannot be measured. This additionally means that the internal consistency can't be measured – the extent to which the questions for each construct measure the same construct. However the direction of responses within each theoretical construct was analysed to support internal consistency.

5.8.3.2 *Results of categorisations*

The results of the categorised factors is presented in Figure 5.11.

The categories *Ease of Access, Anxiety, Perceived Seriousness and Advised to Attend by a Health Care Professional* were all statistically significantly different between those who would (n = 58), and those who would not (n = 47), have found the real-time descriptive information useful for supporting their attendance decision today.

Ease of access is considered to have medium mutability, mostly due to the fact that health services can be relocated to improve access. However in this case, it seems that a significant proportion of patients who prioritise convenience to the facility may be prepared to travel to be seen more quickly, and that knowledge of wait times may support this decision. Similarly, a significant proportion of those motivated by *anxiety* and the need for reassurance would consider the real-time descriptive information useful today. This may be because more information reduces anxiety. Alternatively it may be that anxiety can be reduced by being seen more quickly at a different facility.

Those who consider their *condition to be serious* are significantly less likely to consider the real-time data useful for supporting this attendance decision. This category is considered to have low mutability. Similarly, those who have been *advised to attend by a health care professional* are significantly less likely to consider the real-time descriptive information useful today. Being advised to attend is likely to increase perceptions of severity, reinforcing the belief that the condition is serious.

While *uncertainty about alternatives* was not statistically significant in this sample, there is a strong suggestion that those who would have found the real-time descriptive information useful today are more likely to be uncertain about alternative facilities. Interestingly, those who *prioritise wait times* were only slightly more likely to consider the real-time descriptive information useful or not useful in this sample. Equal numbers of patients who were *advised by friends and family*; or were motivated to attend by *past experience* considered the real-time data useful or not useful today. Previous experience was considered to have low mutability but this would indicate that approximately half of patients have the potential to change given new information.

Categories	Mutability	Useful		Not Useful		
		n	%	n	%	p-value (2-sided)
AdvisedHCP	Low	26	44.8	32	68.1	***0.017**
EaseOfAccess	Medium	38	65.5	20	42.6	***0.019**
UncertaintyAlt	High	36	62.1	21	44.7	**0.075**
WaitTimes	Medium	24	41.4	16	34.0	**0.441**
PerceivedSeverity	Low	42	72.4	47	100.0	***0**
Anxiety	Medium	40	69.0	23	48.9	***0.037**
AdvisedFamily	Medium	47	81.0	37	78.7	**0.768**
PastExperience	Low	17	29.3	17	36.2	**0.455**

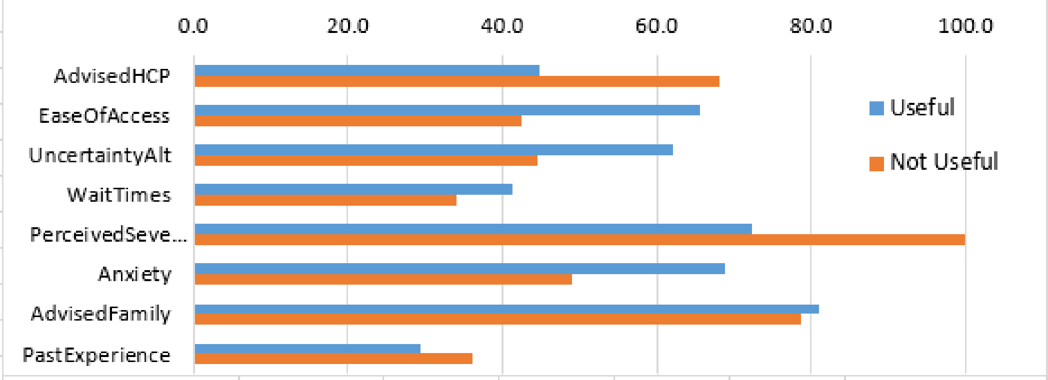


Figure 5-11 Chi squared analysis of categorical themes according to perceived usefulness of real-time information

5.8.4 *Summary of findings: closed questions*

The closed questions characterised low-acuity patients who are more likely to consider real-time descriptive information useful for supporting attendance decisions. Those patients in ED who consider real-time data useful for their own attendance decisions tend to be younger, in better health, less certain whether ED is the most appropriate place to be, have not been referred from another service, prioritise convenience of access, are less certain about what alternative facilities are available, don't like waiting, and have a tendency to be more anxious and in need of reassurance.

In contrast, those who don't consider real-time information to be useful are more likely to be older, in poorer health, more sure that ED is the best place for them, more likely to consider their condition to be serious, are unconcerned by waits and are less likely to be anxious. Dinh et al. (2017) analysed over 10.7 million ED presentations, and classified 45% of these as low-acuity presentations, defined as those who self-presented (were not transported by ambulance), were assigned a triage category of 4 or 5 (semi-urgent or non-urgent) and discharged back to usual residence from ED. They found that one of the strongest predictors of low-acuity presentations was being aged less than 40 years, suggesting that the real-time information is targeting the most appropriate age groups. Additionally, from the literature review in Section 5.6, and the observational data in Section 5.3.1, it seems that anxiety is a characteristic of low-acuity attenders (Lega & Mengoni, 2008; Backman et al., 2008; Agarwal et al., 2012), which appears to strongly influence decision-making (Pearson et al., 2018) and is considered to be of medium mutability (Coleman et al., 2001; Penson et al., 2012; Mason et al., 2017). If knowledge of real-time wait-times can influence the perceived need for services, it may be achieving this by reducing anxiety. This can therefore influence more than health-seeking behaviours, potentially affecting outcomes such as consumer (patient) satisfaction. Feedback loops back into earlier model components (Andersen et al. 2014) indicate that this can have a positive reinforcing effect on contextual and individual characteristics, and ongoing health-seeking behaviour.

Prioritisation of convenience and ease of access are characteristics of low-acuity ED attenders for whom the real-time information is considered valuable. Non-urgent patients are reported to attend for convenience of location, despite being aware that their condition is not urgent or serious, and could wait (Chapman & Turnbull, 2016; Mason et al., 2017). Convenience, proximity, and travel options enable access and utilisation (Andersen et al., 2014) and have medium mutability (Coleman et al., 2001; Penson et al., 2012; Mason et al., 2017). It may be that knowledge of wait-times at nearby locations increases perceived convenience, as travelling further for an overall lower wait-time is preferable to visiting the nearest facility to wait for a longer period for treatment.

Those attendees who are older, in poorer health, consider their condition to be more serious, and have been advised to attend by a HCP are less likely to be mutable in their decision-making, but also, intuitively, perhaps more likely to be appropriate attendances. Both perceived severity and having been advised to attend by a HCP are individual 'need' characteristic (Andersen et al., 2014) with low mutability (Coleman et al., 2001; Penson et al., 2012; Mason et al., 2017). Afilalo et al. (2003) described these as 'inevitable non-urgent visits'. Real time DA is significantly unlikely to influence these attendance decisions.

Finally, patients who value wait-times are neither more nor less likely to be influenced by real time DA for their attendance decisions. The belief by patients that they'll be seen more quickly in ED, and not wanting to wait for a GP appointment are all classified as 'medium' strength reasons (Coleman et al., 2001; Penson et al., 2012). As patients (as participants) had no knowledge of actual wait-times, in Andersen's model wait-times can be classified as a predisposing characteristic representing a patient's beliefs, attitudes or knowledge about health services wait-times. The provision of actual wait-times will change this to an Individual Enabling Resource (Organisation), as knowledge of wait-times then enables a more informed choice. It's possible however that other characteristics of health-seeking behaviour may represent stronger reasons in some patients. For example a patient who believes their condition to be serious, or who has been advised to attend by a HCP may also value being seen more quickly, however the perceived severity will lead to ED attendance rather than attending an alternative facility for treating minor conditions.

The set of open questions provide more insight into these decision processes.

5.9 Questionnaire development: open questions

To understand how patients perceive the value in the real-time data, open questions gave patients the opportunity to consider how the real-time data would be useful today, or in the future, how they could see wait-times and predicted wait-times to be useful for themselves, and how they perceived the NHS could use this information for planning and delivery of services. This provides insight into the pressures that patients, as end-users, observe the NHS to be under, and what they perceive to be potential solutions. As it is a relatively new intervention for patients to have access to real-time hospital operational data, the value of this information for patients is currently unknown, and as end-users, it is valuable to involve patients early in a potential intervention. The following three questions were asked:

a) Why might it be useful for you or your friends/family to know about current waiting times at A&E and other urgent care services near to you?
b) How useful would it be to you or your friends/family if you could have predicted waiting times for the next few hours?
c) *NHSquicker* is meant for those seeking urgent care. However, can you think how the NHS could make use of this information in Devon & Cornwall?

Of 152 completed questionnaires, 128 respondents provided open data. While the previous sections have gathered data which will inform the evaluation of the application of the framework, this section supports the problem definition stage, as it is necessary to understand how the end-users of the system, the patients, perceive the usefulness and expected value of the real-time data.

The open data was manually coded thematically into 6 main nodes (overarching themes), with 9 child nodes (sub-themes, within themes), for analysis. This is done by reading and coding by similarity of theme, for example ‘Good to know for collection’ and ‘Helpful to know when preparing entertainment or to bring food/drink for children’ are both coded in the child node “For planning, e.g. childcare, parking”. Larger data sets can support multiple thematic hierarchies for analysis, however each child node layer contains a reduced number of data points. The nodes are tabulated in Table 5.4.

Table 5-4 Node/child node hierarchical analysis using NVivo

Node	Child node
Balance demand/capacity across UCN	Demand/capacity shaping
	Help plan staffing/resources
Other potential uses of real-time data	
Information other than wait times e.g. other available options	
Save time and travel	
Waiting time knowledge is useful	For “when to go” decisions
	For “where to go” decisions
	To manage expectations

	For planning, e.g. childcare, parking
	Whether it is appropriate to attend
	Reduce anxiety
	Predicted wait times are useful
Would not have changed mind today	

5.9.1 *Summary of open data*

Wait-times were considered useful overall, with 262 references to the value of knowing wait-times to support attendance decisions. These were sub-categorised into how patients perceived the usefulness of the wait-times, with the majority of respondents suggesting it would be useful for making decisions about where best to attend (80) with fewer references to when to attend, either a different time of day or day of week.

Predicted wait-times were considered useful, with 99 references, and a further 47 references to using descriptive and predictive wait time information for planning, for example childcare, transport, parking or work. There were multiple references to the value of additional information about available facilities or opening hours, especially for those who are out-of-area. Twenty-nine patients indicated that the real-time information would not have changed their attendance decision. The reasons given were mixed, for example the perceived seriousness of their condition, the facilities available at ED, and retaining the choice to attend the nearest service. Finally, with regard to prescriptive data, there were a number of references to the need for the NHS to balance demand and capacity, and to consideration of staffing and resources to balance demand. This indicates that patients can see the value at the system level of spreading demand across the UCN.

5.9.2 *Patient characteristics*

NVivo enables the coded themes to be split by attributes, that is, by the closed questions. This supports closer investigation of both open and closed questions. Data is coded into non-mutually independent categories. For example, "*…plan your visit accordingly, maybe decide if your visit is entirely necessary, prepare yourself for a wait and make plans...*" can be simultaneously coded as 'planning', 'manage expectations' and 'consider whether appropriate to attend'.

Coding queries and crosstab queries were conducted using NVivo to drill further into the data. For clarity, coded nodes (open data) will be double underlined, and attributes (closed data) will be dot-dash underlined in the following subsections. An example of crosstab results is in Appendix 2; further results are available upon request. The following subsections investigate the categories identified in Section 5.7 from the literature, and in Section 5.83 in the closed questions, to gain more insight in light of the open question data.

5.9.2.1 *Ease of access*

A statistically significant relationship was found between 'ease of access' measures and the usefulness of real-time information for decision-support (Section 5.83). It is therefore possible that a significant proportion of patients who prioritise convenient access to the facility may be prepared to travel to be seen more quickly, and that knowledge of wait-times may enable this decision.

In the coded nodes 'where to go' and 'to consider whether this service is the most appropriate', 32 participants (49 references) with the 'ease of access' attribute indicated the usefulness of the real-time data, with a number indicating that they would have made a different choice today, e.g. "Maybe would have gone to a different hospital with shorter waiting times" and "Very useful. I wouldn't have come". Nonetheless, 16 patients would not have changed their mind. For example: "Not useful, I would rather waste time than fuel".

Patients who value ease of access, and are not anxious, are more likely to consider when to go, i.e., coming at a different time or different day, but less likely to consider where to go, i.e. going to a different place, than those are simultaneously anxious. This suggests that anxiety is a higher priority in attendance decisions.

Patients with past experience of the facility who value ease of access are more likely to use the real-time data to manage expectations and to reduce anxiety. Patients who are unsure of alternatives and value ease of access are more likely to consider the real-time information useful to consider when to go, but not where to go, again suggesting that 'ease of access' is a strong attendance reason. Even where the urgency is considered low, and patients feel that they can wait, a proportion would rather attend the most convenient facility than find about alternative services.

Ease of access is considered to have medium mutability, and the findings reflect this. Patients who simultaneously are referred by a HCP, consider their condition to be serious or have personal knowledge of ED are less likely to consider the real-time information useful for changing attendance behaviour, but see the value in reducing anxiety and managing expectations. On the other hand, those who are uncertain of alternatives will consider when to go, and those who are anxious will consider where to go, relaxing the strength of the 'ease of access' reason.

5.9.2.2 *Anxiety*

A statistically significant relationship was found between 'anxiety' measures and considering the real-time information useful for decision-support. Of those categorised as 'anxious', there were 147 references to the usefulness of the real-time information.

To determine whether anxious patients are likely to use the real-time data to travel to be seen more quickly, or if having more information reduces anxiety without changing behaviour, 'anxiety' measures were used against the codes that supported the use of real-time data for reasons other than behaviour change: 'planning' and 'to manage expectations'. Forty references to using the real-time information for these purposes were made by 28 participants who had 'anxiety' attributes. The majority (12) were concerned about arrangement transport/lifts; while the remaining were concerned about parking, managing their own expectations, childcare, work and other commitments.

In contrast, 61 references of 36 patients with 'anxiety' attributes indicated that the real-time information would support decisions about where to go, when to go, or whether to go. Of these, 51 references were categorised as 'where to go'. In other words, this group of patients were concerned with being seen now, not at another time, and are prepared to travel to be seen more quickly.

Those who are neither anxious, nor value ease of access, are more likely to be interested in information other than wait-times, and less likely to use real-time information for planning childcare or parking. Those who are anxious and value ease of access are more likely to use the real-time data to reduce anxiety, manage expectations and decide where to go. Similarly, those who are anxious and not aware of alternatives are more likely to use the data to decide where to go, and less likely to consider that they 'would not have changed their attendance decision'.

Anxiety is considered to have medium mutability. It seems that anxious patients are happy to travel to be seen, but would not delay their treatment to be seen at another time. However patients who are both anxious and have been referred by a HCP or value ease of access are more likely to use the real-time data to manage their expectations, and less likely to consider it useful for where to go and when to go decisions.

5.9.2.3 *Uncertainty about alternative facilities*

The relationship between uncertainty regarding available services, and usefulness of real-time information did not reach statistical significance (p=0.075) but given a larger dataset, it is possible it might have done. 16 references to the node 'information other than wait-times' referred to the usefulness of the additional information and not knowing about other options. Those who simultaneously prioritised ease of access or who are anxious are less likely to consider this additional information useful but more likely to consider when to go, suggesting that ease of access is a stronger attendance reason than not knowing about other options. It's possible that 'not knowing' may sometimes mean 'not interested in finding out'.

There were 38 references to 'when to go', indicating that the information can support decisions to go at a different time, for example "Very useful, I would have waited (come at another time)!!" 28 patients indicated that the real-time information provides information about 'where to go', a surprisingly small number. Patients who are simultaneously anxious are less likely to consider when to go than those who are not anxious, while patients who prioritise ease of access are more likely to consider when to go. Patients who are anxious but not uncertain of alternatives are less likely to consider where to go than those are uncertain of alternatives.

While this attribute is considered to have high mutability, it's possible that many patients expressing uncertainty about where to attend are in fact prioritising ease of access.

5.9.2.4 *Severity*

There is a statistically significant relationship between the perceived severity of the condition, and the usefulness of real-time information. The more severe the condition is perceived to be, the less the real-time data is likely to impact on attendance decisions. This category is considered to have low mutability, and of the 34 references in the 'would not have changed mind' node, 32 of those were in the severity group. Patients with perceived severity who are not referred by a HCP are much more likely to consider the value in the data for making where to go and when to go decisions than those are referred. Patients who are serious and referred are more likely to use the real-time information for managing expectations. Patients who are serious, regardless of whether they prioritise ease of access, are more likely to use the real-time information for planning, and to manage expectations, and less likely to use it to help to decide where to go.

Patients who consider their condition to be serious are more likely to consider that the real-time information could be useful in the future if they do not prioritise ease of access. Patients who consider their condition to be serious are less likely to change their attendance behaviour, but still consider the information to be valuable for planning, managing expectations and reducing anxiety.

5.9.2.5 *Referred by HCP*

Patients are significantly less likely to consider the real-time information useful for supporting attendance decisions where they have been referred to the service by a HCP. Of the 34 references in the 'would not have changed my mind' node, 20 of these were in the 'referred by HCP' group.

There is significant overlap between the categories 'referred by HCP' and 'perceived severity', suggesting that patients who have been referred may be more likely to consider their condition to be serious. Patients who are referred, whether or not they consider their condition to be serious, are more likely to reflect the value of considering whether it is appropriate to attend. This may be because they consider their own attendance to be appropriate, and are concerned that others should reflect on the same. Patients who are both serious and referred are more likely to consider the information useful for managing expectations. However patients who are not referred, whether or not they consider their condition serious, are more likely to consider that the real-time data supports where to go decisions.

This attribute is considered to have low mutability, and the findings support this. However patients who have been referred are open to seeing the value in the data for ‘where to go’ decisions.

5.9.2.6 *Other categories*

The categories ‘wait-times’, ‘advised to attend by friends/family’, and ‘past experience’ were statistically similar between the groups who considered that the real-time information would be useful, or would not be useful for today’s visit. This supports the null hypotheses, that value in the real-time information for today’s decision is independent of these factors.

Those who have a previous positive experience of the service are more likely to consider the real-time information useful for reducing anxiety, managing expectations, planning, and considering when best to attend. They appear less likely to consider where to attend, and are less likely to see the value in information other than wait times, in particular where they value not waiting. However a proportion are happy to attend the same facility at a different time.

Similarly, those who have were advised to attend by family or friends are more likely to consider the real-time information useful for planning, reducing anxiety, managing expectations, and considering where to go. They are more likely to be confident that they would not have changed their mind today, than those whose family/friends were not involved in the attendance decision.

Those who prioritise wait times are likely to have attended ED to avoid waiting to see their GP. They are more likely to consider that their attendance decision would remain unchanged, less likely to consider when to go, but open to the possibility of considering where to go, particularly where friends or family are not involved in the decision.

5.9.2.7 *Certainty about today’s decision*

There is a significant difference between those who find real-time data useful, and not useful, and the certainty of patients that they are in the most appropriate place for today’s attendance (p<0.05). This suggests that where there is uncertainty about where best to attend, the additional information provided by the real-time data is considered to be useful for supporting attendance decisions. The open data was used to investigate the value that the real-time information is providing to patients who are certain, or not certain, that they have accessed the most appropriate service.

As might be expected, those who value information other than wait times are more likely to be uncertain, suggesting that they would value knowing where else they could have gone for this attendance decision.

Those who would use the real-time information to reduce anxiety and to manage expectations are more likely to be certain of their decision, and less likely to consider that they would have changed their mind. Again this is to be expected.

Those who consider the value in 'when to attend' decisions are more likely to be certain. This suggests that while they are confident they are in the right place, some patients could have waited to attend at a different time.

Those who value 'where to attend' decisions are spread evenly across all levels of certainty. This suggests that while today's decision might not change, the value in making 'where to go' decisions is still seen as important.

5.9.3 *Patient characteristics: conclusion*

A summary of the analysis is depicted in Figure 5.14.

Observational data found a general concern by managers and clinicians that more information is needed about how people with low-acuity conditions use ED. They agreed that this required engaging with patients about what mattered to them to find out what patients base attendance decisions on. Clinicians were keen that assumptions were not made about who 'needed' to be seen in ED, and that providing information does not necessarily mean facilitating a decision, which may be multifactorial.

The questionnaire set out to do this, and to investigate whether real-time information can support these attendance decisions, and in which patients. Staff felt that the provision of new information would reduce anxiety, but there was some concern that it may lead patients to make suboptimal decisions, as 'low-acuity' doesn't necessarily equate to 'inappropriate', and for a proportion of these patients, ED is the most appropriate service. This knowledge contributes to an understanding of the added value to the system of using real-time data applications for patient decision-support.

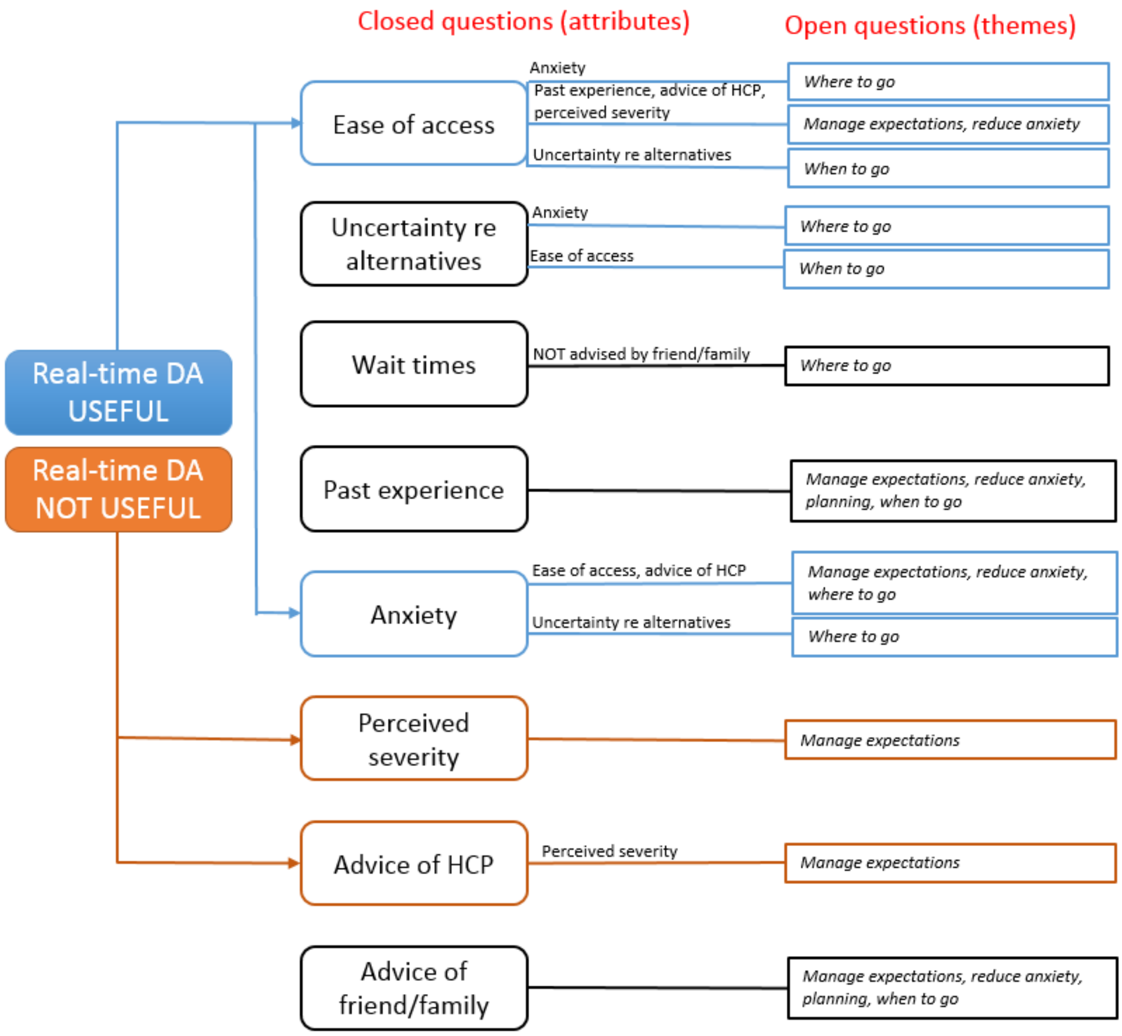


Figure 5-12 Summary of open and closed questions analysis

The closed questions characterised low-acuity patients who are more likely to consider real-time descriptive information useful for supporting attendance decisions. Those patients in ED who consider real-time data useful for their own attendance decisions tend to be younger, in better health, less certain whether ED is the most appropriate place to be, have not been referred from another service, prioritise convenience of access, are less certain about what alternative facilities are available, don't like waiting, and have a tendency to be more anxious. Open questions sought to explore how real time DA is considered to add value for patients.

The attribute 'ease of access' is considered to have medium mutability, and the findings reflect this. Patients who simultaneously are referred by a HCP, consider their condition to be serious or have personal knowledge of ED are less likely to consider the real-time information useful for changing attendance behaviour, but more useful for managing expectations or reducing anxiety. However those who simultaneously are uncertain of alternatives will consider when to go, and those who are anxious will consider where to go.

Anxiety is considered to have medium mutability. It seems that anxious patients are happy to travel to attend a different service, but would not delay their treatment to be seen at another time. However patients who are both anxious and have been referred by a HCP, or value ease of access, are more likely to use the real-time data to manage their expectations, and less likely to consider it useful for where to go and when to go decisions.

The attribute ‘uncertainty about alternative facilities’ did not quite reach statistical significant between those who would, or would not, consider the real-time information useful for today’s attendance decision. While it is considered to have high mutability, it’s possible that many patients expressing uncertainty about where to attend are in fact prioritising ease of access. Those who simultaneously prioritised ease of access or who are anxious are less likely to consider this additional information useful but more likely to consider when to go, suggesting that ease of access is a stronger attendance reason than not knowing about other options.

Patients with the attributes ‘perceived severity’ and ‘referred by HCP’ are less likely to change their attendance behaviour, but still consider the information to be valuable for other uses. Patients who are both serious and referred are more likely to consider the information useful for managing expectations. These categories are considered to have low mutability, and the data supports this.

Overall, a large proportion of patients considered that real-time information is valuable for supporting attendance decisions, and the open data categorised the responses into a number of factors. These are summarised in a word cloud in Figure 5.15 for visualisation.

Figure 5-13 Word cloud summarising open data terms

Twenty-nine out of 128 patients indicated that they would not have changed their attendance decision. However while those who are certain they are in the right place are less likely to consider the real-time information useful for today's decision, they are as likely as those who are uncertain to consider that the information is valuable for 'where to attend' decisions. As the decision about where to attend is left with the patient, it is valuable to understand how and why patients take the decisions they do, and how additional information may support these. These findings suggest that patients are retaining the ability to make appropriate attendance decisions, and are able to use the real-time information for a variety of purposes, both to support decisions about where to attend, but also to add value to their attendance in a number of other ways.

5.9.4 *Value in predicted wait times*

The majority of respondents perceived value in having access to predicted wait times, with 99 references, of which ten indicated that they wouldn't find it useful, for example, "Not relevant - if it's an emergency I have to attend anyway. I don't attend A&E for non-emergencies". Others referenced difficulties with transport, or their proximity to a service, for example, "Exeter A&E is still the closest". The word 'useful' was used 59 times for this question. The majority of respondents who provided a reason for this value were coded as planning, e.g. for childcare, parking, travel or work. For example, "Would have checked wait times today. Would have made easier child care arrangements". However a number of patients were interested in 'when to go' and 'where to go' decisions, for example

“Very useful, I would have waited (come at another time)!!” and “So you can make an informed decision about where to go”. Other individual enabling characteristics can attendance decisions, for example: “Limited use to us as no car/personal transport to reach other areas.” Those who arrived using public transport or ‘other’ (e.g. walk, taxi) still saw the value however, indicating “Would have helped arrange transport”, and “It would be useful to help decide who would be best placed to take the patient and wait with them”, so it seems the additional predictive information can potentially be used to assist with planning alternate transport means.

Those who valued ease of access and were unsure of alternatives are less likely to find the predicted wait times useful. Patients who are anxious, and have also been referred, are more likely to find the predictions useful than those who have not been referred. From the previous analysis, it is likely that this group of patients are more interested in using the data for planning and managing expectations, than for attendance decisions. Patients who consider their condition to be serious are more likely to find the predictions useful, in particular if they have been referred, and in particular if they value ease of access. This is likely to reflect the value that the majority of patients have placed in this data for planning. Patients who consider their condition to be serious are equally likely to find the predictions useful, whether or not they prioritise waiting, and whether or not they are anxious.

It seems that patients perceive value in the predictions, but the questionnaire data indicates that the value it offers patients will be toward improving their experience, rather than changing their attendance behaviour. However some patients indicated that they would use predictions to consider attending a different service, “Would be good to know how long we'd roughly have to wait and whether anywhere nearby would be able to see us sooner,” and “If not a direct emergency it would enable you to make a decision about when to go, hopefully having a knock-on effect on waiting times etc.” This indicates that patients see the value for themselves in making both ‘where to go’ and ‘when to go’ decisions, but that they also see the potential advantages to the system, as the ‘knock-on effect’ to waiting times is an indirect advantage to both other patients, and to system performance.

5.9.5 *Value to the NHS*

Patients were asked to consider how they saw the value of real-time DA using wait time data for the NHS, by answering the question, "*NHSquicker* is meant for those seeking urgent care. However can you think how the NHS could make use of this information in Devon & Cornwall?"

There were thirty-two references to balancing demand and capacity, and thirteen to consideration of staffing and resources to balance demand. For example, "Doctors could better direct and refer patients to share the workload across the county. Predictions can help to work out where resources need to be sent and at what times", and "For all of these, GPs etc. could use this info to inform patients of where to go if a visit to A&E is needed. This could result in a more even spread of patients across hospitals etc. rather than the larger ones having the greater percentage."

This indicates that patients can see the value at the system level of spreading demand across the urgent care network. Patients indicated both that the NHS could use the information to spread demand by redirecting patients, for example "Divert urgent/less urgent people to other health services" and "diverting patients to quieter services"; and that by supporting patients to change attendance decisions, the NHS might benefit, for example "Help myself, help the system" and "Help patients to select the most appropriate place and time for both themselves and for the NHS." Additionally, patients saw value for other patients and the system by changing health-seeking behaviour, for example "hopefully will deter those with minor injuries attending A&E!!" Using the information to redirect or optimise staffing and other resources was considered to be a good use of the descriptive and predictive information, for example "Service planning i.e. staffing at peak times" and "could help facilities prioritise staffing levels". Patients also saw value for the NHS in having information about the performance of other services nearby, for example, "By creating greater flexibility and also for giving further information to other services".

There were 19 references to additional uses of real-time information, including expanding to include GP wait times, NHS111 using the wait-time information, and expanding the availability of the information outside of its current geographical coverage. This was an open question, without reference to ED crowding, however it seems that patients could very clearly see the potential value for the NHS in using the data to support decisions about crowding.

5.9.6 *Limitations*

This questionnaire study has several limitations. Firstly, while a full set of decision variables aimed to be collated from the literature, the use of a single database (CINAHL) may have restricted the search. CINAHL is limited to medical and allied health journals. The variables may have been strengthened by using a second researcher to validate identified variables, however efforts were made to validate categorisations using NHS staff.

As eight theoretical constructs - categories of ED attendance reasons - were identified in the literature, the questionnaire was long. To ensure a high response rate, researcher facilitation was required to explain and demonstrate the real-time information, introduce the questionnaire, gain consent and remain nearby to collect completed questionnaires. The risks in using this approach include introducing self-report bias, where patients seek social desirability or social approval (Donaldson & Grant-Vallone, 2002). For example following recent media emphasis on ED pressures, patients may reinforce their own beliefs that they are in the most appropriate place for their condition rather than admit to contributing to unnecessary ED demand. Few people indicated that wait-times were a motivation for attendance. This may be discomfort at the suggestion that they were not prepared to wait a few days for a GP appointment, or it may be a genuine belief that their condition is too urgent to wait. Additionally, it might seem more socially desirable to be uncertain of other services nearby, than to admit that ED is not the most appropriate place today. The extent of these uncertainties remains unknown.

An additional risk of long questionnaires is no-saying and yes-saying: some respondents answer yes or no to all questions (Choi & Pak, 2005). Multiple response questions in particular may have suffered from no-saying, as ticking the box is likely to indicate a definite yes. Finally, open-ended questions were left until the end of the questionnaire, where response fatigue may have limited the richness of the data collected (Choi & Pak, 2005). For example, a large number of responses asking about the perceived usefulness of predictive wait-time data were along the lines of ‘yes, very useful’.

Multiple response questions were chosen for the main data collection rather than forced choice 'yes-no' questions. While the implicit assumption is that checked items correspond to 'yes' and unchecked items correspond to 'no', higher endorsement rates are observed for 'yes-no' questions, which might overstate the response, by forcing respondents to choose among limited options (Meyners & Castura, 2014; Callegaro et al, 2015). However with multiple response items, interpretation of the unchecked box could be 'no', a 'maybe yes', a missed entry, uncertainty, or not wanting to answer the question (Choi & Pak, 2005; Callegaro et al, 2015). While this makes answering the questionnaire faster for patients, a limitation to this style of question is that it is an insensitive measure. There may be insufficient discriminating power to differentiate the respondents, compared, for example, with a Likert scale.

Due to the intensive nature of data collection, it wasn't possible to collect a larger sample in the limitations of this research. However a larger sample would have increased the statistical power and supported stronger interpretations between the open and closed codes, as some of the cross-matrices contained small numbers. For this reason, the comparative open/closed analysis is interpreted as indicative, and percentages were not included in the results. An example of an NVivo cross-tab analysis is included in Appendix 2b.

Finally, the questionnaire design makes construct validity and internal reliability difficult to assess. This means that conclusions drawn from the questionnaire have limitations. However triangulating the results with the literature, staff direct observations, and in Chapter 7 with staff interviews, aims to increase the validity of the findings.

5.10 Chapter Summary

This chapter has addressed the problem definition phase of the second aim of RQ2, to apply IHAF within a case study in a hospital ED (Table 5.4).

Table 5-5 Research Question 2

Research Question	Aim	Objectives
2. *How can an integrated hybrid approach using real-time simulation and data analytics support short-term*	To test and evaluate the potential of an integrated hybrid approach for short-term decision-support in healthcare	**1.** To propose a generic framework supporting an integrated hybrid approach for short-term decision making in healthcare

operational decision-making?	combining real-time simulation with analytics approaches.	**2.** To apply the framework within the case study in a hospital ED **3.** To evaluate the framework in this context

A study using HM requires a conceptual framework to consider the constituent stages of a conventional M&S study and to explore complementary techniques (Mustafee & Powell, 2018). IHAF provides such a conceptual framework to support the HM for its intended purpose. It includes an evaluation component to both determine the value of this approach in its applied setting, and to provide knowledge for future iterations. To support evaluation, the problem definition phase seeks to determine criteria and influencing factors for evaluating the HM. At this stage, a formative evaluation may be considered necessary (Venable et al., 2017), determined by considering the specific purpose of the HM, all stakeholder groups, and possible evaluation methods. For this case study, direct observations and patient questionnaires were used as a formative evaluation. This chapter presented the development, implementation and analysis of the patient questionnaires which, as an evaluation, contributes to the problem definition, and will be subsequently synthesised with the final evaluation in Chapter 7.

The results show that real-time DA have the potential to contribute to reducing ED crowding by influencing both patient health-seeking behaviour (through availability of real-time and predicted wait-time data), and staff decision-support (through the HM).

For patients, with consideration of Andersen's conceptual model (Andersen et al., 2013), it is clear that there is the potential for the real-time DA to impact on *Individual Characteristics*, in particular by providing enabling information and by reducing anxiety. *Contextual characteristics* are considered to be fixed, that is, not influenced by the real time DA. In the model, a combination of *contextual and individual characteristics* result in health-seeking behaviour and health outcomes. By providing relevant information, including real-time and predicted waiting times, about alternative facilities (*Individual Enabling Resource: Organisation*), and reducing anxiety (*Individual Predisposing Characteristics: Social and Beliefs*), the 'perceived need' for health services can be influenced. The model suggests that this results in health-seeking behaviour, and the questionnaire data has indicated that a subset of patients are likely to use the real time DA to make 'where to attend' and 'when to attend' decisions.

Those who are more likely to consider 'where to go' are patients who are anxious, prioritise wait-times and ease of access, and who are uncertain of alternative facilities. Those who are more likely to consider 'when to go' are patients who are not anxious, who are certain they are in the most appropriate place, and who value ease of access. Patients who consider their condition to be serious, or who are referred by a HCP are more likely to use the real-time DA to support planning their visit, managing their expectations, and reducing their anxiety. This suggests that for this group of patients, the real time DA is bypassing the 'Health Behaviour' Component of Andersen's model and influencing 'Outcomes' directly, in particular patient satisfaction.

While in the closed questions, patients who value wait-times indicated that they are neither more nor less likely to be influenced by real-time DA for their attendance decisions, the open questions provided more context. As patients (as participants) had no knowledge of actual wait-times, in Andersen's model wait-times can be classified as a *Predisposing Characteristic* representing a patient's beliefs, attitudes or knowledge about health services wait-times. The provision of actual wait-times may change this to an *Individual Enabling Resource (Organisation),* as knowledge of wait-times then enables a more informed choice. Changing a *belief* to an *enabler* is important if the health behaviour being enabled is the ideal behaviour for both the patient and the NHS. The questionnaire data has indicated that those who are more likely to require ED treatment are less likely to be influenced by the real-time DA, indicating that this is expected to be the case.

As Andersen's model is for individual health-seeking behaviour, system-level outcomes are not predicted or explained (Andersen et al., 2014). However information from patients about the potential benefits to the system indicate that they support levelling demand across the system, both through their own (and others) attendance behaviour, and through the NHS using the information to manage demand and resources. In this case, the real-time DA has the potential to influence the component *Contextual Enabling Resource (Organisation).* These are conditions that facilitate or impedes use of services, such as the number and distribution of services, staffing, structure in the community, resources, opening hours, and facilities. While not directly influencing any of these factors, where demand is managed across an urgent care network by both patient behaviour and staff processes, a reduction in crowding, and subsequently in wait-times in ED, acts as a *contextual enabling condition* for those individuals whose attendances are appropriate and necessary.

5.11 Implications for IHAF

Within IHAF, the questionnaire study presented in this chapter has two purposes. The first is to assist with defining the problem, within the principles of QI. This requires considering all stakeholder groups, including patient experience and outcomes, as well as cost savings and operational efficiency. Involving patients as end-users, who are part of the system under investigation, enables an understanding of the current requirements and perceived value for patients to be considered in the design and function of decision-support interventions. This was addressed by asking patients how they perceive that the NHS, as well as themselves, might benefit from a real-time DA application across the urgent care network. Patients were able to see the need to balance supply and demand across the network. They indicated that patient decision-making can contribute to shaping demand which benefits both patients and the NHS, and also that the NHS might use the information to manage demand, for example by appropriately diverting patients to alternative facilities and by managing resources. This has contributed to understanding the problem, and the desire for solutions.

Secondly, the questionnaire explored the implications and the added value to the system of using real-time data applications for patient decision-support, for later synthesis with the formative evaluation using staff interviews (Chapter 7). Observational data found that managers and clinicians agree that it is necessary to understand more about how people with low-acuity conditions use ED. Clinicians were keen that assumptions were not made about who 'needed' to be seen in ED, and that attendance decisions are not coerced by the new real-time information. However patient anxiety and perceived urgency were also believed to influence inappropriate attendance to ED. One finding is that subsets of patients, in particular those who perceive their condition to be serious, and those who were advised to attend ED, recognise value in other uses of the real-time information. This includes planning toward attendances, managing their own and others' expectations, and reducing anxiety. This knowledge may assuage the fears of clinicians who are concerned that real-time information might support sub-optimal attendance decisions, as it seems that patients may use the information to improve their own experience of attendance, without changing their attendance decision. Many studies have suggested that patient self-triage and decision-making regarding ED attendance is reasonable and appropriate based on the information available with which to make a decision (e.g. Nelson, 2011; Breen & McCann; 2013; Chapman & Turnbull, 2016; Cheek et al., 2016; Krebs et al., 2017; Weber, Hirst & Marsh, 2017). The questionnaire has indicated that real time DA, as additional information, can appropriately support these decisions. This addresses early potential risks in implementing IHAF.

The questionnaire results suggest that patients see crowding as a problem which both impacts patients, and to which patient behaviour is also a contributing factor. They see value both for themselves and the NHS in using predictive DA for supporting decisions to reduce crowding. For patients, predictive DA are more likely to enable 'when to go' than 'where to go' decisions. Patients also indirectly have indicated that prescriptive DA can support crowding by augmenting staff decisions which help to manage demand. This informs the next stages of IHAF. This will involve examining and processing the real-time data currently available to patients to support decisions regarding crowding in ED by implementing predictive and prescriptive analytics. The next chapter develops the integrated HM supported by the IHAF framework.

Chapter 6: Application of IHAF – use-case NHS Trust ED

B. Hybrid Model

6.1 Introduction

This chapter applies and tests the hybrid modelling (HM) component of the Integrated Hybrid Analytics Framework (IHAF, refer to Chapter 4), illustrated in Figure 6.1 with the HM components highlighted. IHAF is proposed as a conceptual framework to support the development of a real-time decision-support tool in healthcare. Chapter 5 introduced the case study and the real-time data which is made available from *NHSquicker.* It outlined the problem definition stage, which has a dual-purpose in the framework: to define the problem, and to identify criteria for evaluation. While problem definition is an important stage in the process of all modelling and simulation (M&S) studies, Design Science emphasises the importance of evaluation to support iterations, improvements and similar future work, to ensure that the modelling process starts with the assumption that the model will be useful in practice. For this reason, the problem definition stage must also consider the criteria and influencing factors for evaluation. For example, it should start with a system-level understanding of what matters in practice, to reduce the risk of unintended consequences in a complex system. The design of the model and its output needs to be considered, such that it is not just useable, but useful in practice. It should support situation awareness (SA), which is an important component of short-term decision-making, and it should consider barriers to implementation of such an approach early in the design process.

Chapter 5 concluded that real-time data analytics (DA) has the potential to contribute to reducing ED crowding by influencing both patient health-seeking behaviour (through availability of real-time and predicted wait-time data), and staff decision-support (through the HM). Information from patients about the potential benefits to the system indicate that patients and the public support levelling demand across the urgent care system. This can take place through their own (and others) attendance behaviour, and by providing information to the NHS to manage their demand and resources. Knowledge gained from patients about health-seeking behaviour may be able to assuage the fears of clinicians who expressed concern that real-time information might encourage patients to make sub-optimal attendance decisions.

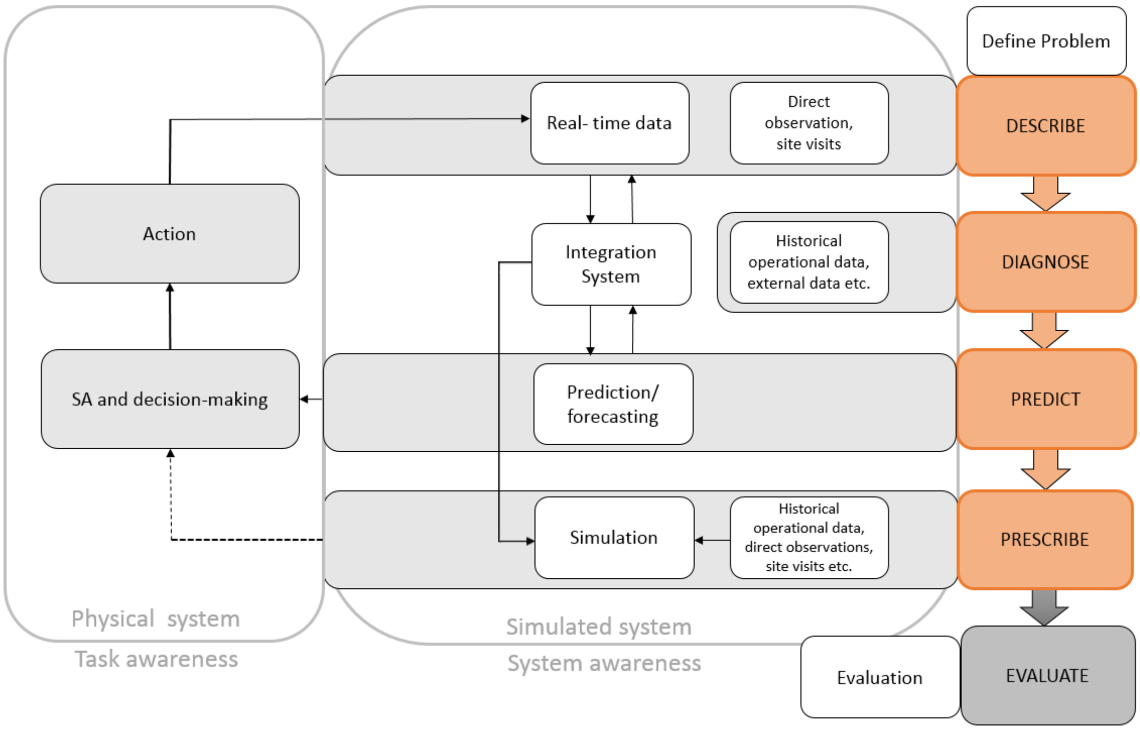


Figure 6-1Integrated Hybrid Analytics Framework (IHAF) with HM components highlighted

Combined with Chapter 5, the HM in this chapter addresses the second objective of the second research question, to apply the framework within the case study at an NHS ED. The stages of the HM will be described and applied. The sections following address each of the highlighted components of IHAF in Figure 6.1.

Section 6.2 describes the architecture of the hybrid model. Section 6.3 outlines the *Describe* component of IHAF, the real-time data and it's pre-processing. Section 6.4 describes the *Diagnose* component of IHAF, and looks at measuring crowding, given the existing data, and defining the simulation trigger. Section 6.5 describes the *Predict* component, outlining the development of the forecasting methodology, SARIMA time-series forecasting. Section 6.6 describes the *integration* of the components and Section 6.7 is the *Prescribe* component of IHAF, the simulation model, and a set of defined scenarios for decision-support. The next section conceptualises the architecture of the hybrid model application.

6.2 Conceptualisation of the hybrid model

Figure 6.2 illustrates conceptually the HM component, encompassing an urgent care network with at least one ED and one MIU to support knowledge about capacity in alternative facilities, and access to real-time data feeds and historical data. These are comparable to those provided by *NHSquicker* and ED operational data, for creating forecasts and populating the simulation model. The architecture consists of:

(a) The implemented near real-time data component (*NHSquicker*), with historical data for developing forecasting models. This approach has been applied in healthcare for forecasting ED crowding in real-time up to 8 hours ahead (Hoot et al., 2009), while Barnes et al. (2015) showed how real-time predictions of inpatient length-of-stay might be used for discharge prioritisation.

(b) Historical operational data from the urgent care network to populate the simulation model, and data inputs that are not available in real-time, such as patient acuity. Model constraints can be imposed using this data (Adra, 2016).

(c) Data pre-processing for moving window analyses as new data is received. For example, Boriboonsomsin et al. (2012) integrated historical and real-time traffic information from multiple sources to reduce the environmental impact of road travel.

(d) Time-series forecasts creating predictions up to four hours into the near-future. This short window allows the forecasts to retain maximal accuracy, while providing adequate time to trigger the execution of intervention scenarios through the real-time DES model. Xu and Chan (2016) found using an analytical approach that even noisy predictions of ED arrival counts can successfully be used to improve ED performance through patient re-direction. Lin and Chia (2017) used ARIMA forecasts of patient arrivals as inputs into a DES model to optimize staff rosters, which improved patient waiting times in the simulation results.

(e) A simulation trigger, given a specific decision rule. Most applications of real-time simulation use a reactive approach for triggering scenario what-if analysis, however Aydt et al. (2008b) described triggering based on forecasts. Bae et al. (2004) showed how the automatic execution of processes using Event-Condition-Action rules can be automatically triggered by an active database without user intervention.

(f) A set of predefined scenarios, including diverting low-acuity patients to alternative facilities. This approach was found to be successful using analytical methods (Xu & Chan 2016). However other scenarios, for example derived from ED escalation policies, can also be explored. For example, with the aim of reducing ED overcrowding, Nahhas et al. (2017) used simulation to explore a range of scenarios, such as flexible treatment rooms, flexible staff activities and flexible shifts.

(g) The DES model to test scenarios, which is initialised using both real-time and historical data (e.g. Espinoza et al., 2014; Oakley et al., 2020).

(h) Information provided to decision-makers to support short-term planning for reducing overcrowding.

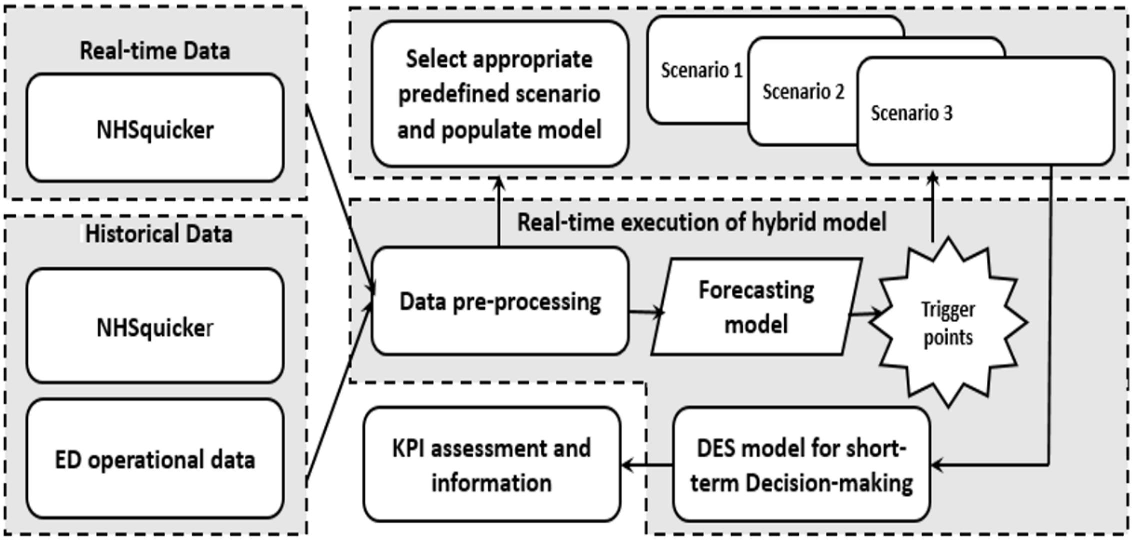


Figure 6-2 Conceptual framework of the HM component

6.3 Descriptive analytics: Real-time data

The *Descriptive* component of IHAF describes the real-time and historical data required by the application, its presentation and pre-processing, as described in Chapter 4 (Section 4.5.1). Chapter 5 outlined the real-time data made available for this application by *NHSquicker* (Mustafee et al., 2018). The use-case is one hospital which contributes real-time data to *NHSquicker* from its ED and its three MIUs: Newton Abbott MIU, Totnes MIU and Dawlish MIU. These four facilities form an urgent care network (UCN). The public have access to near real-time *NHSquicker* data (updated between 5-15 minutes) for supporting attendance decisions. The data has been made available historically with three variables: (i) The total number of patients in each department ('*Total Patients*'); (ii) The number of patients waiting to be assessed by a clinician ('*Patients Waiting*') (iii) The maximum wait time to be assessed by a clinician ('*Maximum Wait*'). The data from each facility in the UCN is pushed to a URL, where it is downloaded every 30 minutes, to be made available for analysis. Data is available from 03/01/2018.

Examining *NHSquicker* data, data blackouts, either at the hospital level (an error in the hospital computer sending the data) or by the client computer downloading the data, require data pre-processing to deal with missing data. For historical data, this was interpolated using the average of the previous four equivalent times of day and day of week (Figures 6.3 – *Total Number*; Figure 6.4 – *Patients Waiting*; Figure 6.5 – *Maximum Wait*. Available data is blue; interpolated data outages are orange).

Maximum Wait data contains significant data quality outliers of up to 1222 minutes. These were concluded to be data errors from the hospital sending the data, and an upper limit of 400 minutes was fixed. Rarely would a wait of this duration occur; in one year the data indicated that it happened 10 times. This was validated in consultation with one senior NHS manager, who indicated that waits longer than 60 minutes were a cause for concern in practice.

However, from the data, it is clear that that waits longer than 60 minutes are very common (Figure 6.5), and the actual upper limit is difficult to determine. This is because performance reporting isn't required for this measure, meaning data quality issues are not a priority. Nonetheless, the number of 12-hour waits doubled in 2019 compared with 2018 at the national level, a clear sign of EDs under pressure (House of Commons, 2020).

A further possible explanation is that 2-3% of patients (nationally) leave before treatment commences without informing staff, hence their departure is not recorded in data systems at the time of leaving (NHSDigital, 2019a). As these patients are never actually seen by a clinician, their recorded wait-time may run to many hours.

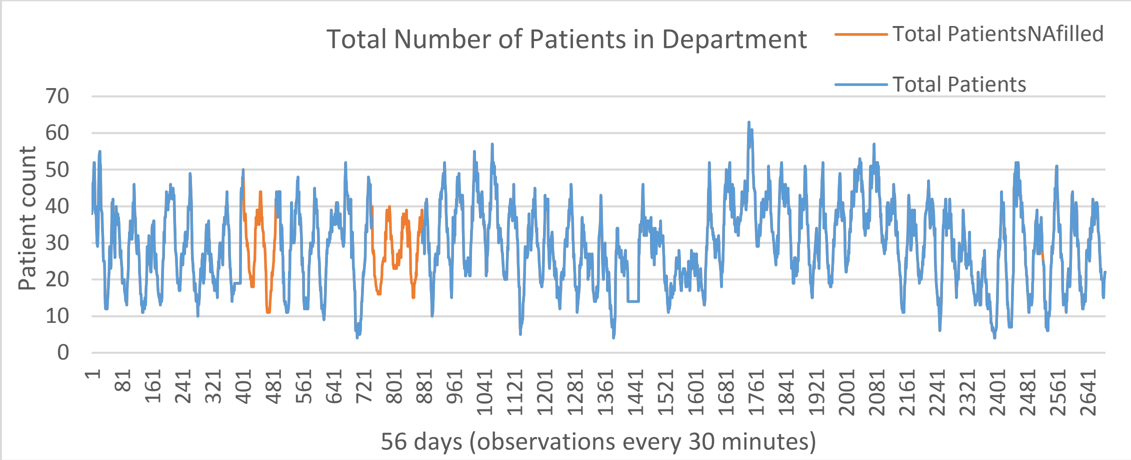


Figure 6-3 Time series plots for 56 days of data for Total Patients with missing data filled. Blue = available data. Orange = data outages interpolated using moving averages

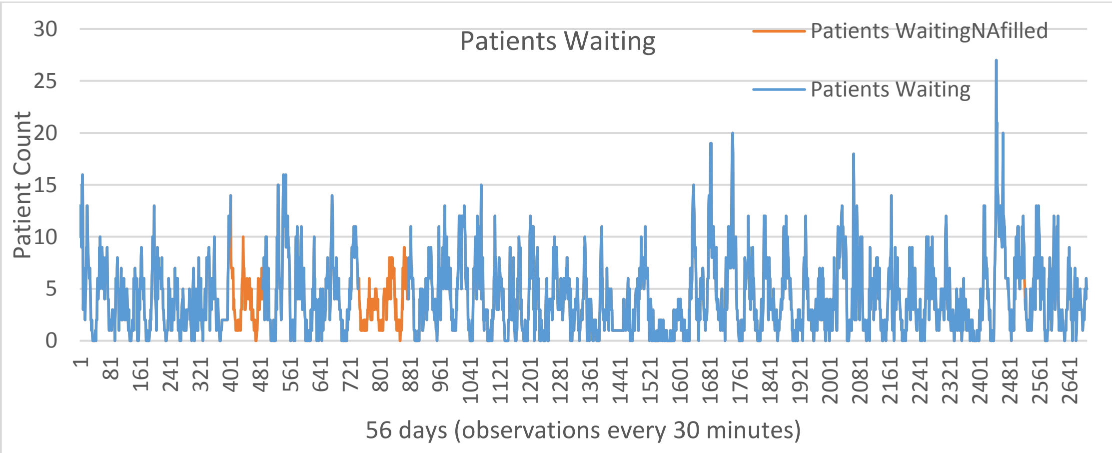


Figure 6-4 Time series plots for 56 days of data for Patients Waiting with missing data filled. Blue = available data. Orange = data outages interpolated using moving averages

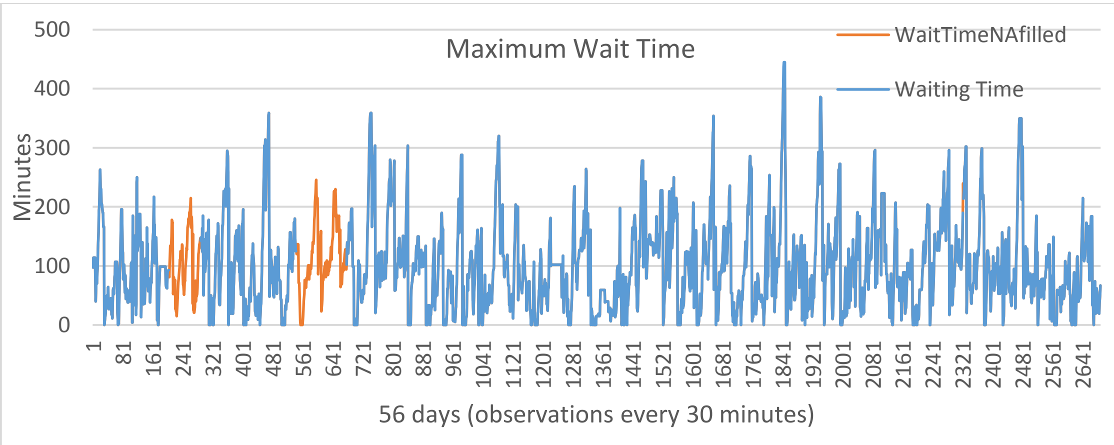


Figure 6-5 Time series plots for 56 days of data for Maximum Wait Time with missing data filled. Blue = available data. Orange = data outages interpolated using moving averages

For later exploratory analysis, historical *NHSquicker* data was partitioned into day of week, week in year, week in month, day of week in month, day in year, and day in month. To determine a trigger for the simulation model, historical hospital ED arrivals data (described in Section 6.4 and Appendix 3) was aggregated into 30-minute batches and mapped with *NHSquicker* data. Lags (a fixed amount of passing time) were created of 1, 2 and 4 hours for half-hourly arrivals and *Total Patients* data. The next section forms the '*Diagnose*' component of the IHAF framework, and explores the data to determine how to forecast crowding, and trigger thresholds for the simulation.

6.4 Diagnostic analytics: What to forecast, and when to trigger

Diagnostic methods are exploratory, and focus on processes and causes, using methods such as correlation, cluster analysis, and root cause analysis. A key use of diagnostic analytics in this framework is determining the conditions for triggering the real-time simulation. In this application of the framework based on a use-case, a preventative trigger is used (Chapter 4, Section 4.5.3), such that the purpose of the HM is to prevent the critical situation from arising in the first place. A preventative trigger is observed in forecasts, and therefore is limited to conditions that can be forecasted (Aydt, 2008b), however a proactive trigger (at regular intervals) or a reactive trigger (in real-time) are alternative methods.

Crowding measures usually take into account triage categories and resources such as beds and staff (Hoot et al., 2007). Using *NHSquicker* data, a proxy measure of crowding is required. Diagnostic analytics are used in the following subsections to explore this and determine an appropriate trigger for the simulation model using the available real-time data. Like predictive and prescriptive analytics, diagnostic analytics requires domain knowledge, and may require additional external information.

In addition to *NHSquicker data*, three pseudonymised historical hospital datasets were made available by the use-case hospital (summarised in Appendix 3 Table A3.1, and in Figure 6.6):

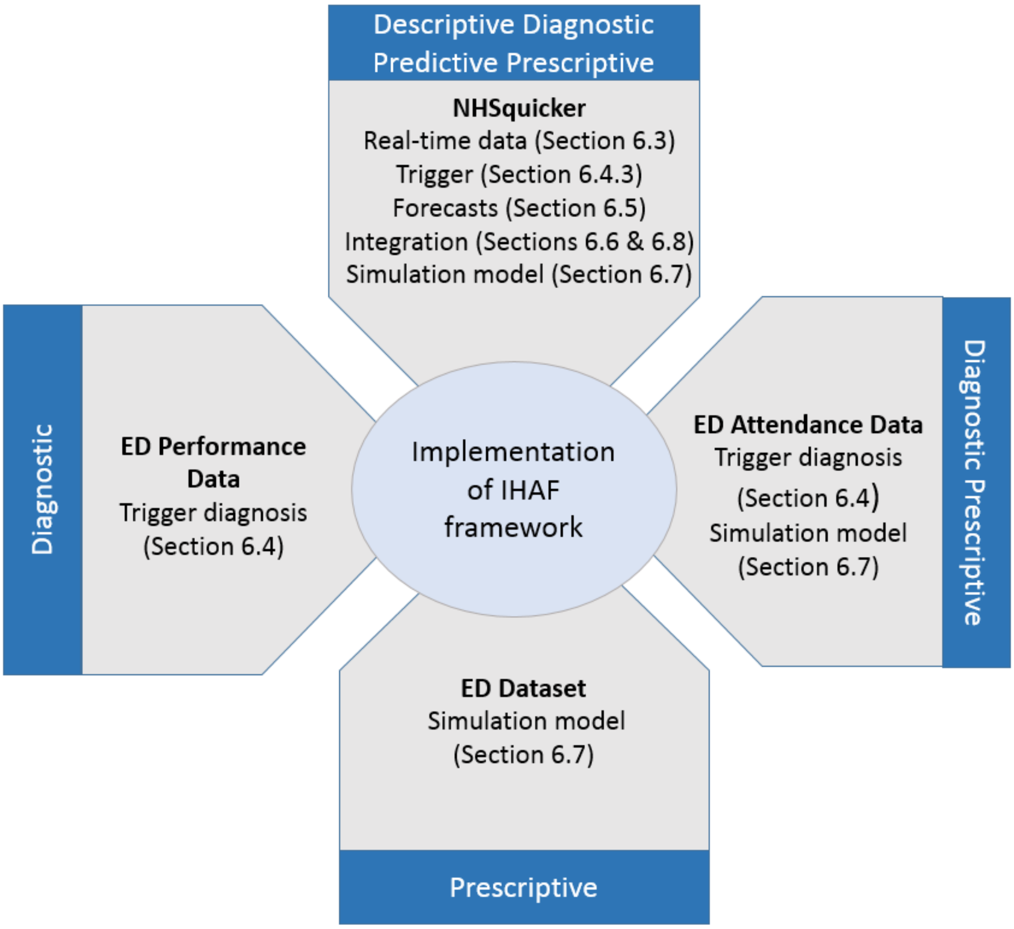


Figure 6-6 ED datasets available for IHAF implementation

(i) *ED Performance Data*:- Daily ED attendance, daily compliance with the 4 hour target, patient overall stays under 2 hours, under 4 hours, and over 4 hours, daily admissions, bed delays, trolley delays, and GP referrals, from 01/04/16 to 19/09/18;

(ii) *ED Attendance Data:-* Attendances dates and times, age, gender, triage category, disposal, diagnosis from 01/04/16 to 29/10/18;

(iii) *ED Dataset:-* (01/08/15 – 31/07/16) Fields include triage category, visit duration, place of discharge/admission, reason for departure, reasons for departure delay, arrival by ambulance/walk-in, and details of treatments and investigations and was used to support the simulation model development and validation.

The next section (6.4.1) examines the ED Performance data for insights, and Section 6.4.2 examines *NHSquicker* data and ED attendance data to determine an appropriate trigger. Appendix 3 contains additional material, and is indicated where appropriate.

6.4.1 *ED performance data*

As part of the *Diagnostic* component of IHAF, the first hospital dataset was examined initially to explore the relationships between patient numbers and wait times. For each day, a 4-hour compliance (%), total attendances, number who waited <2 hours, <4 hours, and >4hours is provided. Numbers in ED "length of stay" (LoS) categories were converted to proportions, and a 2-4 hour LoS was calculated. Relationships in the datasets were examined using correlations.

As the daily proportion of those with LoS less than 2 hours, and those with LoS 2-4 hours increases, the compliance with the 4-hour target increases. A correlation measures the extent to which two variables are related. This is done by calculating the standardised covariance, using Pearson's coefficient (r), which requires only that the data are interval for it to be an accurate measure of the linear relationship between two variables. Without assuming causality, a linear relationship is observed between performance against the 4-hour target and the proportion waiting less than 2 hours (r = 0.46), and the proportion of patients waiting between 2-4 hours (r = 0.7). In other words, performance against the target decreases, as the proportion of LoS less than 4 hours decreases. This is to be expected, as the remainder wait for greater than 4 hours.

Looking at actual counts of patient length of stay data (Figure 6.7), the same pattern can be seen; again the proportion waiting 2-4 hours has a stronger relationship with performance against the 4 hour target. Figure 6.7 is a scatterplot of patients who have waited less than four hours, against compliance with the four hour target. Patients who have waited less than four hours are subdivided into busy days and quiet days, and again into those who waited less than 2 hours and those who waited 2-4 hours.

An average of 205 patients were attending the department daily during the time period covered by the data. Filtering the data into above average (>204 patients) and below average (<205 patients) attendance finds that the counts of those attending between 2-4 hours on busy days (above average attendance) is most strongly correlated with performance against the 4 hour target (r = 0.73). The implications of this are that the busier the department gets, the more impact this group of patients have on the performance target. This confirms that the number of patients in the department directly impacts on waiting times. This is further investigated using ED attendance and *NHSquicker* data.

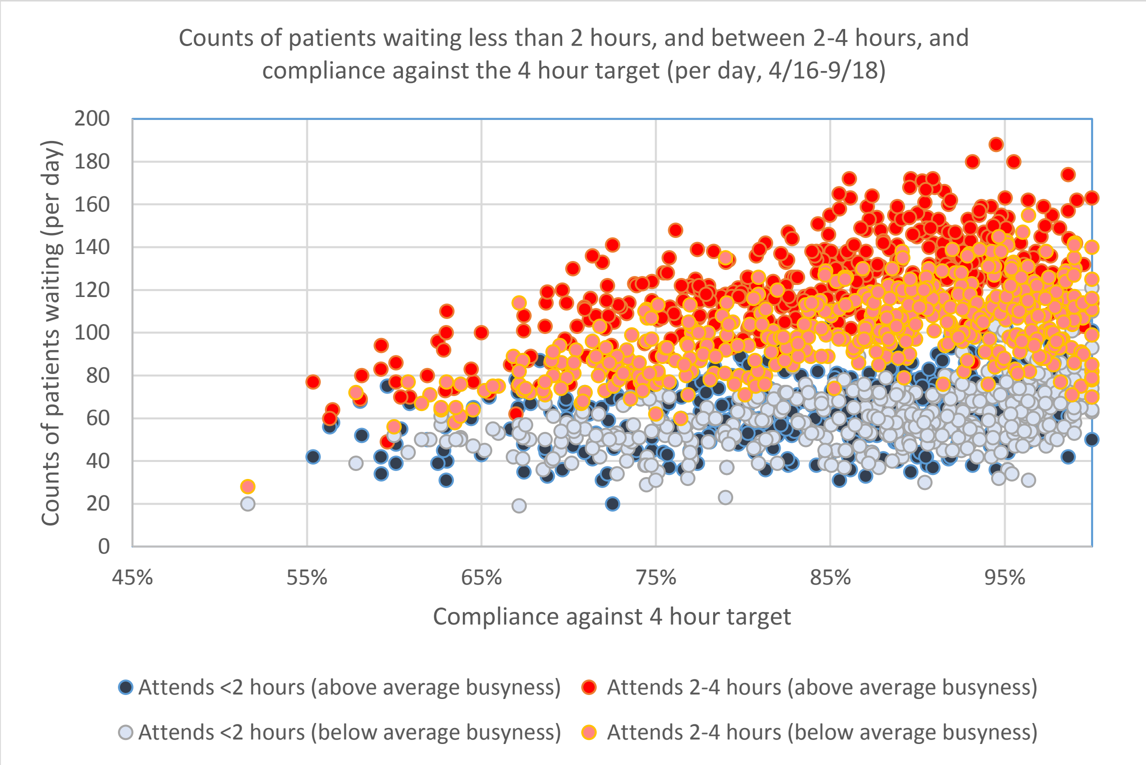


Figure 6-7 Counts of patients waiting less than 2 hours, and between 2-4 hours, and compliance against the 4 hour target

6.4.2 *NHSquicker data and ED attendance data*

Figure 6.8 shows a visual correlation between *NHSquicker* data *Total Patients* and *Patients Waiting* (03/01/18 – 17/01/18). The plot shows 14 days of 48 data points per day, plotted every 30 minutes.

A seasonal pattern exists when a time-series is influenced by calendar-related factors, for example, the month, day of the week, or hour of the day (Hyndman, 2011). Seasonality is always of a fixed and known period. The seasonal (24 hour) variance in *Total Patients* is daily (range = 3-63 patients), and has a larger magnitude than *Patients Waiting* (range = 0-27 patients).

As *Patients Waiting* contains smaller discrete numbers (mean = 4) compared with *Total Patients* (mean = 28), there is a less clear seasonal pattern in the *Patients Waiting* time-series as a change of one patient represents a relatively much larger shift. The implications of this are that time-series forecast models from *Total Patients* data are likely to be more accurate than a model using *Patients Waiting* data.

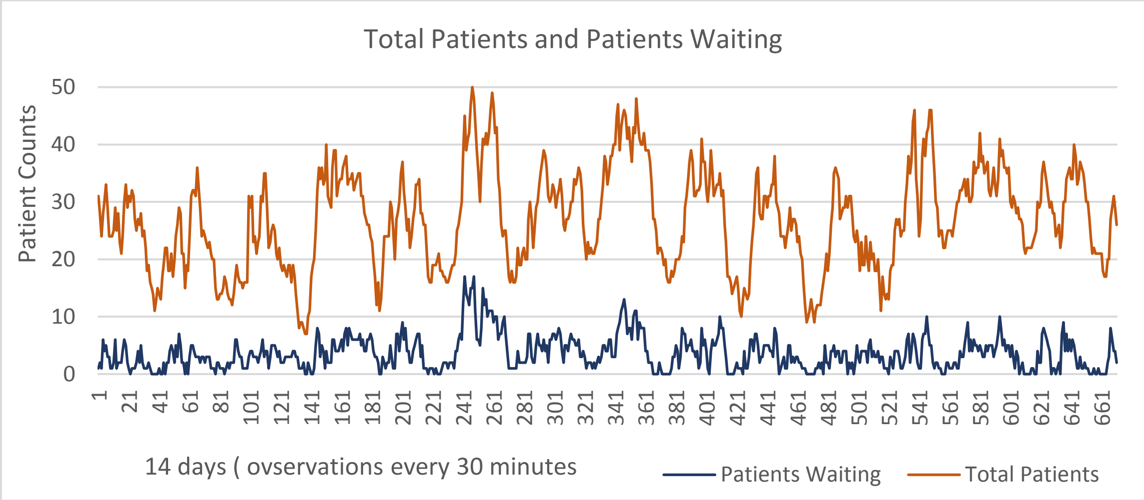


Figure 6-8 Time series of Total Patients and Patients Waiting over 14 days (30 minute observations)

Although a daily seasonal pattern is exhibited, *Maximum Wait Time*, as a maximal measure, is highly volatile, with frequent extreme drops as individual patients are assessed for treatment. While some correlation is visualised with the *Total Patients* series over the same time period (Figure 6.9, <u>note dual axis</u>), a lag is present such that wait times appear to peak a short time after total patient numbers.

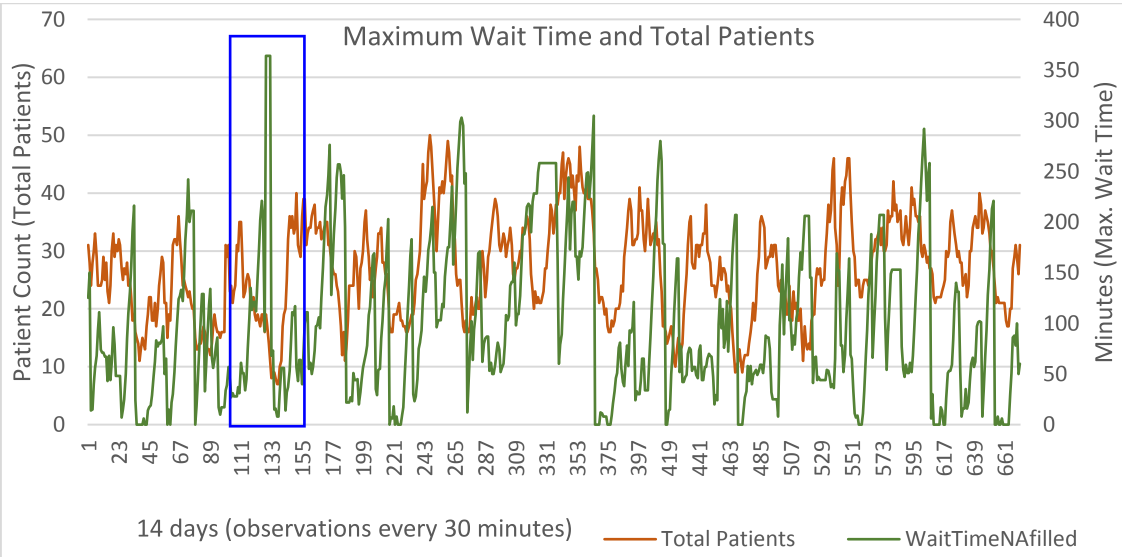


Figure 6-9 Time series of Total Patients and Maximum Wait Time over 14 days (30 minute observations)

In Figure 6.9, the blue rectangle illustrates this. *Total Patients* is peaking at approximately observation 115 with 36 patients, while wait times remain below 50 minutes. However by observation 130, the wait time has increased to 360 minutes.

Half-hourly arrivals (historical ED arrivals data, batched into 30 minute arrivals), (03/01/18 – 17/01/18), mapped with *NHSquicker data* (Figure 6.10, note dual axis) have a similar regular daily seasonality but are volatile in the short-term, and *Total Patients* peak sometime after arrivals. The rectangle in Figure 6.10 illustrates this, for example arrivals peak at observation 142, with *Total Patients* peaking at observation 150. Although *NHSquicker* does not have access to arrivals data, the relationship between these two variables is of interest, as patient arrivals directly lead to crowding.

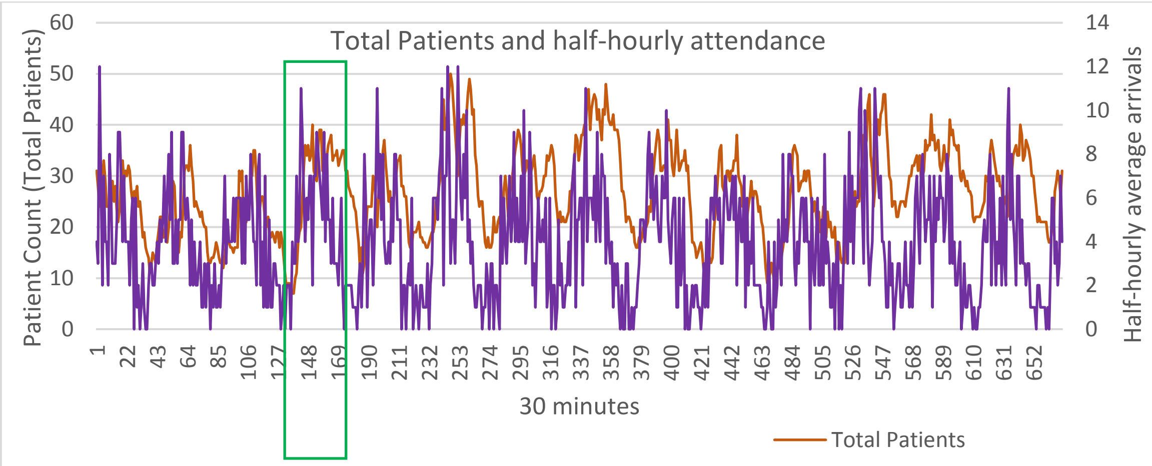


Figure 6-10 Time series of Total Patients and ED arrivals (30 minute observations) over 14 days

In summary, the time-series data show that patient arrivals peak before the *Total Patients* in the department, which in turn peaks before the *Maximum Wait* time. The number of *Patients Waiting* is correlated with *Total Patients*.

The relationships are further explored. Pearson's correlation coefficient *r* for *Total Patients* and *Patients waiting* is *r*=0.67. The effect size of *r* varies from 1 to -1, and 0.67 is considered a large effect size, with a strong positive correlation (Cohen, 1992).The coefficient of determination, R^2, is a measure of the amount of variability in one variable that is explained by the other, and is obtained by squaring the correlation co-efficient *r.* A simple linear regression model (using open-source software, R 3.6.1) investigates the relationship between *Patients Waiting* and *Total Patients*. A linear regression model is in the form:

$$Yi = linear\ model\ plus\ noise = (\beta 0 + \beta 1xi) + \epsilon i$$

The error term (ϵi) represents variables not considered in this simple model. Regression models assume that: a) the errors are normally distributed and, on average, zero; b) the errors all have the same variance (they are homoscedastic), and c) the errors are unrelated to each other (they are independent across observations). While the third assumption can be relaxed slightly (Stewart, 2016), the first two assumptions are checked using a histogram and Q-Q plot (a graphical technique for determining if two data sets come from populations with a common distribution) of the residuals. The scatterplot is seen in Figure 6.11 and the output from R is in Appendix 3 (Figure A3-1):

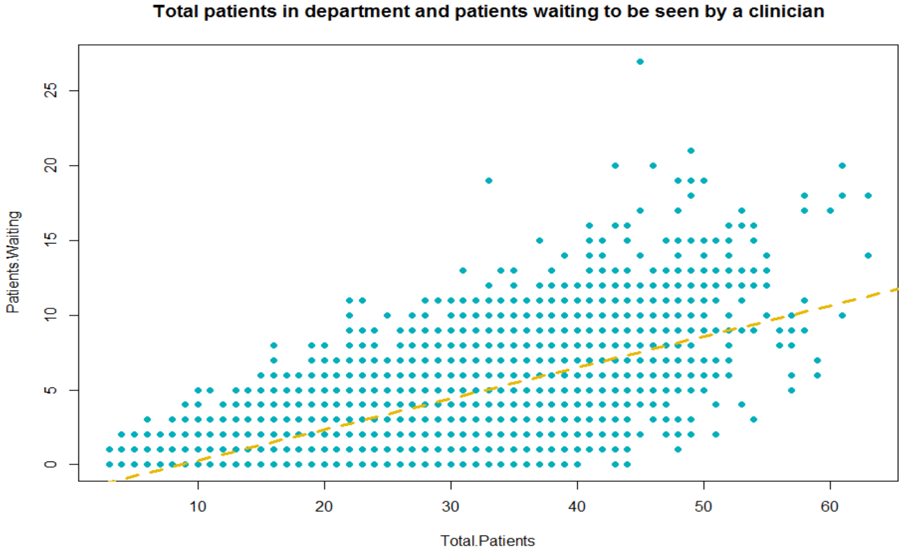


Figure 6-11 Scatterplot and line of best fit of Total Patients and Patients Waiting

From the output, *Patients Waiting = 0.208*Total Patients + -1.836*, and this is significant to $p<0.01$. The R^2 is 0.4554. This provides confidence that information about the *total number of patients* in the department is an indicator of the *number of patents waiting to be seen by a clinician*. However the model violates the assumptions of a regression model (see Appendix 3), so while it indicates a relationship between the two sets of data, it doesn't allow predictions to be made. The residuals are checked (Appendix 3, Figure A3.2) and are not normally distributed, with a systematic departure from normality in the maximum quartile of the Q-Q plot.

This effect can be explained by the busyness in the department. As the total number of patients in the department rises and the demand-capacity mismatch increases, a build-up of low-acuity patients in the waiting area can occur if patients of higher urgency are present, as they will be treated with priority. The unpredictable nature of urgency levels in the department shifts the variance away from homoscedasticity as the total patients increase. Nonetheless, there remains a clear relationship. The relationships are broken down at different thresholds of *Total Patients* in Table 6.1.

Table 6-1 Correlations between Total Patients and Patients Waiting at different partitions of Total Patients

Total Patients	Correlation coefficient *r*	R^2 (sig)	Range of residuals
<31	0.5	0.245 (p<0.01)	-3.9870 - 8.3077
>30	0.5	0.251 (p<0.01)	-7.9685 - 18.9082
<36	0.54	0.296 (p<0.01)	-4.9552 - 14.3869
>35	0.45	0.205 (p<0.01)	-8.1104 - 18.8772
<41	0.58	0.344 (p<0.01)	-5.9408 - 14.3021
>40	0.34	0.111 (p<0.01)	-8.1094 - 18.7893
<46	0.63	0.392 (p<0.01)	-6.9814 - 19.8265
>45	0.24	0.0556 (p<0.01)	-8.3649 - 11.3900

Correlations are measured at different partitions of *Total Patients*. By examining the table, and both the main scatterplot (Figure 6.11) and the scatterplot of the residuals (Appendix 3, Figure A3.3), increasing variance begins at around 40 *Total Patients*. However these may be due to the effects of having a natural 'floor' on the data, such that the larger the dataset (at <46 *Total Patients*), the more likely it will detect the correlation in the dataset as a whole, while there is clearly increasing variance as the *Total Patients* increase.

While the range of the residuals of the linear regression model is -0.496 – 14.39 at <36 *Total Patients,* there is one obvious outlier; without this the range would be similar to <31 patients. This is the same at <41 patients. However by <46 *Total Patients*, several more outliers have crept into the residuals at the higher numbers of *Total Patients* (plotted in Appendix 3, Figures A3.4 and A3.5). This indicates that an appropriate trigger is between 40 and 45 predicted *Total Number of Patients* in the Department. Histograms and Q-Q plots of the residuals again confirm heteroscedacity and non-normality. However the increasing variance visualised between 40-45 *Total Patients* on average suggests that problems with crowding start around here.

Total Patients from 3/1/18 - 2/2/18 are mapped against daily 4-hour compliance data to look for insights that might confirm this (Figure 6.12). While this is a crude measure, it is possible to visualise that on days where *Total Patients* stayed below 40, compliance tends to be high, but on days where *Total Patients* exceeded 40 at any point, compliance appears to drop, often significantly, illustrated in the red rectangles.

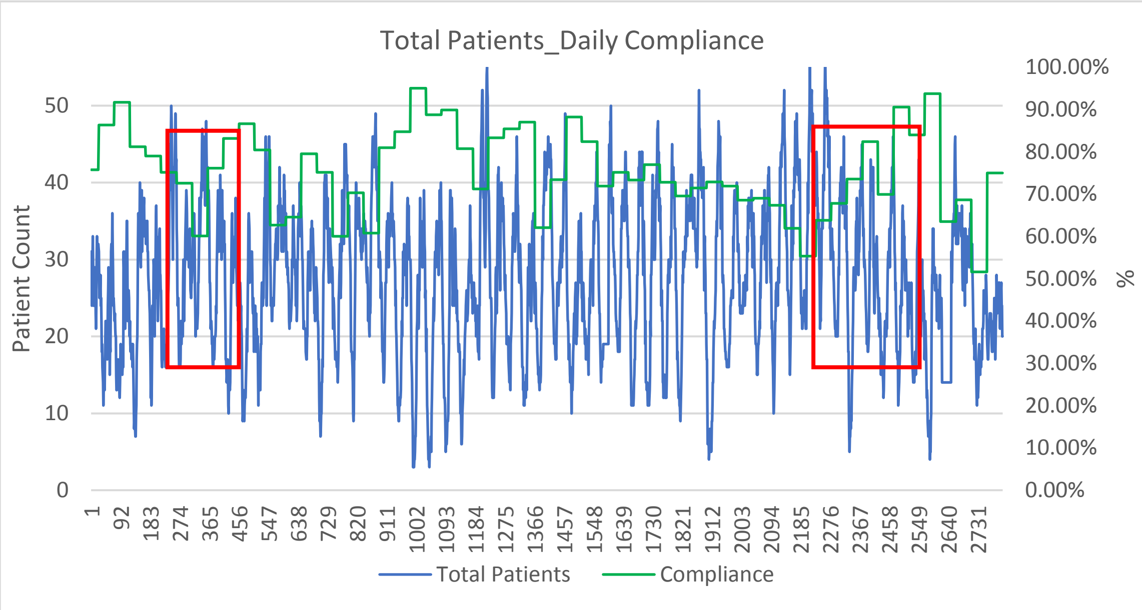


Figure 6-12 Total Patients (every 30 min) and Daily Compliance with the 4-hour target (24 hourly)

Table 6.2 shows correlation coefficients (*r*) between each of *Maximum Wait, Total Number,* and *Patients Waiting,* and *half hourly Arrivals.* These are calculated with 1, 2, 3, and 4 hour arrival lags. The correlation between *Total Number* and *Arrivals* at 1, 2 and 3 hour lags are similar (r = 0.55-0.57) while *Patients Waiting* is most closely related to the arrivals 1 hour ago (r = 0.55).

In Appendix 3, Figure A3.6 shows a scatterplot of *Total Number* of patients with *arrivals* 1, 2 and 3 hours ago, for visualisation. This corresponds with the data in Table 6.2 (columns 1 and 3), showing a correlation between the *Total Patients* and arrivals at 1-3 hours previously (shown in bold).

Table 6-2 Correlations between Maximum Wait, Total Number, and Patients Waiting, with lagged half hourly Arrivals at 0-4 hour lags.

Patient arrivals (all triage categories)	Maximum Waits (*r*)	Total Number (*r*)	Patients Waiting (*r*)
Arrivals	-0.14	0.39	0.42
Arrivals 1 hour previously	-0.06	**0.55**	**0.55**
Arrivals 2 hours previously	0.03	**0.57**	0.47
Arrivals 3 hours previously	0.11	**0.55**	0.41
Arrivals 4 hours previously	0.16	0.49	0.32

Table 6.3 shows correlation coefficients (*r*) between *Maximum Wait* and *half hourly Arrivals*, with lags of 1-4 hours. The correlation between *Maximum Wait*, and *Total Patients* at 1, 2 and 3 hour lags are similar (r = 0.35) while *Patients Waiting* is most closely related to the current arrivals (r = 0.67), as seen previously. In Appendix 3, Figure A3.7 is a scatterplot of *Maximum Wait time* and *Total Number* of patients in the department 2, 3 and 4 hours ago for visualisation, corresponding with columns 1 and 2 in Table 6.3 (shown in bold).

Table 6-3 Correlations between Maximum Wait, Patients Waiting, and lagged Total Patients at 0-4 hour lags

Total number of patients in department	Maximum Wait (*r*)	Patients Waiting (*r*)
Total	0.22	**0.67**
Total 1 hour previously	0.31	0.53
Total 2 hours previously	**0.35**	0.40
Total 3 hours previously	**0.35**	0.32
Total 4 hours previously	**0.35**	0.16

Patient Arrivals influence the *Total Patients* in the department with a one-to-three hour lag. In turn, *Total Patients* influence the *Maximum Wait* time with a two-to-four hour lag. The number of *Patients Waiting* is correlated with *Total Patients*. *Total Patients* data appears to be the most useful data for acting as a proxy for crowding and will therefore be used for the predictive component in this implementation of IHAF (Section 6.5). The predictions will be used to trigger the simulation (Section 6.7). Considering the average of *Total Patients*, crowding appears to occur around 40 *Total Patients* in the department. However this will vary throughout the day, as resources also vary. For this reason, the next section will investigate how the simulation trigger can be time-varying across a 24-hour period.

Implications for IHAF implementation
Using *NHSquicker* data, *Total Patients* appears to be the most useful data for acting as a proxy for crowding and will therefore be used to make predictions (Section 6.5), and as a predictive trigger to initiate the simulation (Section 6.7).

6.4.3 *Time-dependent trigger*

A trigger defines an action or set of actions that are executed when an event occurs. The previous section identified a simulation trigger at an average of 40 total patients in the department. However this is likely to vary across a 24-hour period, as resources vary. This is investigated initially by calculating the mean and the standard deviation (StD, +/- 1, 1.5, 2) for each *Total Patient* dataset per hour of day [5500 observations (115 days)]. These are shown in Table 6.4.

Table 6-4 Standard Deviations (StD) of Total Patients per hour

Hour	00.00	01:00	02:00	03:00	04:00	05:00	06:00	07:00
1StD	38	35	32	29	27	25	24	24
1.5StD	41	39	36	33	30	29	28	27
2StD	45	43	40	37	34	32	32	31
Hour	08:00	09:00	10:00	11:00	12.00	13:00	14:00	15:00
1StD	25	28	32	35	38	40	42	43
1.5StD	28	31	35	39	42	44	46	47
2StD	32	35	38	42	46	48	50	51
Hour	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
1StD	42	42	43	44	44	44	43	41
1.5StD	46	45	47	48	48	48	47	45
2StD	50	49	51	52	52	52	51	49

To illustrate, *Total Patients* are plotted as a continuous line graph to assist visualisation (Figure 6.13) on a subset of data from 00:00 to 00:59 over a 115 day period, and as a scatterplot (Figure 6.14) against *Patients Waiting*. As with the full dataset (all hours), the scatterplot exhibits increasing variance around the mean as *Total Patients* increase.

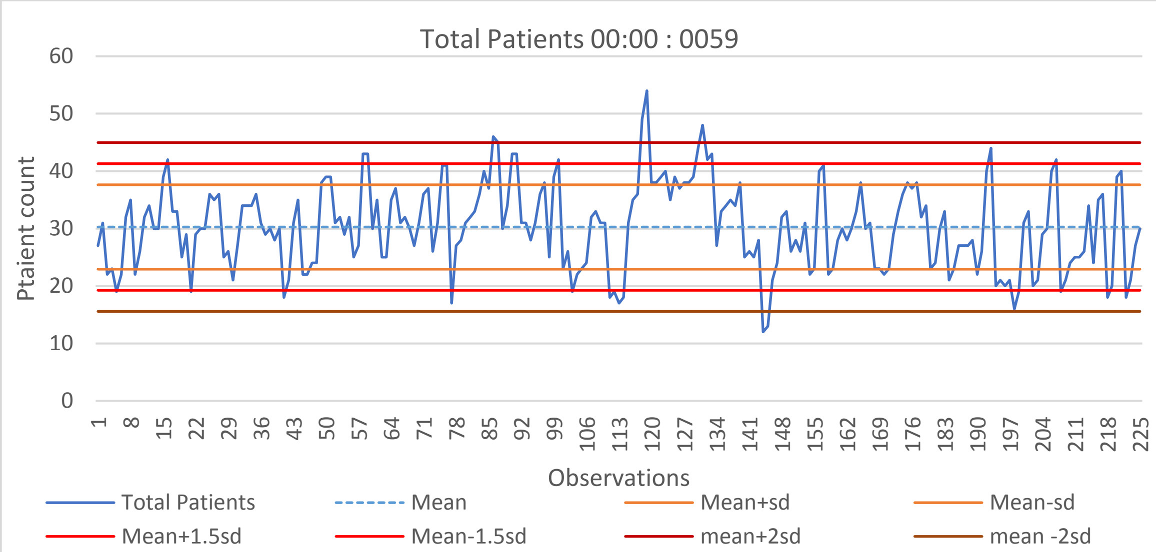


*Figure 6-13 115 days of 00:00 to 00:59 with mean and SD, 1.5*SD, 2*SD*

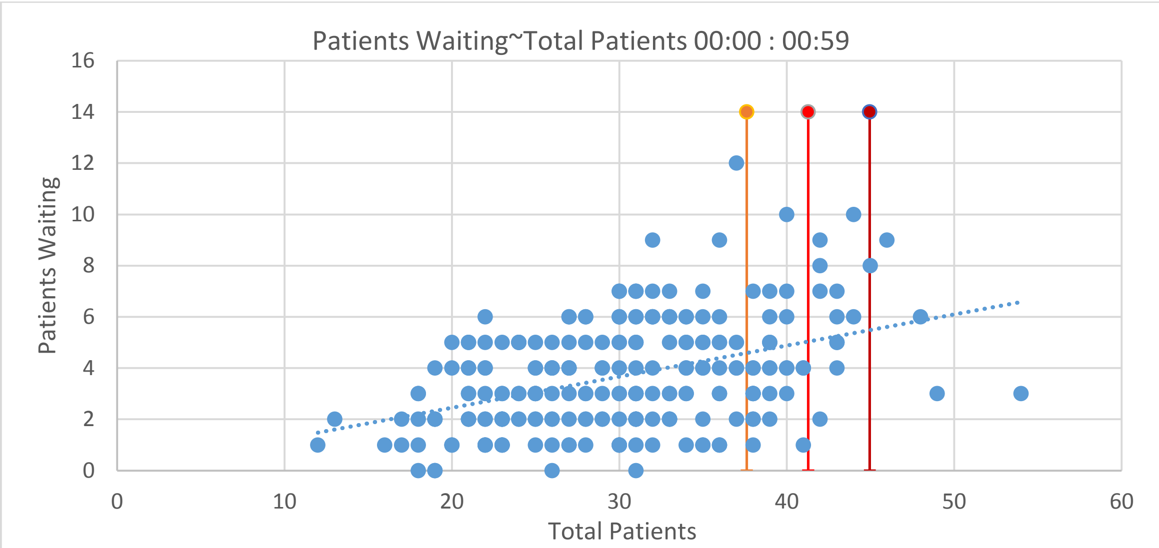


*Figure 6-14 Scatterplot of Patients Waiting and Total Patient 00:00 to 00:59, with SD, 1.5*SD, 2*SD for Total Patients*

These plots are replicated in Appendix 3 for 12:00 to 12:59 (Figures A3.8 and A3.9). In both cases, it's possible to see that the variance starts to increase at 1-1.5 standard deviations from the mean.

The time period 00:00 to 00:59 is plotted in Figure 6.15 against the daily compliance data to look for further insights. This is replicated for the time period 12:00 to 12:59 in Appendix 3 (Figure A3.10). Even taking these one hour snapshots (plotted as a single line to assist visualisation, <u>note dual axis</u>) a relationship between *Total Patients* and *daily 4-hour compliance* is clear.

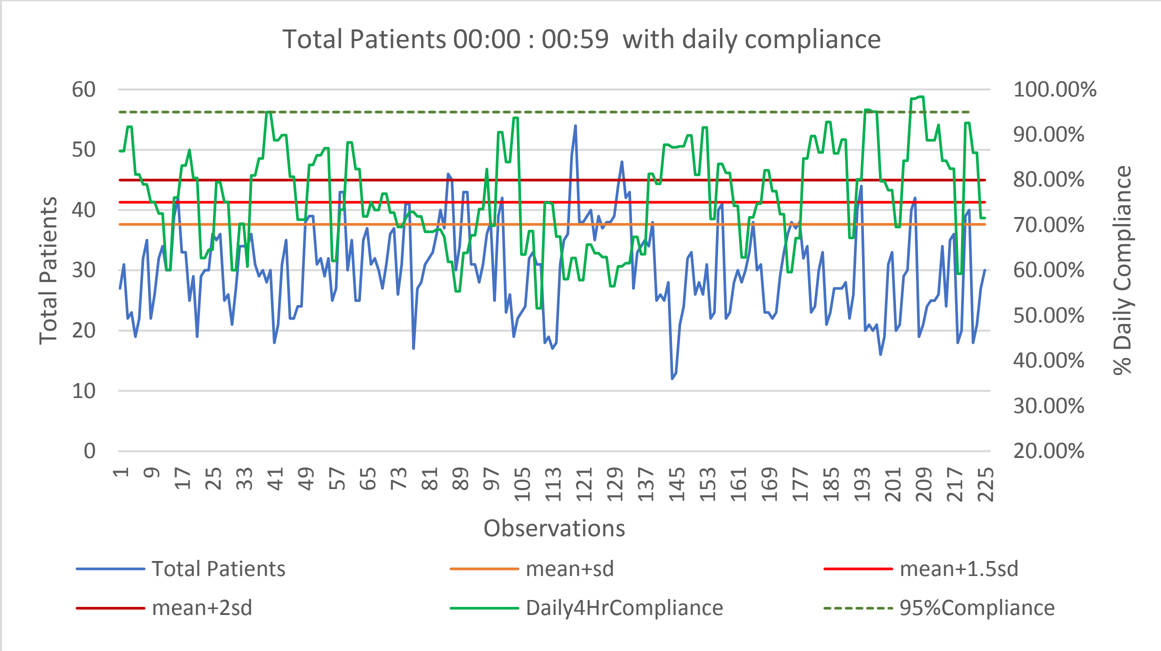


Figure 6-15 Total Patients (00:00 to 00:59) and Daily Compliance with the 4-hour target (24 hourly)

Based on the visual information in these plots, a time-dependent trigger of 1.5 StD is chosen. This is an average trigger of 39 across a 24-hour period, which is realistic, and is represented in Table 6.5.

Table 6-5 Hourly trigger at 1.5 standard deviations

Hour	00.00	01:00	02:00	03:00	04:00	05:00	06:00	07:00
1.5StD	41	39	36	33	30	29	28	27
Hour	08:00	09:00	10:00	11:00	12.00	13:00	14:00	15:00
1.5StD	28	31	35	39	42	44	46	47
Hour	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
1.5StD	46	45	47	48	48	48	47	45

Based on the analysis in this section, *Total Patients* data is chosen for forecasting as a proxy for crowding, and a 24-hour time dependent trigger is proposed. This forms the *Diagnose* component of IHAF. The next section (Section 6.5) implements the forecasting methodology, and forms the *Predict* component of this application of the IHAF framework.

Implications for IHAF implementation

Using *NHSquicker* data, *Total Patients* is the most useful data for acting as a proxy for crowding and will therefore be used to make predictions (Section 6.5). These will form the predictive trigger to initiate the simulation (Section 6.7). In this implementation, the predictive trigger is time-dependent over a 24-hour period, as available resources vary throughout the day.

6.5 Predictive analytics: Time-series forecasting of Total Patients

In the IHAF framework the purpose of the predictive component is to predict the onset of a critical event, such that subsequent decisions are ‘preventative’ rather than ‘reactive’. The previous diagnostic stage has identified that *Patient Arrivals* influence the *Total Patients* in the department for one to three hours afterward, and in turn, *Total Patients* influence the *Maximum Wait* time for the next two to four hours. This means that arrivals still have the potential to affect crowding up to seven hours later, and that accurate forecasts of *Total Patients*, and appropriate action on the basis of these forecasts, can reduce wait times for up to four hours ahead (Figure 6.16). For these reasons, Total Patients data is used for the forecasting model.

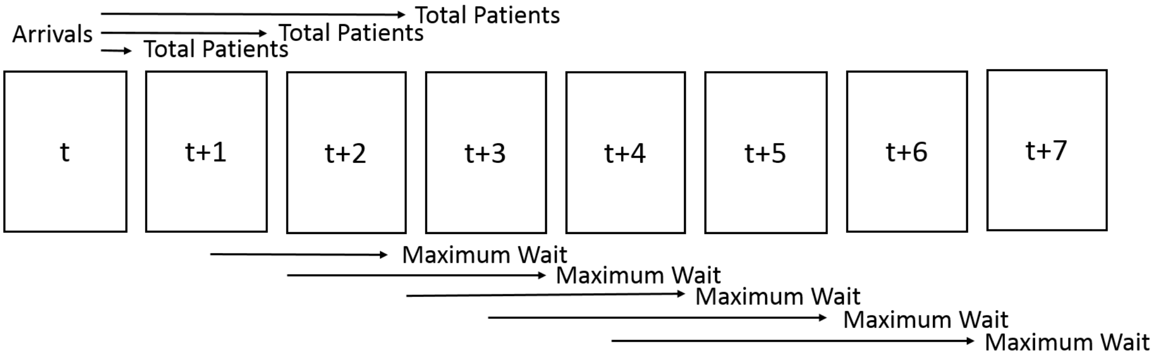


Figure 6-16 Conceptualisation of impact on KPIs of Total Patients forecasts

In this thesis, the terms ‘forecasting’ and ‘prediction’ are used interchangeably. There is an extensive body of work predicting demand for emergency services, and Appendix 3, Section 3.1.3 contains a short review. Time-series methods are part of a suite of predictive analytic methods which have shown considerable success in predicting emergency demand, in particular variations of auto-regressive moving averages (ARMA) as developed by Box and Jenkins (1976).

SARIMA (seasonal autoregressive integrated moving averages), a generalisation of ARMA models, is an approach for modelling univariate time series data that contains a seasonal component. For this case study, SARIMA modelling has been chosen for creating short-term forecasts 2 and 4 hours ahead, as the ED data has a strong daily seasonality. Due to the availability of forecasting libraries, Python 3.7 is used for the forecast modelling. The following sections outline the development of the forecasting model for the *Predict* component of this application of IHAF.

6.5.1 *Characteristics of time-series*

Total Patients data is a univariate time series, that is, a sequence of measurements of the same variable collected over time, at regular 30 minute intervals. Figure 6.17 indicates that there is no consistent trend over the time span plotted (115 days), and no obvious outliers. The variance appears constant over this time span. As observed previously, there is a daily seasonality in the data-series. As the data is plotted every 30 minutes, the seasonality is every 48 data points. These features will be investigated in more detail in subsequent sections. Appendix 3 (Section A3.1.3) provides some additional information, including decomposition of the data into its components (trend, seasonal component, residuals, Figure A3-11).

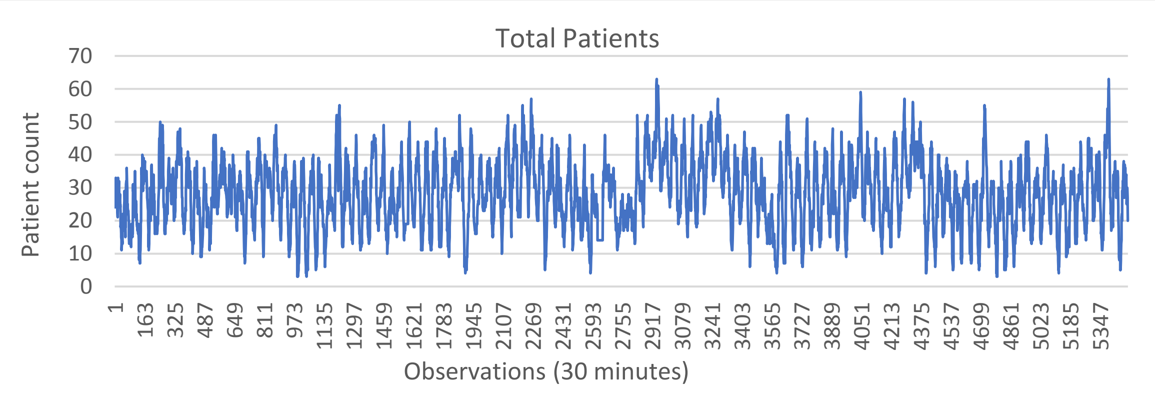


Figure 6-17 Total Patients (every 30 minutes) for 115 days from 3/01/2018

6.5.1.1 *Autoregression*

Autocorrelation is a feature of most time series, as the observations close together tend to be correlated, or serially dependent. Section A3.1.3 in Appendix 3 provides additional details about autocorrelation, and autoregressive models.

In the *Total Patients* data, a quick visual check is performed to look for autocorrelation in the data set by plotting t with t-1 (a lagged value of 1). This has an *r* value of 0.96, indicating a very high linear relationship (Figure 6.18).

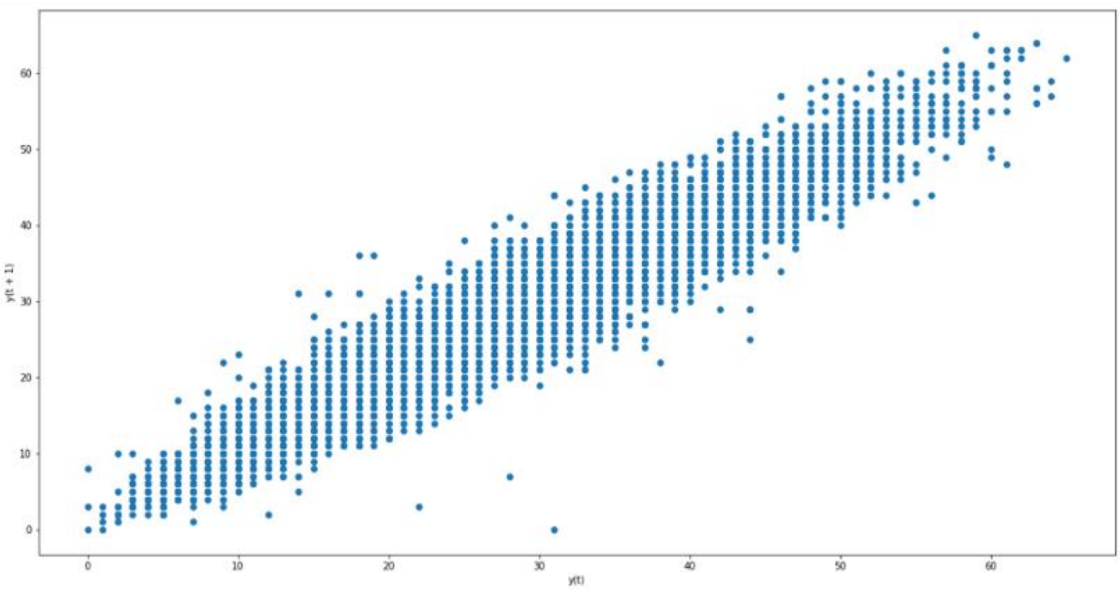


Figure 6-18 Scatterplot of Total Patients and Total Patients -1

One of the simplest ARIMA models is AR(1), or naïve forecast, which uses a linear model to predict the value at the present time using the value at the previous time. This is an autoregressive model of order 1, where the order indicates how many previous lags are used to predict the current time. This can provide a baseline performance as a point of comparison, to give an indication of how well other models will perform on the forecasting problem.

The naïve forecast (AR(1)) reflects the autocorrelation, with a RMSE (root mean squared error) of 3.204 (Figure 6.19).

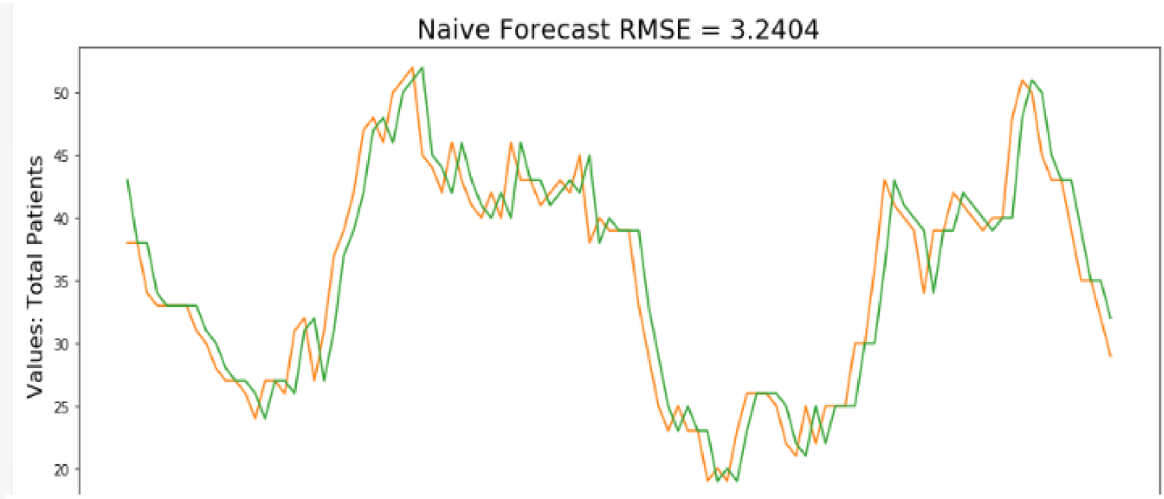


Figure 6-19 AR(1) model on Total Patients dataset for one step ahead using 30 minute observations.

RMSE is one measure of evaluating the accuracy of forecasts. This, and other measures of accuracy, are discussed in more detail in Appendix 3 (Section A3.1.3.2). Due to the high granularity of the data, rolling 30 minute forecasts look fairly accurate (Figure 6.19). However it is clear that forecasts using this method with observations more widely spaced would not perform so well.

A scatterplot could be repeated to look at other lags, for example a visualisation of the time-series shows a strong daily seasonality, hence a lag of 48 (24 hours) could be investigated for the same time one day ago. Alternatively, an *autocorrelation plot (ACF)* will show the correlation coefficients for each lag variable (see Appendix 3, Section A3.1.3.1).

It is apparent that there is significant positive and negative autocorrelation as the *Total Patients* vary throughout the day. The ACF plot (Figure 6.20) provides 96 lags (2 days). The correlations do not appear to diminish over time, suggesting that there is no trend. They remain highly significant with a clear daily seasonality, with peaks and troughs at 12 hours and 24 hours, as expected. A Partical ACF (PACF) (Figure 6.21) also shows seasonality, and the two plots suggest a combined AR and MA process.

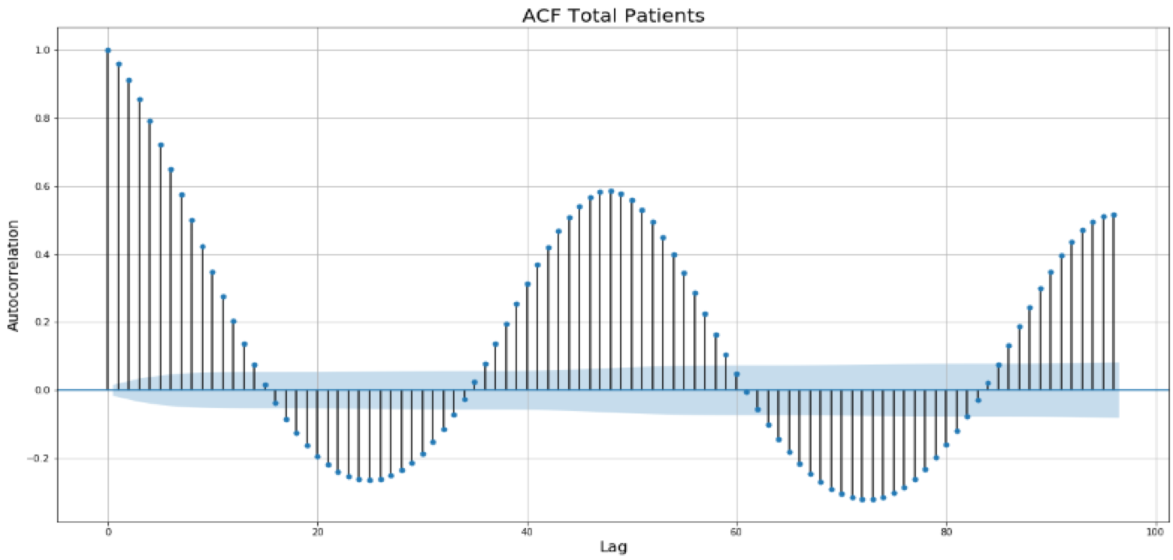


Figure 6-20 ACF for Total Patients to lag = 96

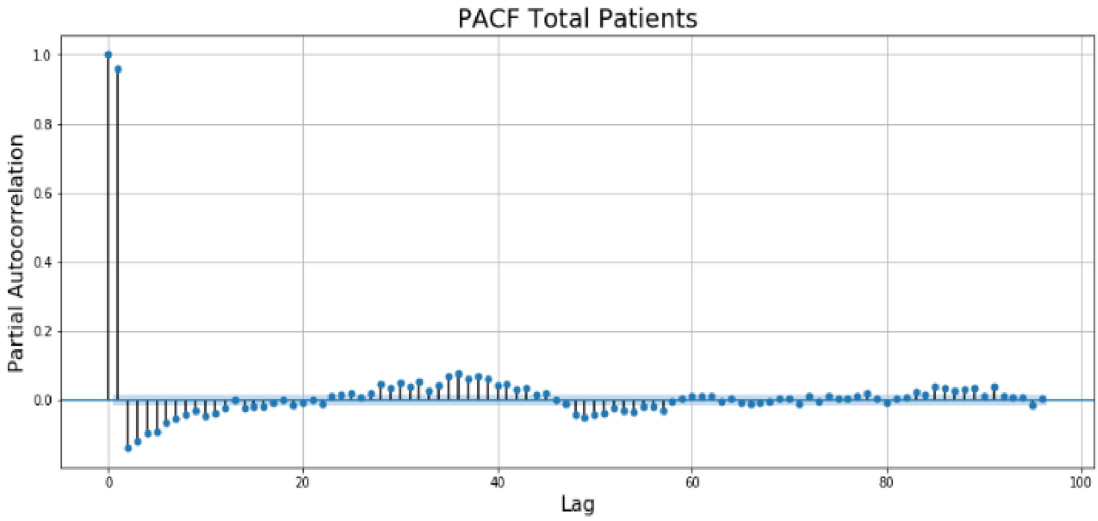


Figure 6-21 PACF for Total Patients to lag = 96

6.5.2 *ARIMA*

Given the above characteristics of *Total Patients* data, ARIMA modelling is investigated for the *predictive* component of IHAF. ARIMA is a generalisation of the simpler ARMA method, which adds integration (I), the use of differencing of raw observations (i.e. subtracting an observation from an observation at a previous time step) to make the time-series stationary, in particular to remove a trend or seasonality. Stationarity is discussed in more detail in Appendix 3, Section A3.1.3.1, and ARIMA models are discussed in more detail in Section A3.1.3.3. For an ACF to make sense, the series must be 'weakly stationary'. This means that the mean is the same for all of *t,* the variance is the same for all of *t*, and the covariance (and correlation) between x_t and x_{t-1} is the same for all *t,* hence the statistical properties of a process generating a time-series do not change over time.

The first step in fitting an ARIMA model is to determine the order of the differencing needed to make the time series stationary, without over-differencing, which can introduce negative autocorrelation and increase the standard deviation (Nau, 2019). For non-seasonal data, first-order differencing may be sufficient. For seasonal data, a seasonal difference is recommended (Hyndman & Athanasopoulos, 2015), while a first order difference may also be required. A seasonal difference is the difference between an observation and the previous observation from the same season. For *Total Patients* data, this requires subtracting each observation from the same time in the previous cycle (48) to create a new time series. This is necessary, otherwise the model assumes that the seasonal pattern will fade over time (Nau, 2019). The data is seasonally differenced: *value(t)* = *obs(t)* - *obs(t* - 48*)* (Figure 6.22).

A statistical test can determine whether the differenced series is stationary. If not, a first order difference may also be necessary. Figure 6.22 illustrates a sample (1000 observations) of the *Total Patients* series. Seasonal differencing reduces the StD from 10.91 to 9.33, and sets the mean to zero. A first order difference applied to the seasonally differenced data further reduces the StD to 3.91.

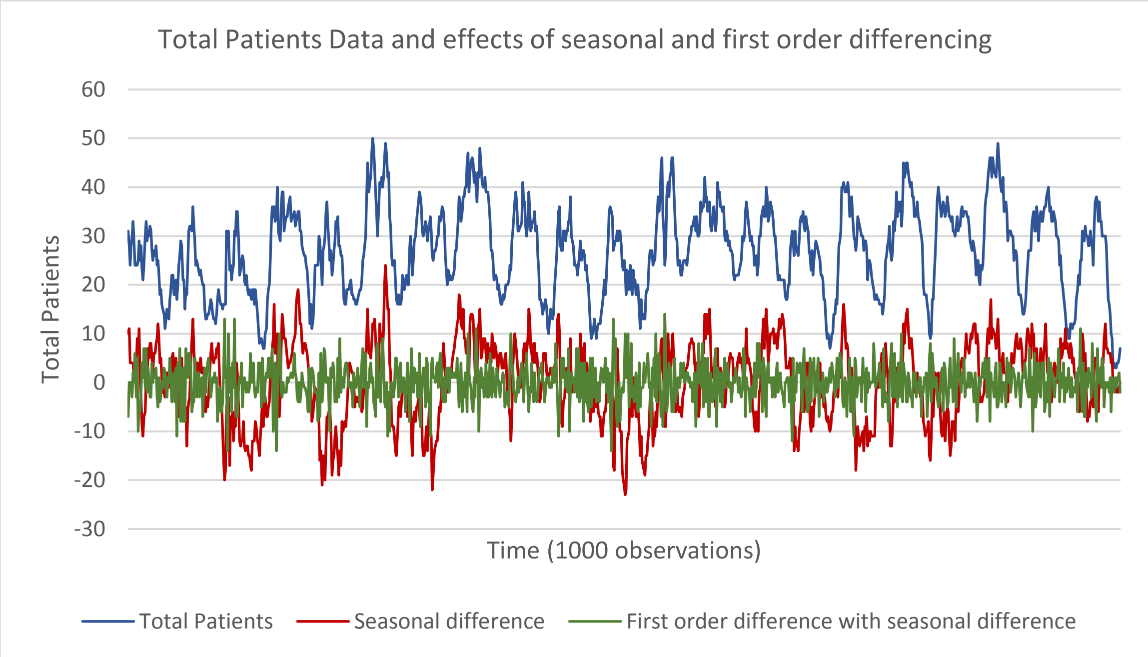


Figure 6-22 Sample of Total Patients, with seasonal, and seasonal with first order differencing

However over-differencing can be problematic. The Augmented Dickey Fuller Test (ADF) is a unit root test for stationarity in a time-series. This is done in statsmodels using adfuller, for analysis of a univariate process in the presence of serial correlation (Statsmodels, 2019b), on the seasonally differenced data, to determine the need for first order differencing. A time-series has stationarity if a shift in time doesn't cause a change in the shape of the distribution; unit roots are one cause for non-stationarity. The test statistic using the seasonally differenced data has a value of -12.02. The more negative this statistic, the more likely the null hypothesis can be rejected (i.e. the dataset is stationary). As part of the output, running an ADF returns a look-up table to help determine the ADF statistic. The ADF results (Appendix 3, Figure A3-12) show that the test statistic value -12.0237 is smaller than the critical value at 1% of -3.432. This suggests that the null hypothesis can be rejected with a significance level of $p<0.01$, and that the time-series is stationary or does not have time-dependent structure. Using the converted dataset means that no further differencing is required, and the p parameter can be set to 0.

Using the newly created stationary data set, the AR (p) and MA (q) parameters now need to be selected. This ACF and PACF plots can give some indications, and in this case there is still daily seasonality present in the data (Appendix 3, Figure A3-13). This indicates that it may be worth considering a better model of seasonality, such as modelling it directly, rather than attempting to remove it from the model using seasonal differencing. Appendix 3 (Section A3.1.3.3) describes further exploration of the ARIMA model parameters, while the next section outlines the SARIMA model chosen for the *Predict* component implemented in IHAF.

6.5.3 *SARIMA*

In Appendix 3 (Section A3.1.3.3) ARIMA(1,0,2) performed well on the data with a RMSE of 3.812 on an unseen test data using one-step forecasts, however it is a method which doesn't support a seasonal component. ARIMA expects data that is non-seasonal, or has had the seasonal component removed; in this case through seasonal differencing. However the ACF and PACF show that there is still some seasonal autocorrelation in the data which could be used to improve the forecasts.

The Seasonal Autoregressive Moving Average (SARIMA) model is an approach for modelling univariate time-series data that contains a seasonal component. It contains additional seasonal terms which are similar to those in the ARIMA (p,d,q) model, but involve backshifts of the seasonal period. It is specified as SARIMA (p,d,q)(P,D,Q)*s*, where *s* is the seasonality (in this case, 48) and (P,D,Q) are the parameters influenced by the seasonal component. P uses the seasonally offset observation in the model, D is the order of seasonal difference, and Q is the order of moving averages or errors in the model, as the seasonal lag differencing introduces a moving average term.

Domain expertise can be used to configure the parameters (Brownlee, 2018), or a grid search can also be used to search a suite of configurations to discover the best model. SARIMA can potentially have a very large number of parameter configurations, so it was considered appropriate to test a wide range of models to choose the best fitting model. A grid search can reveal non-intuitive model configurations with lower model forecasts it can have a very large number of possible configurations.

SARIMA parameters were selected for the combination with the best performance by searching a set of combinations of (p,d,q)(P,D,Q)s. The seasonal component was set to 0 or 48, and the other parameters to 0, 1, or 2. This made over 1400 possible combinations. Joblib was used to speed up the process through parallel processing, however it was still a lengthy process. The model parameters were selected by minimising Akaike's Information Criterion (AIC). AIC is an estimator of out-of-sample prediction error and the relative quality of statistical models for a given set of data. In estimating the amount of information lost by a model once fitted, AIC trades off between the goodness-of-fit of the model and its simplicity. This means that AIC deals with both the risk of over-fitting and the risk of under-fitting (Zajic, 2019).

The selected model is SARIMA(1,1,2)(1,0,1)[48]. The RMSE is 2.771 using one-step ahead forecasting (Figure 6.23). The trend term is selected as 'nc', which is 'no constant' and indicates no trend (Statsmodels, 2019a). Figure 6.23 plots a 21-day rolling (cross-validation) forecast on an 80:20 test set, and for visualisation, Figure 6.24 plots a two-day rolling forecast on two days of the test set with a sample training set of *Total Patients* data (6 days). Better performance was achieved using a first order differencing (d = 1), as suggested by the plotted transformations of *Total Patients* in Figure 6.22, where the StD reduced markedly using a first order differencing on the seasonally differenced time-series.

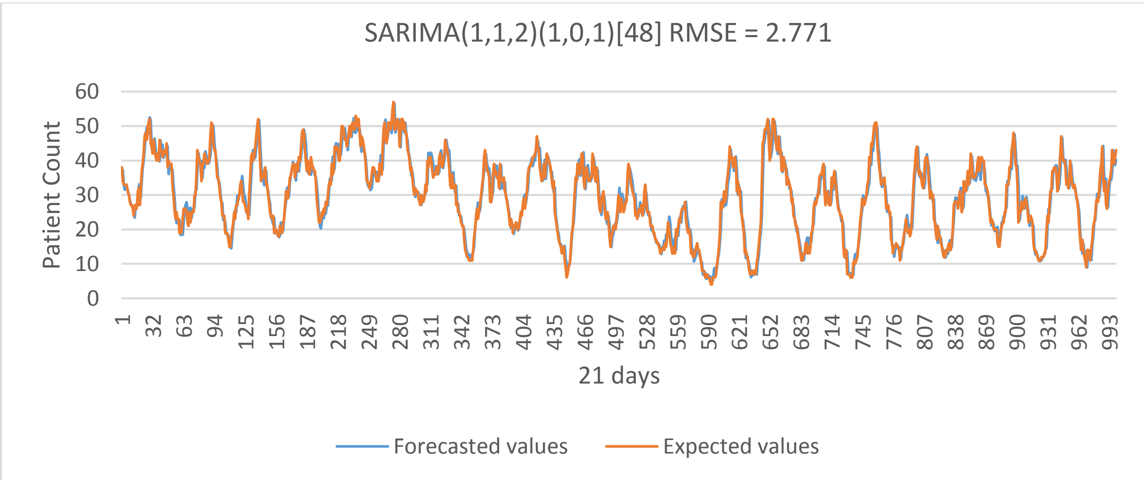


Figure 6-23 SARIMA (1,1,2)(1,0,1)[48] with one-step ahead forecasts.

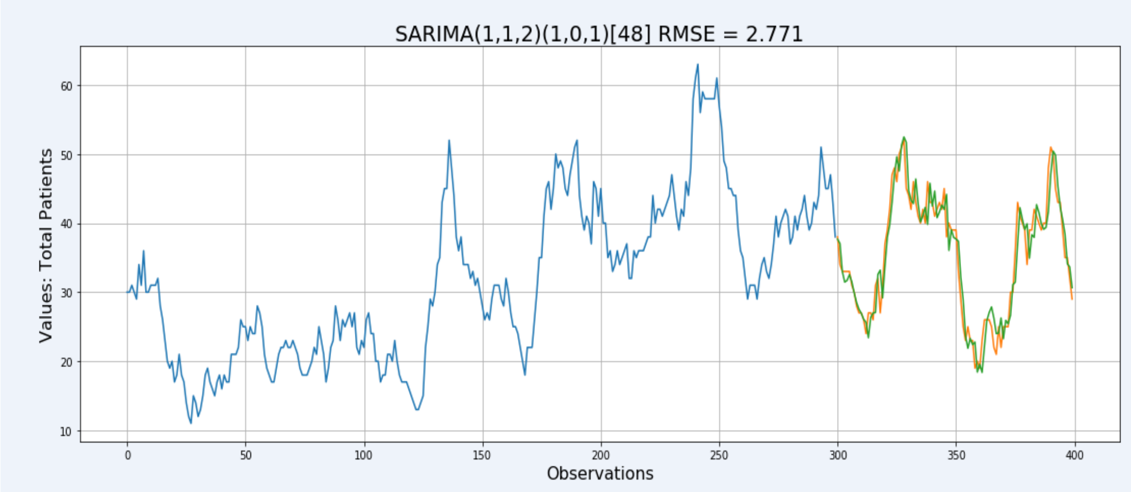


Figure 6-24 SARIMA (1,1,2)(1,0,1)[48] with one-step ahead forecasts. ---- Predicted values ---- Expected values

A prediction interval (PI) provides an interval within which the predicted value is expected to lie with a specified probability. Assuming that the forecast errors are normally distributed, a 95% PI for a one-step ahead forecast is: $y \pm 1.96\sigma$ where y is the forecasted value, and σ is the SE of the forecast distribution. For an 80% PI, the multiplier is 1.28. When forecasting one step ahead, the SE of the forecast distribution is almost the same as the StD of the residuals (Hyndman & Athanasopoulos, 2015). A common feature of PIs is that they increase as the forecast horizon increases. The further ahead the forecast horizon, the more uncertainty is associated with the forecast, and thus the wider the PIs. For a sample of data, the 80% and 95% PIs are plotted below for 48 one-step forecasts (Figure 6.25), alongside the actual observations.

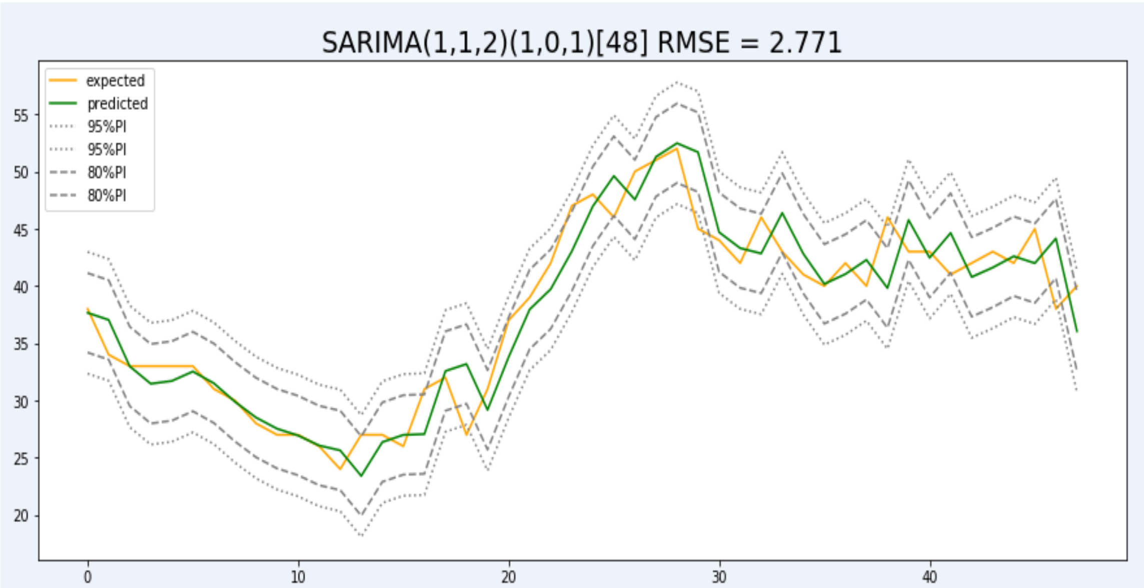


Figure 6-25 One step-ahead (rolling) forecasts for 48 data points with 80% and 95% PIs

This is done by calculating the StD of the residuals, and using the above formula to calculate and plot the PIs. The model is trained on a sample of the historical data. The regression co-efficients learned by the model are extracted and used to make predictions in a rolling manner across the test dataset. As each time step in the test dataset is executed, the prediction is made using the co-efficients and stored. The actual observation for the time step are then made available and stored to be used as a lag variable for future predictions. This is the cross-validation method. A summary of the model fit is presented in Appendix 3 (Figure A3-19).

The residuals are examined in Appendix 3 (Figure A3-20) to confirm the fit of the model. They approximate a normal distribution, which is a useful confirmation of the PIs, and there is no significant autocorrelation, suggesting the chosen model is a good fit for the seasonally differenced *Total Patients* data.

6.5.4 *Resampling for 2 and 4 hour forecasts*

SARIMA is used in the previous subsection to select a model using one-step ahead 30-minute rolling forecasts, however 2-4 hour predictions provide a better possibility of reacting to reduce crowding (Section 6.4). To illustrate using 80% prediction intervals (PI), predictions four hours ahead are reproduced in Table 6.6.

Table 6-6 Multi-step forecasting using 30 minute seasonally differenced Total Patients Data with SARIMA (1,1,2)(1,0,1)[48] up to 4 hours ahead with 80% upper and lower PIs

Forecasts	Prediction	SE	80% UL PI	80% LL PI
30 min	27	2.739089	30.739939	23.727872
1 hour	28	3.766248	32.758610	23.117015
1.5 hours	29	4.464295	34.652110	23.223514
2 hours	28	4.987426	34.321718	21.553906
2.5 hours	36	5.397586	42.846722	29.028902
3 hours	29	5.727762	36.269347	21.606277
3.5 hours	28	5.998221	35.615535	20.260089
4 hours	30	6.222542	37.902665	21.972959

The granularity of the data means that multi-step forecasts 2-4 hours ahead underperform. Using 30 minute data, forecast standard error (SE) increased from 2.74 up to 6.22. The PIs are calculated by using the following formulas:

80% UL PI = prediction + (1.28 * SE)

80% LL PI = prediction – (1.28 * SE)

An alternative approach is to resample the data to reduce the granularity, and provide one-step ahead forecasts. This is done as multistep forecasting places a significant burden on existing data by assuming the accuracy of each forecast (Figure 6.26).

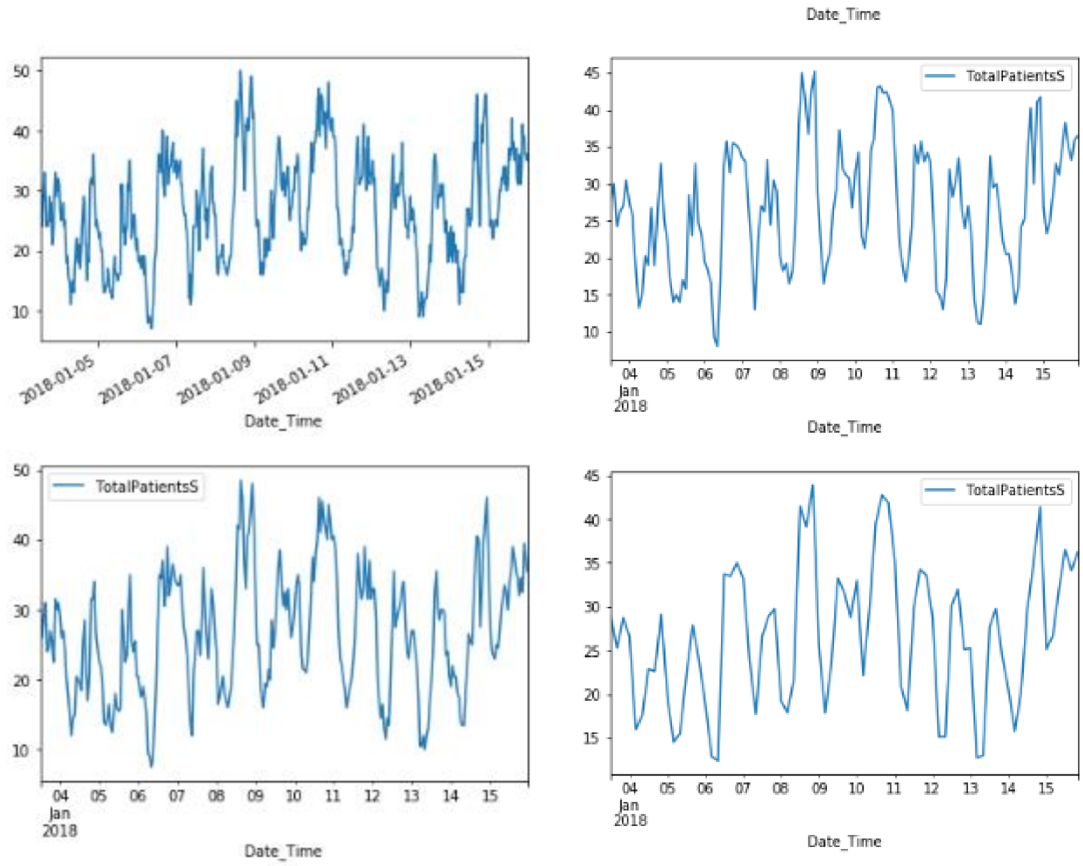
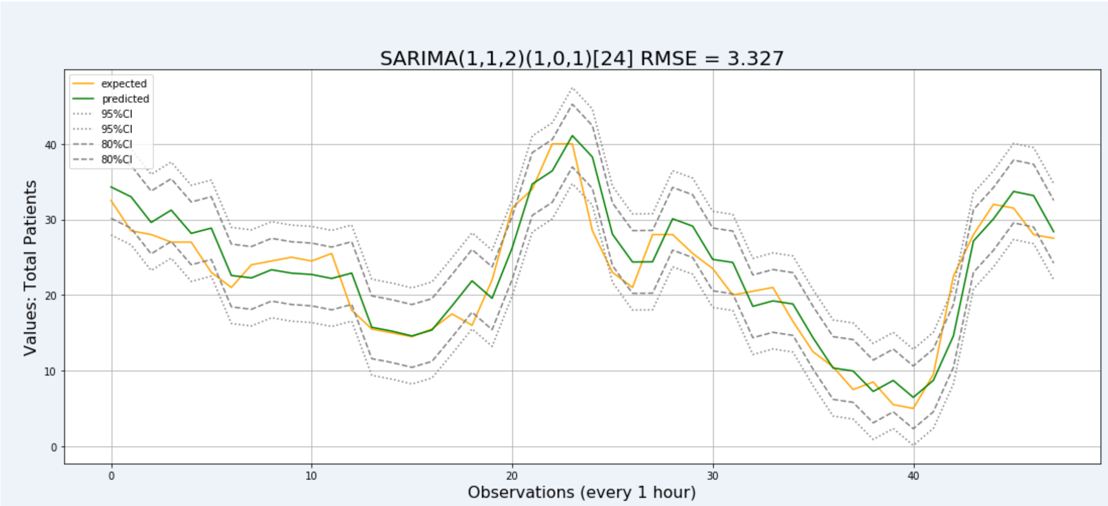


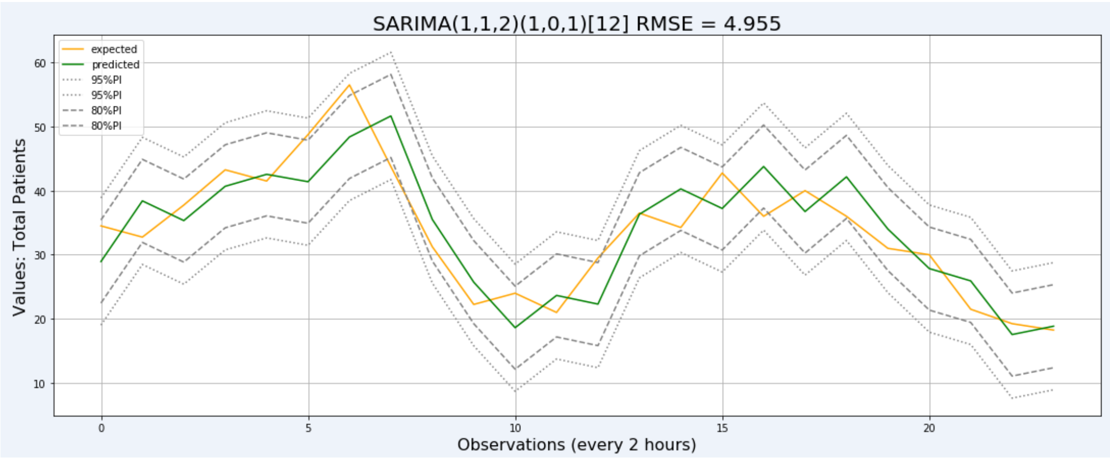
Figure 6-26 Resampling Total Patients: top left: half hourly. Bottom left: hourly. Top right: 2 hourly Bottom right: 4 hourly

Resampling converts a dataset time-interval into a new time-interval. The 30 minute data is down-sampled to hourly, 2-hourly, and 4-hourly by calculating the average *Total Patients* over these time periods. An advantage to this approach is that less data is needed for training the model each time it is called, while the seasonal periods are retained. Naïve forecasts are repeated as a baseline on each resampled dataset. As can be expected, as the granularity of the data reduces, the naïve forecasts lose accuracy. These are plotted in Appendix 3, Figures A3-21 – A3-24. The naïve forecasts on 4-hour resampled data RMSE = 9.480.

The resampled data is seasonally differenced and fitted with SARIMA(1,1,2)(1,0,1)[s] for comparison with the baselines. To illustrate, these are visualised in Appendix 3, Figures A3-25 – A3-27. Each outperforms the baseline, though depending upon the selection and size of the training/test sets, slightly different RMSEs are returned, as will be seen in the next section. Due to the length of time required to train the model and cross-validate it, subsets of data were used in this phase. Note the seasonal period must be adjusted for each dataset. These are plotted in Figure 6-27 (a-c) with 80% and 95% PIs.

The 2-hourly and 4-hourly resampled data is examined in more detail to determine the goodness of fit of SARIMA(1,1,2)(1,0,1)[s] for 1, 2 and 4 hour datasets.





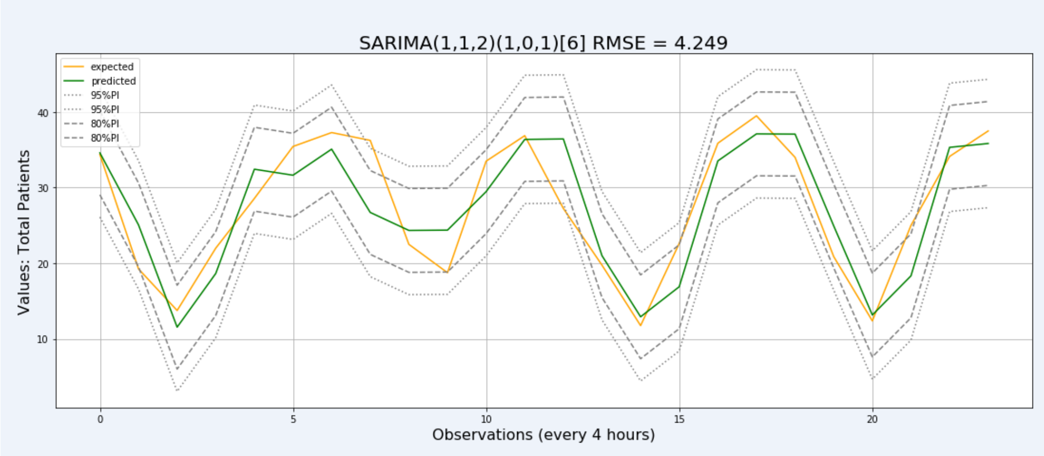


Figure 6-27 (a-c) SARIMA(1,1,2)(1,0,1)[s] on seasonally differenced Total Patients (a) 1, (b) 2, and (c) 4 hour resampled data with 80% and 95% prediction intervals.

Examination of the residuals and ACFs using the 2-hour resampled data shows that the residuals are approximately normally distributed with a small peak of remaining autocorrelation at 1, which may be relevant. The diagnostic results of the residuals on the 2-hourly resampled data are in Appendix 3 (Figure A3-28). This allows forecasts to be updated every 30 minutes, for 2 hours and 4 hours ahead.

PIs can be estimated using standard error (SE), which can be returned as summary statistics with the forecast() function in statsmodels. The model is fitted and evaluated (Appendix 3, Section 3.1.5). Cross-validation of predictions against the test set shows that the model does under-estimate many of the peaks and troughs, leaving some room for improvement. However for a 4-hour (multi-step) forecast, the RMSE = 3.852, which is much improved by training on the full dataset compared with the subsets of data investigated earlier.

Multi-step forecasts on the resampled data up to 4 hours are presented below in table 6.7. As with Table 6.6, the PIs increase with multi-step forecasts. However for one-step ahead forecasts on the 2-hour and 4-hour resampled data, the SEs appear to be acceptable. *As the 4-hour multi-step predictions using 2-hourly data have an SE that is fairly close to the 4-hour resampled one-step prediction, 2- and 4-hour forecasts from the 2-hourly re-sampled data are chosen for making forecasts.*

Table 6-7 Multi-step forecasting using seasonally differenced Total Patients Data with SARIMA (1,1,2)(1,0,1)[s] on resampled data with 80% upper and lower PIs

Resampling	Forecasts	Prediction	SE	80% UL PI	80% LL PI
1 hour	1 hour	31	3.553757	35.62	26.52
	2 hour	31	5.028984	37.35	24.48
	4 hour	30	6.404020	37.60	21.20
2 hours	2 hour	29	4.512923	23.73	35.30
	4 hour	32	6.095951	24.18	39.80
4 hours	4 hour	32	5.391947	38.90	25.10

6.5.5 *Forecasting*

The forecasts are now created for the *Predict* component of IHAF. Having selected the model and the model parameters that best fits the relationships in the historical *Total Patients* data, the salient information captured by the model must be saved so that it does not need to relearn the regression coefficients each time a prediction is needed. The Statsmodels module in Python has built-in functions to save and load models by calling save() and load() on the fitted SARIMAX Results object (Statsmodels, 2019a). The model is trained on all available data and saved. The training data is also saved, for knowledge of the number of observations seen, which is required by the predict() function of the Results object. The model used is SARIMA (1,1,2)(1,1,1)[12] using the seasonal differencing component built into SARIMA and a seasonal period of 12, with the 2-hour dataset.

6.5.5.1 *Two hourly data*

The full dataset has 17196 observations. The resampled 2-hourly dataset contains 4299 observations, but retains the same number of seasonal periods. The coefficients are printed in Figure 6.28.

```
[ 0.72278257 -0.81471212 -0.18408407 -0.03559139 -0.97723887 20.36640598]
```

Figure 6-28 SARIMA (1,1,2)(1,1,1)[12] parameters (p,d,q,P,D,Q)

The model is fitted to be used later for making predictions (Appendix 3, Figure A3-29). The entire training sets of 30-minutes and 2-hourly resampled data are saved as numpy arrays. The load() function will be used to load these later

(Figure A3.30). The coefficients are cross checked to ensure they have saved and load correctly.

The predictions can now be made using ARResults.predict. This requires a start and end for making in-sample or out-of-sample predictions (Statsmodels, 2019). Unfortunately ARResults doesn't support PIs, so these are estimated using the SE returned from the get_forecast function earlier (Appendix 3, Figure A3-29). The SE were stable across the test sample, which would be expected on a stationary dataset. 80% intervals are chosen as realistic.

Now the forecast model needs to be kept updated, once the next real observation is made available by *NHSquicker*. This requires updating the data set used as inputs to make the subsequent prediction. The following steps are required (Appendix 3, Section A3.1.5, provides additional detail):

- The new observation is recorded. In Appendix 3 (Figure A3-31), the unrealistic figure 120 is manually inputted.
- The 30-minute dataset and 2-hour dataset are loaded. The 30-minute dataset is indexed with the original date-time index so that an additional row can be added.
- The new observation is inputted. For the integrated model, the new observation can be added in real-time and the index retained.
- To control the size of the 30-minute dataset, it is saved with the first value removed as a new value has now been added to the end.
- The new 30-minute dataset is saved.
- The new 30-minute dataset is resampled 2-hourly and saved as the new 2-hour dataset.

The code can now be re-run every time a new observation is sent from *NHSquicker*, every 30-minutes to update the 2-hour and 4-hour forecasts. This is conceptualised in Figure 6.29.

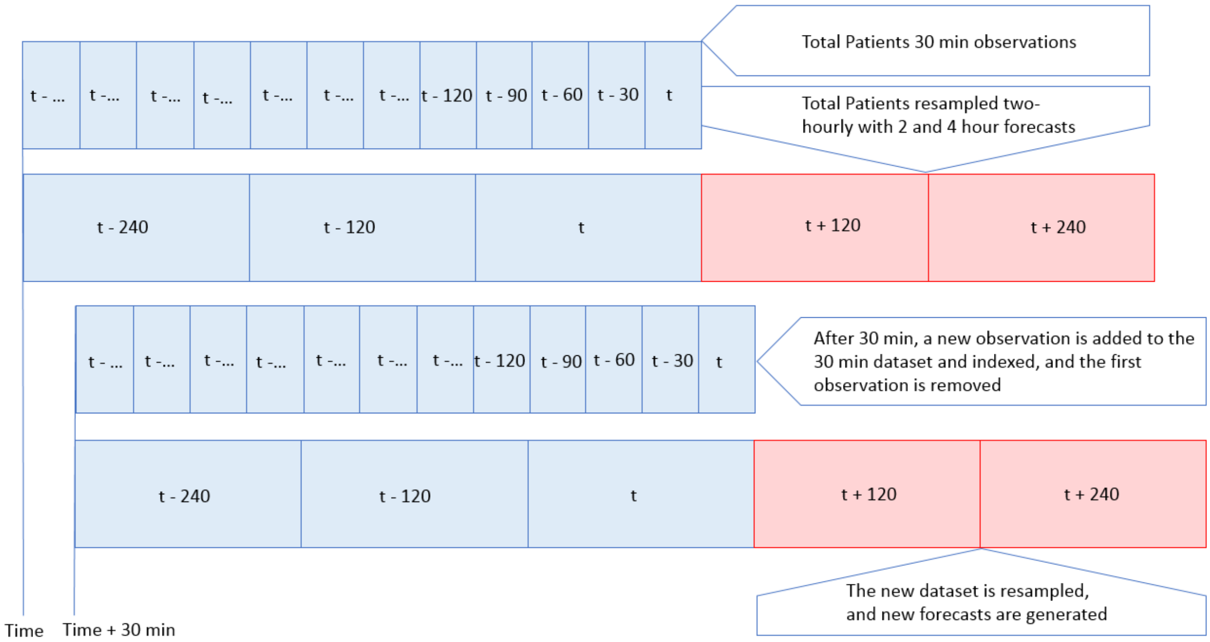


Figure 6-29 Conceptualisation of data processing and forecasts on Total Patients data

The integration of the real-time forecasts in IHAF is discussed in the next section.

Implications for IHAF implementation

Using *NHSquicker* data, *Total Patients* data is used as a proxy for crowding. Forecasts using SARIMA time-series forecasting are generated 2 and 4 hours ahead, on 2-hour resampled data. These are updated every 30 minutes as a new observation arrives. This forms the *Predict* component of IHAF. The predictions are used as a predictive trigger to initiate the simulation, which will be used to support system recovery and task- and system-level awareness (Section 6.7).

6.6 Integration Component: forecasts

It is considered to be an important element of the HM using IHAF, that the data, the predictive model, and the simulation model are integrated for usability in terms of its effectiveness and efficiency (Karsh, 2004; Middleton et al., 2013). As the HM is designed to be a recurrent-use decision-support tool during busy periods, minimal manual interaction is important. This is required, regardless of the methods used in the implementation, hence synergies between the methods require early consideration. As discussed in Chapter 4 (Sections 4.2 and 4.6), the number of methods, the frequency of interaction, the number of points of interaction, and type/frequency of triggers will be determined by the specific application.

In this case, the forecasting model must receive the updated data, process it, make predictions, and if it reaches the threshold value, trigger the simulation. This procedure is illustrated in Figure 6.30. Following the arrows: (1) Using Java programming, the real-time data from *NHSquicker* is downloaded, parsed, and saved at set intervals (Mustafee et al., 2016); (2) The Java component sends the relevant values to the forecasting model (Section 6.5); (3) The forecasts are sent back to determine if the thresholds reach the hourly trigger (Section 6.4.3); (4) If the hourly trigger is reached, the simulation model is triggered, and *NHSquicker* values are injected into the simulation at initialisation; (5) Short-term scenarios provide decision-support; (6) Where action is taken based on the decision, this will be reflected in the real-time data. Information updates system-level and task-level awareness to augment decision-making, however the autonomy of the decision-maker is retained. For this reason, control of the system is indirect, represented as a dashed line in Figure 6.30.

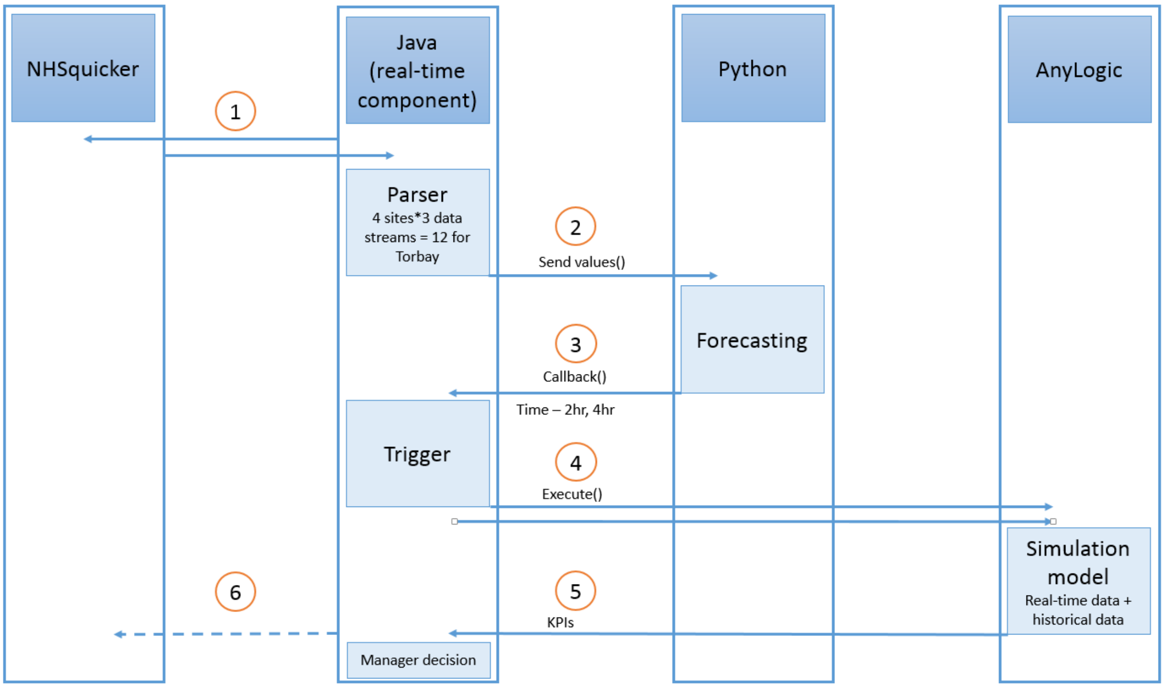


Figure 6-30 Integration of the hybrid model

As per Section 6.5.5, the date-time index is updated in the 30-minute dataset in the Python forecast model code, while the new observation is inserted directly into the forecast code as a variable (Figures A3-32 and A3-33). This is done using Java, which can call the Python forecast model and insert arguments directly. This creates a new row in the 30-minute dataset. This new dataset is then used to update the 2-hour dataset for making forecasts (see Figure 6-29). More detail is provided in Appendix 3, Section A3.1.5.

The elements for this are in place to integrate the model for future work. A similar procedure is required for returning the forecasts to Java, to trigger the simulation, which is discussed in the next section.

6.7 Prescriptive Analytics: Discrete-Event Simulation

The final stage of the HM development in the IHAF framework is the *Prescriptive* stage, the real-time simulation. In this implementation of IHAF, DES is used due to the stochasticity and queue-based structure of emergency care, however other methods can be used, as discussed in Chapter 4. The purpose of the DES in this case is to provide solutions toward preventing crowding, given a forecast of *Total Patients* exceeding the thresholds defined in Section 6.4.3. This demonstrates how short-term demand forecasts can be used for planning for recovery. The use of forecasts enables more recovery time, by preventing the critical situation from occurring in the first place. In the HM, the predictive and prescriptive components work in synergy, combining the benefits of each method (conceptualised in Figure 6.31).

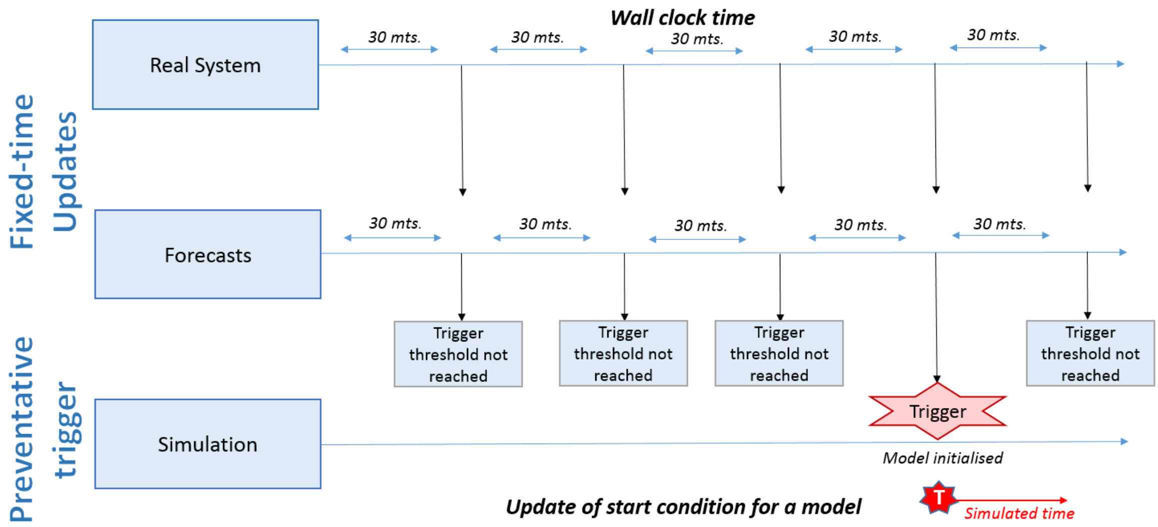


Figure 6-31 Synergy of the real-time data, forecasts and simulation (descriptive, predictive, prescriptive)

As discussed in Chapter 2, DES is the most commonly used simulation method in ED, while crowding is the most common problem investigated (Gul & Guneri, 2015; Paul et al., 2010). The pace and unpredictability of ED adds a specific challenge for M&S studies (Jurishca, 2005), hence real-time simulation has been proposed for short-term decision-support in ED where circumstances change hour-to-hour (Tavakoli et al., 2008; Bahrani et al., 2013). Improved access and quality of healthcare operational data makes automated data capturing systems for real-time simulation increasingly feasible in hospitals. The next sections outline the development of the DES for IHAF.

6.7.1 *Stages of a simulation study*

A clear, stepwise method is recommended for approaching a modelling and simulation (M&S) study, ensuring that all main activities are addressed throughout the study lifecycle using a methodical approach. This can assist the researcher in formulating the problem, understanding the system, investigating appropriate methods for addressing the problem, building the simulation model, designing experiments, and presenting the results (Robinson, 2004). It also assists with verification and validation activities throughout the lifecycle (Balci, 1989). A number of DES M&S frameworks have been proposed (e.g. Shannon, 1998; Robinson, 2004; Law, 2009; Balci, 2012) with a view to supporting the conduct of simulation studies. Most frameworks start with problem formulation, as the problem communicated from a decision-maker or stakeholder to a researcher or analyst is rarely clear and well-defined (Balci & Nance, 1985).

The second stage is that of investigating solution techniques. Once the problem is understood, the technique/s with the highest cost-benefit ratio should be selected (Balci, 1989). The boundaries of the model must be determined, a conceptual model of the problem formulated, and input data identified and prepared. The IHAF framework provides a high-level starting point for these issues. The final stages include programming the model, verifying and validating the model, and experimentation and analysis, followed by communication of the results, or in this case, evaluation of the model in its real-world setting (Chapter 7). The model is developed using AnyLogic 8.5.2 Personal Learning Edition (PLE). AnyLogic is developed in Java, and the download and parser scripts for the *NHSquicker* model are written in Java (Mustafee et al. 2017b), aiming to support integration of the HM components.

The stepwise method used in this thesis is that outlined by Martin, Depaire and Caris (2018) for conducting M&S studies based on a critical synthesis of existing frameworks from the literature (Figure 6.32). The starting point is problem formulation, aligning with the problem definition stage of IHAF. Additionally, continuous assessment is a central feedback mechanism in the method, including all evaluative actions that can cause the simulation project to iterate back. The steps are discussed below. However as this method was developed to encompass the full lifecycle of an M&S study, some of the stages have been previously addressed, in particular, aspects of steps 1-3.

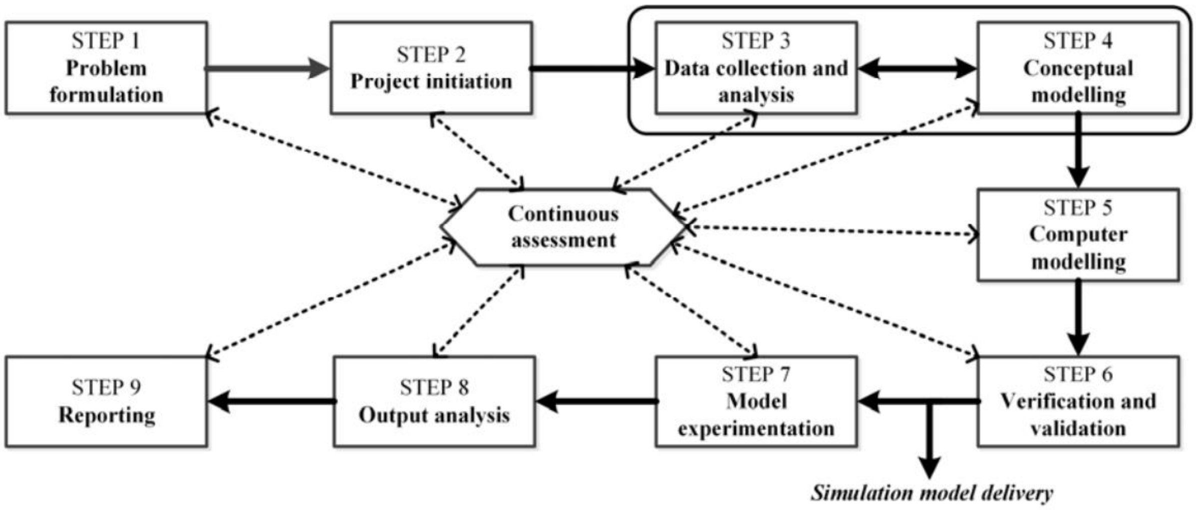


Figure 6-32 Method for conducting an M&S study, from Martin, Depaire and Caris (2018)

STEPS 1 & 2: Problem formulation and Project initialisation

The modeller should understand the problem and its context and have basic insights into the process containing the problem situation. As the simulation model is a component of a HM, the problem formulation stage has been previously addressed in Chapter 5.

The problem formulation step clarifies and specifies the problem. The goals of the study are specified, alongside any sub-objectives, and the questions to be answered. In the case of an autonomous simulation, which questions the model will answer are clarified. The method is chosen, and model boundaries are selected. To support flexibility, Kelton et al., (2015) suggests that model boundaries should not be rigid. Preliminary input parameters and the level of model detail can be considered, aiming for the simplest model possible to address the problem.

In this application, relevant aspects of model scope include:

- Entities (patients) enter the system according to hourly arrival distributions, and are assigned a triage category.

- The model initialises at the current date-time, less a warm-up period, as its runtime is 2-4 hours, and arrival distributions vary according to time of day and day of week.
- A proportion of patients leave the system without being treated. This is considered a safety risk.
- Staff treat patients based on priority.
- Based on ED data, patients may have one, two or three treatments, and may have internal (within the department) or external (outside the department) investigations performed.
- Once treatment is completed, patients leave the department (are discharged) or are admitted. There may be delays for admission (e.g. no bed available) or discharge (e.g. transport delays).
- A proportion of patients die in the department.

Outputs are:

- Length-of-stay (LoS) in ED
- Total patients in the department
- Number of patients in the waiting room by triage category
- Time spent in waiting room
- Patients who leave without being seen (LWBS) by triage category

While *NHSquicker* provides limited information for updating a simulation model, it contains real-time information about the entire urgent care network (UCN), which can be helpful for decision support. As discussed in Chapter 2, demand management (managing demand at source) is one approach to managing crowding. Currently, in ED this involves closing ED to minors (Triage categories 4 and 5) when crowding occurs and reaches 'OPEL 4', i.e. the highest level of operational pressure. This is a measure of operational pressure across the whole hospital, and considers ED wait times and numbers, as well as available bed capacity and expected discharges. Low-acuity patients are asked to attend the nearest MIU at this point.

The geography of the use-case area means that all three MIUs are roughly equidistant from the ED (Figure 6.33).

Figure 6-33 Geography of the use-case area: MIU ● ED ●

The road to MIU(1) is the most direct route, therefore patients are generally advised to attend here. However, if MIU(1) is also at capacity, this risks creating a demand-capacity mismatch. For this reason, it is advantageous to be able to consider the capacity across the UCN when managing demand. Given the real-time information available from *NHSquicker*, and the focus of this study on patients with low-acuity conditions and crowding, the following scenarios are proposed for initial investigation:

- Baseline – proportion of patients who LWBS (leave without being seen) per triage category, calibrated to 2018 data.
- Scenario 1 – Redirect all Category 4 and 5 patients when the number of patients in the department reaches hourly trigger (reactive trigger).
- Scenario 2 – Redirect a proportion of Category 3, 4 and 5 patients to MIU when the number of patients in department is forecasted to reach the hourly trigger in 2-4 hours' time (preventative trigger).
- Scenario 3 – Redirect a proportion of Category 3, 4 and 5 patients to MIU when the number of patients in the department is forecasted to reach the hourly

trigger in 2-4 hours' time (preventative trigger), and given sufficient MIU capacity.

STEP 3 Data collection and analysis; and STEP 4 conceptual modelling

Steps 3 and 4 can occur interactively, to avoid the conceptual model becoming too complex for the data available to support it (Onggo & Hill, 2014). In DES, this often starts with a basic process flowchart, which is helpful to guide data collection. Figure 6.34 shows a flowchart of the ED basic processes mapped with the real-time data. Patients may enter either via ambulance or they may walk in. Patients brought by ambulance will go straight to treatment, and will be allocated a triage (severity) category using the following severity category descriptors, taken from the ED dataset:

- 1: Immediate
- 2: Very urgent - within 10 minutes
- 3: Urgent - within 1 hour
- 4: Standard - within 2 hours
- 5: Non-urgent – within 4 hours

Patients who walk-in will register with a receptionist. At this point patients will enter the waiting room, to wait for triage (ideally within 15 minutes), where a triage category will be allocated. Alternatively, a patient may be sent directly to the treatment area. From the waiting area, a patient may wait for triage, may go straight through to the treatment area, or may choose to leave the department without waiting for treatment. Patients who leave without being seen (LWBS) represent a quality and safety concern, and thus LWBS rates are used as an ED performance metric (RCEM, 2019). LWBS is associated with perceptions of excessive waiting times and poorer patient experience.

Following triage, a patient may be discharged, sent back to the waiting area, or sent for immediate treatment. At this point, patients could also be potentially sent to a MIU if a significant delay until treatment is predicted, and the patient's condition is non-urgent. Treatments will occur in different areas of the department (minors – generally categories 4 and 5; majors – generally categories 2 and 3; and resus – category 1). Although the ED datasets provided give some indication of where in the department treatment took place, patients may move, for example back to the waiting area, into a clinical decision-support area, or into another treatment area. From the data, triage category is a more stable characteristic for determining individual behaviour in the model, as it is fixed. Patients may have zero, one, or more treatments, and zero, one or more investigations, which may take place at the bedside (internal investigations, e.g. blood test) or outside of the ED (external investigations, e.g. Xray). Patients may be admitted or discharged; a small number of patients will die in the department. Prior to admission or discharge, a patient may be 'admitted' to a Clinical Decision Unit (CDU) to await a decision. At this point, the clock stops in terms of four-hour monitoring. Figure 6.34 maps the *NHSquicker* real-time data to a conceptual flowchart. As the data has been validated for a different purpose, patient decision-making, it is not the 'ideal' data for supporting real-time simulation, however it can be used to initialise parts of the simulation model.

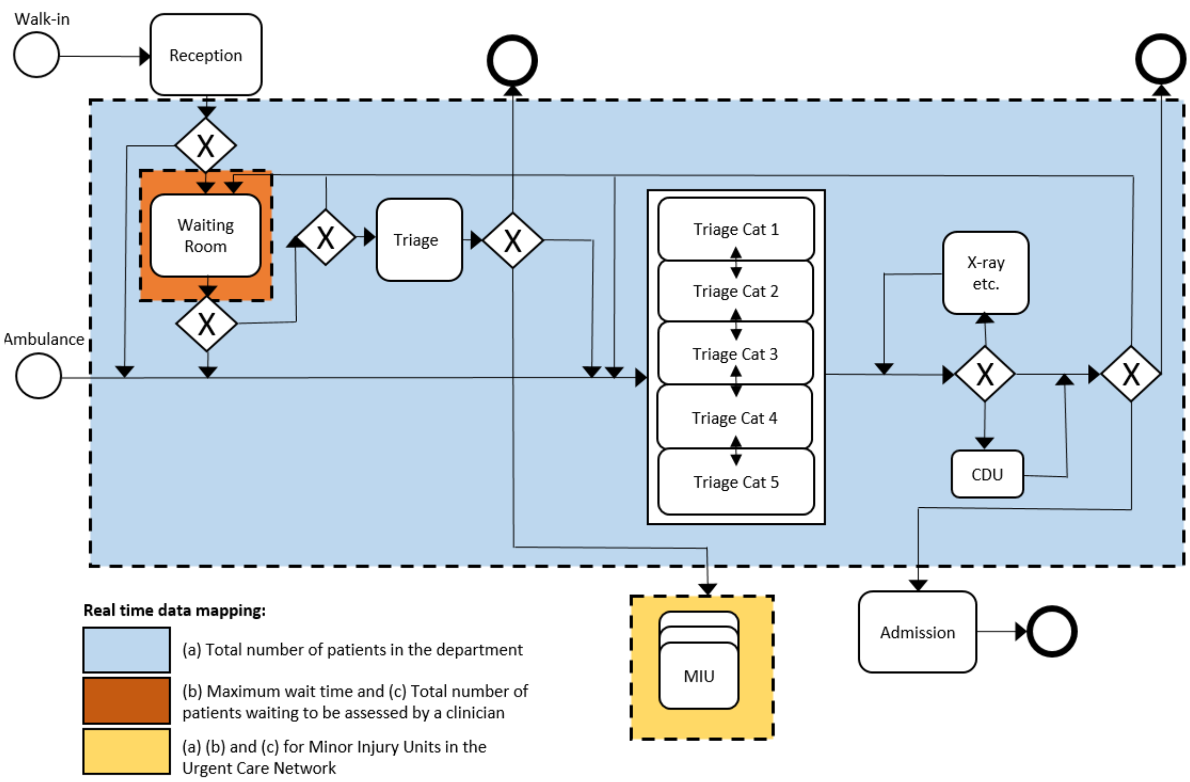


Figure 6-34 The use-case Emergency Department mapped with NHSquicker real-time data

To convert the flowchart into a computer model, both structural data and data to model input parameters are needed. Validation data is also required (Robinson, 2004). Data collection sources can include interviews and observations of the process, alongside data from information systems (Martin et al., 2018). The historical and real-time datasets used for building and validating the model are described in Sections 6.3 and 6.4, and Appendix 3. Additionally, site visits and observations were needed to model the processes.

Arrivals at the ED vary by hour-of-day and day-of-week, so an arrival schedule was constructed for each (Appendix 3, Section A3.2.1, Table A3.2). This means that when the simulation is initialised at the current date-time, patient arrivals will continue to be generated using an appropriate distribution as the model runs, which is in minutes. The arrival rate was calculated using 2018 data by dividing the average hourly arrival by 60. Inter-arrival times were calculated, however AnyLogic supports the use of arrival rates, and applies a Poisson distribution for each calculated arrival rate, which is seen in the ED dataset (Figure 6.35, example of distribution of arrivals from 1200:1300, Monday). This enables entities to enter the simulation model using a distribution for each hour-of-day and day-of-week.

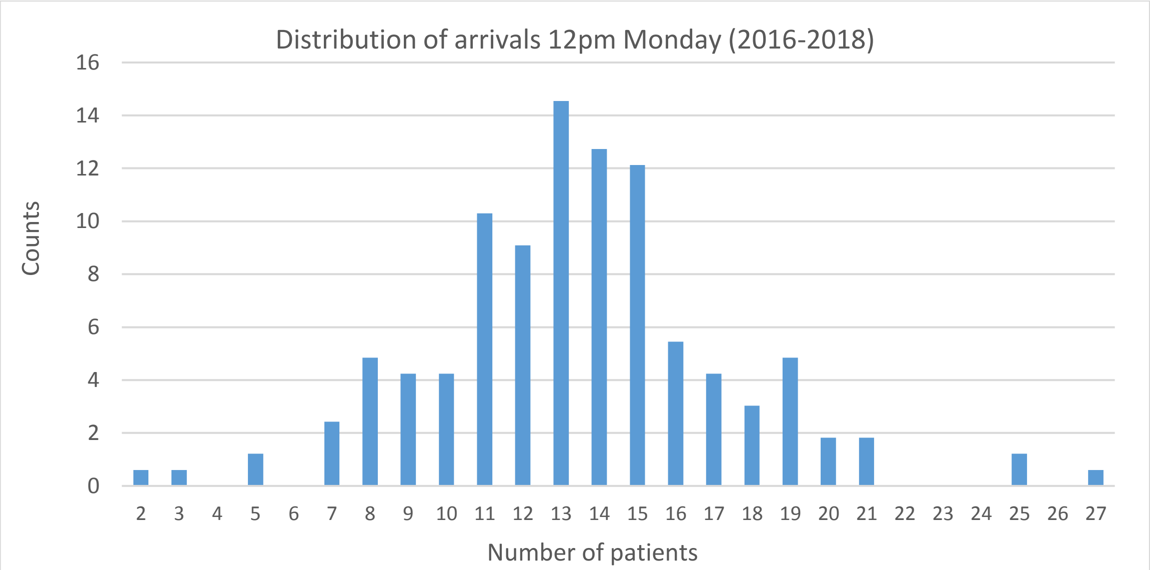


Figure 6-35 Example of distribution of arrivals in one hour (12pm) on one weekday (Monday) 2016-2018

Triage categories were defined at the start of Step 3. A triage category distribution was calculated. This was found to be relatively stable per year of available data from 2016-2018 (Appendix 3, Figure A3-34), and per hour-of-day (Figure 6.36). A small number of un-coded (null) observations were found to be evenly spread across triage categories. Figure 6.36 shows that the daily arrival patterns per triage category follows the overall arrival pattern. This enables entities to be allocated a triage category upon arrival into the system according to a probability distribution, regardless of time-of-day or day-of-week. The probability distribution is in table 6.8.

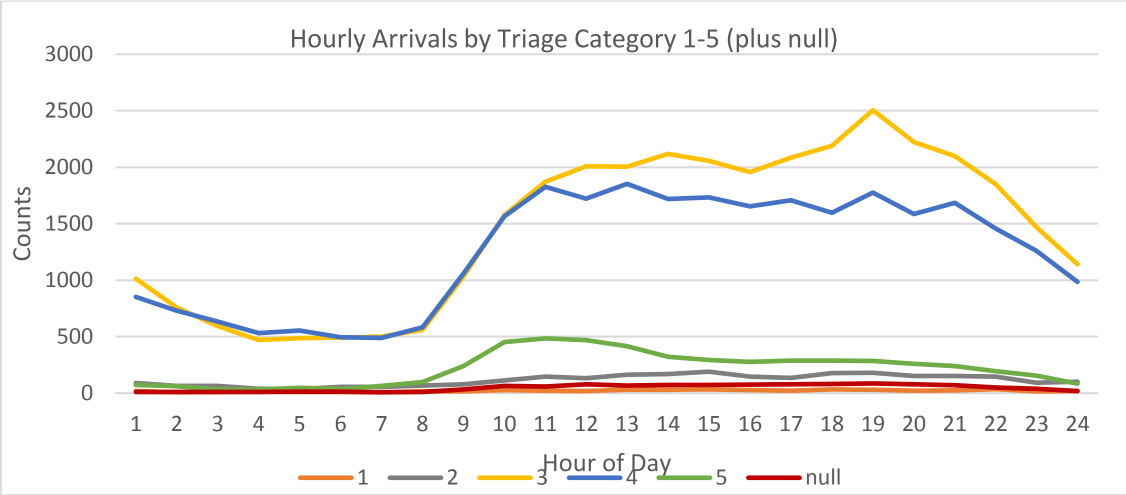


Figure 6-36 Hourly arrivals per triage category, ED Attendance data, 2016 - 2018

Table 6-8 Triage category probability distribution

Category 1	Category 2	Category 3	Category 4	Category 5
0.007	0.037	0.484	0.406	0.066

Although the ED dataset contains duration of stay data per triage category, these distributions are wide and flat, and vary little between triage categories, as they do not differentiate between time spent in treatment, and time spent waiting for treatment. For behavioural reasons (i.e. working to targets), ED lengths-of-stay all peak sharply at the four-hour mark, distorting the distribution (e.g. Gruber et al., 2018). For illustration, the use-case ED Category 5 LoS are plotted in Appendix 3, Figure A3-35.

A better approach is to determine the proportion of patients who had no treatment, one treatment, two treatments, and three treatments, for each triage category, as captured in the ED dataset. This includes all treatment options, including resuscitation, drug administration by all methods, splints, plaster, dressings, and minor operations. These are calculated in Appendix 3, Section A3.2.1, Tables A3.3; A3.4. Table A3.5 is used as conditional probabilities in the simulation model. A staff nurse provided estimates of treatment durations for first, and subsequent treatments per triage categories, in triangular distributions. These are in Table A3-6 in Appendix 3. It is a well-recognised problem that ED data is frequently incomplete, such that service times distributions are unobtainable (e.g. Kuo et al. 2016), and the wide variability in ED activities can make this a significant challenge for simulation modelling. In this case, while the ED dataset includes timestamps, it contains considerable missing data and errors. However, the use of estimates presents a weakness.

From the ED data it is possible to determine the proportion of patients who required internal and external investigations. Again, discussions with a staff nurse divided the investigations into those that could be done at the bedside (e.g. blood tests, urine tests, physical examination) and those that required leaving the department, potentially freeing up the trolley and staff (e.g. Xray, ultrasound scan, MRI, CT scan, bone scan). For simplicity, these were combined into zero, and one or more investigations, displayed in Table A3.7 (Appendix 3), and are used as conditional probabilities. Table A3.8 shows estimated distribution of service times for investigations, and the resource requirements by a staff nurse.

Table A3-9 (Appendix 3) displays patient discharges as a proportion of all discharges, by triage category. These are divided into those who are admitted, those who died in the department, those who were discharged to any destination, and those who left without treatment or refused treatment (Leave without being seen [LWBS]). Patients who die in the department, and who LWBS are included in the simulation model. Additionally, those who 'could have gone to MIU' are estimated by adding those who were coded as any of the following: 'Discharge – follow up treatment by GP', 'Discharge – no follow-up', 'Left department before being treated/Did not wait', and left department having refused treatment/self-discharged'. This was for later use in developing simulation scenarios.

Patients who walk-in are triaged, usually by a triage nurse (nurse practitioner) but occasionally by a consultant, when 'minors' are busy and 'majors' are quiet. Estimated triage durations are in Appendix 3, Table A3-10. These are used in the simulation for triaging walk-in patients. The proportion of patients who arrive by ambulance/helicopter and those who walk-in to the department is displayed below in Table 6.9. Those who arrive by ambulance or air-ambulance will by-pass triage, whereas those who walk-in will enter the waiting area and wait for triage.

Table 6-9 Probability distribution for walk-in and ambulance arrivals

Arrival mode	Probability
Ambulance/Air ambulance	0.36
Walk-in	0.64

The average number of patients per hour who leave without being seen (LWBS) was calculated from the ED data. It was found that these correlate highly with the average *Maximum Wait* time calculated from the *NHSquicker* data at r = 0.947. This is plotted in Figure 6.37 as averages per hour-of-day (note dual axis), and as a scatterplot in Figure 6.38.

This is useful, as *Maximum Wait* peaks 2-4 hours after *Total Patients*, as found in Section 6.4, where a correlation was found between *Total Patients* and *Maximum Wait* times in 4-hours. This would suggest that there is also a relationship between *Total Patients* and LWBS.

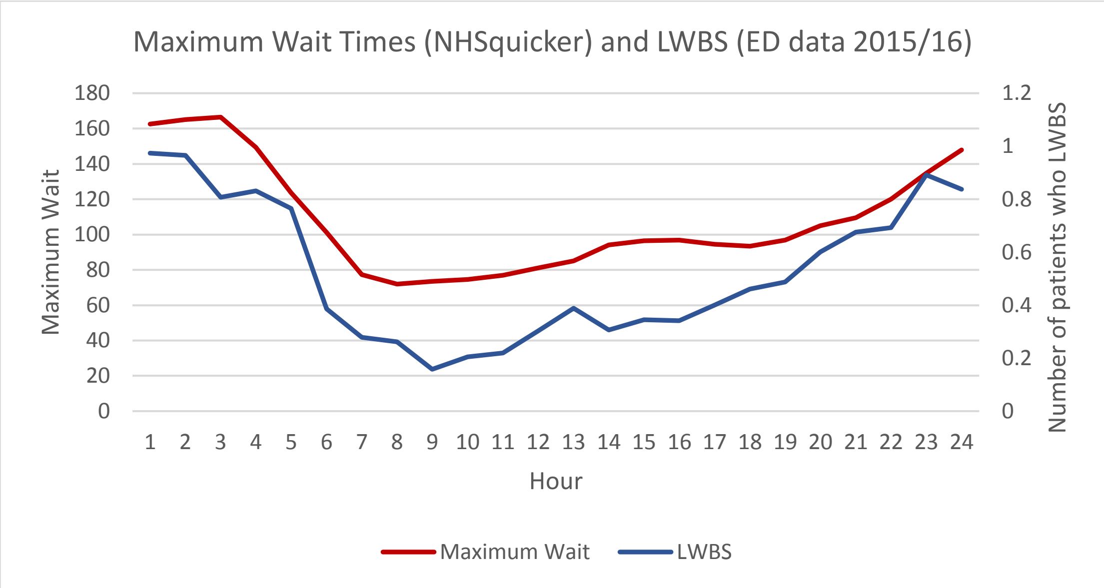


Figure 6-37 Average maximum wait and LWBS per hour of day

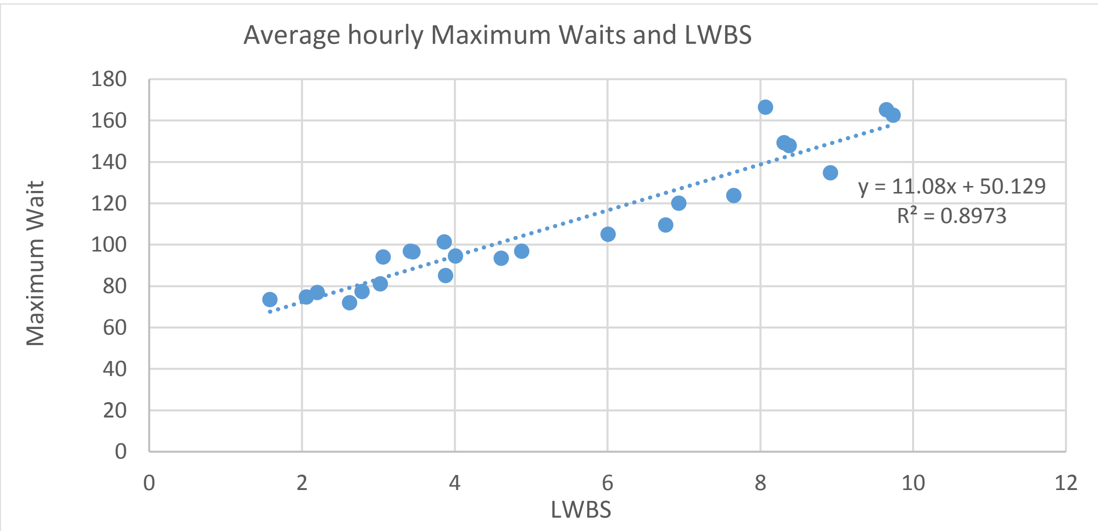


Figure 6-38 Scatterplot of Maximum Waits (NHSquicker) and LWBS (ED historical data)

Plotting average hourly LWBS with a 4-hour lead with average hourly *Total Patients* is therefore similarly correlated, r = 0.833, as expected (Figure 6.39). This suggests that a reduction in *Total Patients* should proportionally reduce the number of patients who leave before treatment. In the simulation model at this time, 'wait-time tolerance', a user-input, is calibrated against the current percentage of patients who LWBS per triage category. However the real-time *Maximum Wait* time can be used to predict the number of patients who LWBS. This is future work, currently not implemented in the model, presenting a further, potentially valuable, use for the real-time data.

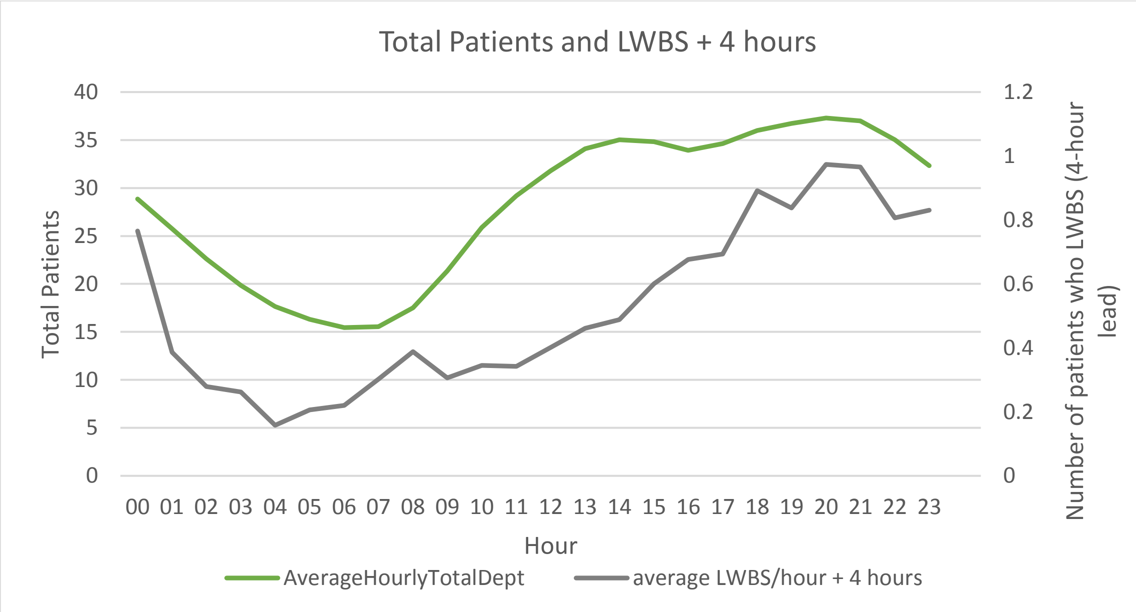


Figure 6-39 Total Patients mapped to LWBS in 4 hours' time

LWBS is examined per triage category, however *NHSquicker* data is not currently available by triage category. Figure 6.40 compares triage category distribution with LWBS distribution by triage category.

Proportionately more patients in Categories 4 and 5 LWBS. This is implemented in the model as a behavioural component. A wait-time tolerance is set per triage category to calibrate LWBS according to Figure A3.9 in Appendix 3, rather than a proportion per triage category. This can remain static, or can be manually inputted, but is a crude measure; using real-time data to calculate LWBS is hoped to be a better approach.

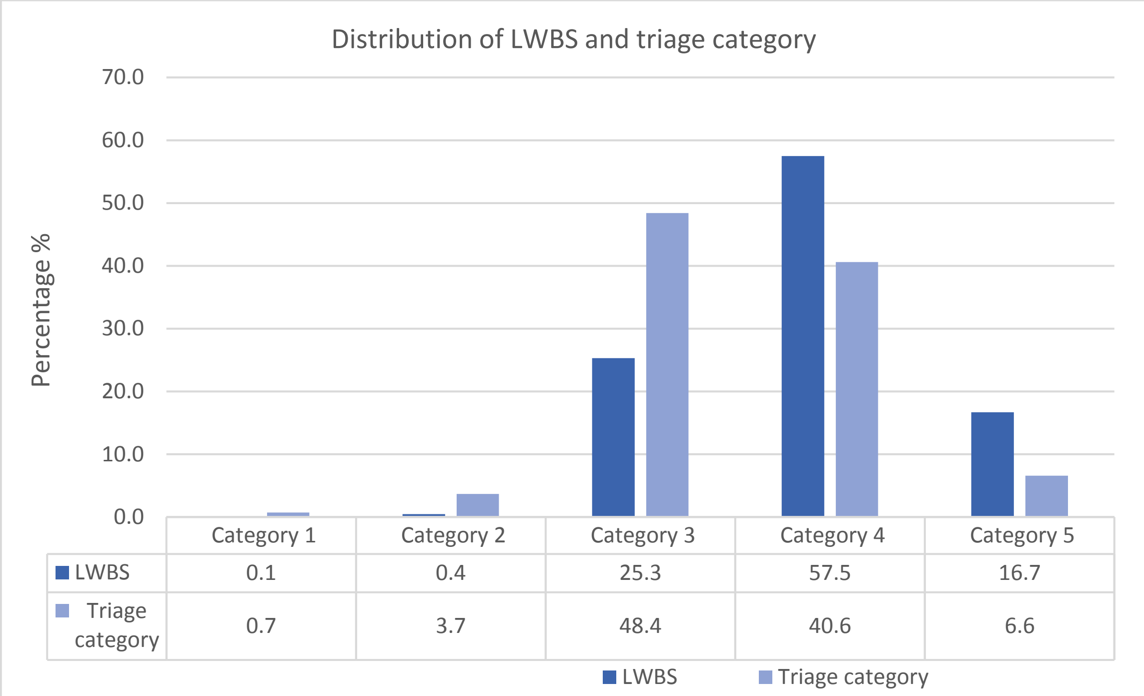


Figure 6-40 Probability distribution of LWBS and triage category

Finally, a field in the ED data provides information about delays for discharge or admission. These include bed delays, theatre delays, waiting to see a specialist, and waiting for transport. As *NHSquicker* currently does not provide real-time information about admission or discharge, these delays are calculated from the ED data by proportion of triage category as a mean percentage increase by comparing the mean LoS in ED of those given a 'delay reason' with the mean LoS for those without a delay.

This table can be found in Appendix 3, Table A3.11. This is important because downstream (hospital) delays will increase the ED LoS, and numbers in the department. The percentage increase, and the percentage of patients affected (per triage category) are used in the model in a 'delay' to replicate downstream delays for the appropriate proportion of patients. At the time of data collection, there were 8 chairs in the Clinical Decision Unit (CDU) that can accommodate a portion of these patients, however the remainder will contribute to crowding in the department.

The last stage of data collection required access to staff rotas, which proved difficult to access. With the help of a staff nurse, these were estimated, however there was a lot of uncertainty due to staff shortages (NHS, 2018). Further discussion with staff indicated that these changed week-by-week, and were unpredictable due to staff shortages, illness and last-minute changes. Rotas were estimated for consultants, junior doctors, nurse practitioners, and nurses across three shifts. Static figures were given for trolleys in triage (2); minors (14); majors (7); resus (3); and CDU (8).

Table 6.10 summaries the data which is available for use in the DES. With time, more real-time data is anticipated to be made available, for example real-time patient arrivals to initialise the model, admissions and discharges, and triage categories of arrivals. The next subsection (Step 5) outlines the development of the simulation model.

Table 6-10 Summary of data available for DES at initialisation and execution

Model element	Data	
	Model initialisation	**Model execution**
Entity arrivals schedule	Historical (distributions can be refreshed with real-time data)	Historical (distributions can be refreshed with real-time data)
Triage category	Historical data	Historical data
Resource availability	Historical or manual (distributions can be updated at initialisation)	Historical or manual (distributions can be updated at initialisation)
Number of services	Historical data	Historical data
Service processing time	Historical (distributions can be refreshed with real-time data)	Historical (distributions can be refreshed with real-time data)
Entity behaviour (LWBS)	Real-time data	
Queues	Real-time data	
Global variable values	Real-time data	
Decision-rules	Real-time data	

STEP 5 Computer modelling

This stage involves converting the conceptual model to an executable model (Robinson, 2004), using a programming language or a commercial package, in an iterative and stepwise approach. AnyLogic 8.5.2 PLE is used to build the model. This provides a visual interface and a flexible method, however the PLE has limitations to its functionality for implementing real-time simulation. In particular, it is not possible to call AnyLogic from a third-party application without exporting the model or uploading it to the cloud to send/receive data using an open API. This is discussed in more detail in Section 6.8.

There are two ways to ensure a simulation model collects data without being biased by an inappropriate starting state. The first is to set a warm-up period, the second is to set initial conditions. Generally, initial conditions are used to save time without the need for a warm-up period. However real-time simulations require at least some initial conditions to be set, using the real-time parameters. In this case, mixed initial conditions and a warm-up period are required (Robinson, 2004). The initial condition specifies the real-time starting condition for the model, while the warm-up period is required to initialise the rest of the model. The warm-up period is built into the simulation such that it initialises with the current date-time less the warm-up period, which can be changed if required without affecting the start time. The simulation is non-terminating, as ED is open 24 hours, 7 days. However its use here is for short-term decision-support, so it is intended for very short runs of 2-4 hours. This is important for the real-time data initialisation, as the effect of the real-time data will degrade quickly as the model runs. The model is transient, meaning that the distribution is constantly changing, as patient arrivals change throughout the day.

Figure 6-41 illustrates a flowchart of the DES model. Previously, Figure 6-34 mapped the real-time data to a higher-level process map. A screenshot of the DES is in in Appendix 3 (Section 3.3). Additionally, a detailed description of the DES is provided, as the model is used for demonstration in the *Evaluation* component of IHAF. The remainder of this section provides this information in summary format.

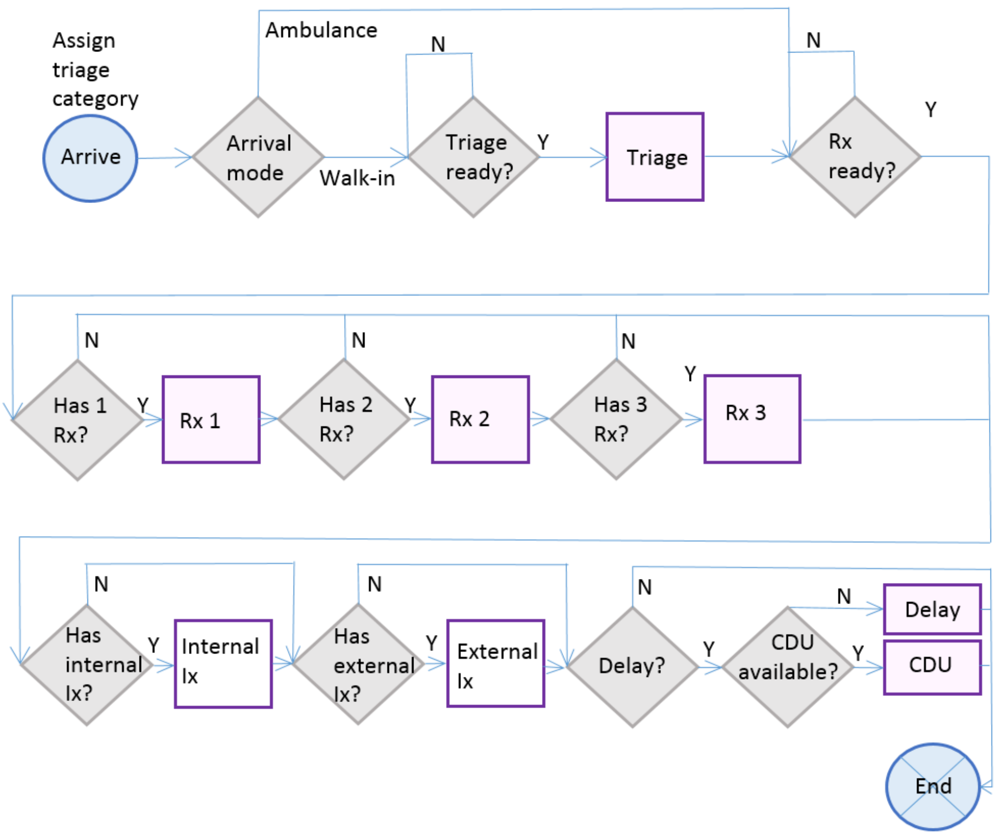


Figure 6-41 Flowchart of ED processes for DES. Rx= Treatment; Ix = Investigation.

As outlined in the Data Collection stage (Step 3), patients enter the model according to an historical hourly rate schedule, and are allocated a single parameter, a triage category, according to a historical distribution. Data collection starts at the current date-time (less the warm-up period). This allows the model to start collecting data using the appropriate arrivals distribution for the hour-of-day and day-of-week, and the appropriate resource schedule for all staff. As explained in the previous section, this is important because the simulation is intended to run for a very short time (2-4 hours).The capacity for each staff type is defined using estimated schedules of three shifts/day as provided by a staff nurse, and described in the previous section. Trolleys are static resources, with fixed numbers.

Upon entry, a conditional block allows patients to either enter the system or be sent to another hospital (not illustrated in flowchart). For the baseline model, this is switched off; for scenarios it is accessed at runtime using a user-control button and sliders.

At this point, patients are defined as 'walk-in' or 'arrive by ambulance'; those who arrive by ambulance or air ambulance bypass triage and go straight to treatment. Those who walk-in are triaged. This is acceptable as ambulance delays aren't captured in this model due to the focus on low-acuity patients, however these are an important part of system performance and capturing real-time ambulance handovers/delays would support a more flexible model for future work. Patients who walk-in are triaged.

Following triage, patients wait for treatment and investigations. These waits will form the real-time component '*Patients Waiting'*, which can be updated at initialisation. These are prioritised according to triage category, and defined according to the probability tables derived from historical data (Appendix 3, Tables A3-5 and A3-7) which defines the number of treatments and investigations per triage category. In consultation with staff, the treatment distributions, and the resource requirements tables were developed and are in Appendix 3 (Tables A3-6 and A3-8). External investigations (e.g. Xray, ultrasound scan or other scans) are assumed to use no resources, hence the external waiting time is built into the delay time distributions. This means that lower categories with lower prioritisation have longer service times. The updating of real-time values in '*Total Patients'* can be distributed in treatment and investigations by triage category, and treatment/investigation probability. For example, five additional patients in the department are most likely to be categories 3 or 4, and most likely to be attending their first treatment, or first investigation.

To capture the behavioural component 'Leave without being seen (LWBS)', patients can be in either a 'waiting' state or a 'not waiting' state. The transitions between these states occur when patients are waiting for a trolley (for initial treatment) and for triage categories 4 and 5, who may return to the waiting room between treatments and investigations. As patients undergoing treatment may be waiting for staff resources, they can also enter a 'waiting state' at this point. A wait-time tolerance per triage category is set. If it is less than the time already spent waiting, the patient leaves. Currently wait-time tolerance is set using user-control slider bars, calibrated with Table A3-9. However future work will investigate setting the LWBS function as a linear relationship with the real-time *Maximum Wait Time*, or *Total Patients*, as described in the previous section. Note in the flowchart in Figure 6-41, delays between treatments and investigations are not captured, however the simulation model calculates these queues to create cumulative waits, which can be updated with real-time *Maximum Wait*. At initialisation, the *NHSquicker* values can update the model using a Java timer and download program, which is implemented as a Class in AnyLogic. This is discussed further in Section 6.8. Additionally, the real-time status of the MIUs can be implemented as decision-rules, within Scenario 3.

A discharge/admission delay is built into the model. In the ED dataset, these are coded, for example, as *waiting for transport, waiting for a specialist, waiting for a bed*. The proportion of patients who are delayed are according to Table A3-11 in Appendix 3. It is anticipated that this data could be made available in real-time in the future for model initialisation. If a CDU chair is available, Categories 4 and 5 will take one, otherwise all patients retain their trolleys. Finally, patients exit the system.

STEP 6: Verification and validation

Verification involves checking for errors to ensure the model is operating as intended, while validation determines the correspondence of the model's behaviour with reality (Sargent, 2013). Most M&S frameworks position verification and validation (V&V) throughout the lifecycle of the modelling process (Balci, 1990; Rabe et al., 2008; Robinson, 2004). The model was checked after each change to ensure it was behaving as expected. Although intended for short-term use of 2-4 hours, validation was done using 7-day outputs (a 'long run').

The following outputs are plotted after a single run of 7 days, with a 3-day warm-up time. The warm-up time was chosen through visual inspection of time-series outputs (Robinson, 2004), which reached a steady state before 3 days using multiple replications. Final calibration was done after 150 replications of 7-day long runs. The number of replications was chosen using a simple graphical approach (Robinson, 2004), but with further model refinements, confidence intervals can better determine the accuracy of a mean average of a value being estimated, and the required number of replications.

Figure 6.42 plots the hourly arrivals for one week, to confirm that hourly arrivals are sampling accurately. Single runs using different random number seeds were undertaken to unsure that the hourly arrivals are not deterministic, that is, that the arrivals schedule is sampling from an hourly Poisson distribution.

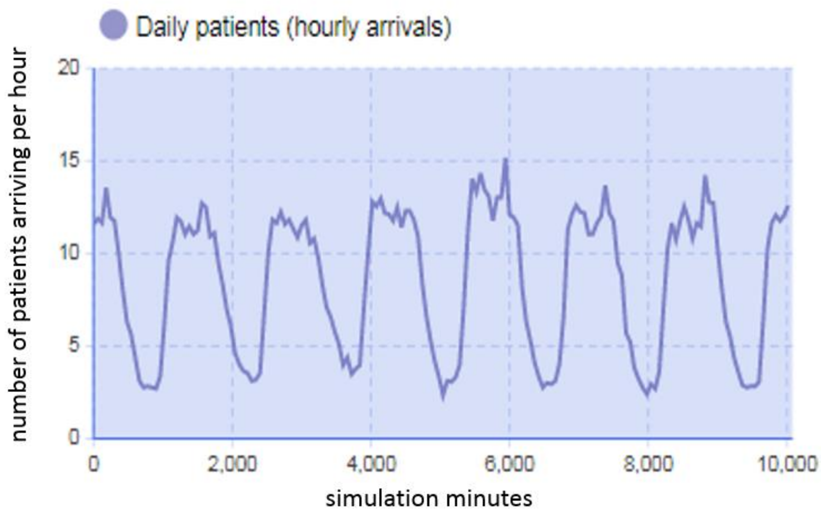


Figure 6-42 Daily patient arrivals (one week, minutes).

Figure 6.43 is reproduced from Section 6.3, and plots a subset of the series *Total Patients* and *Patients Waiting* (14 days = 678 observations each).

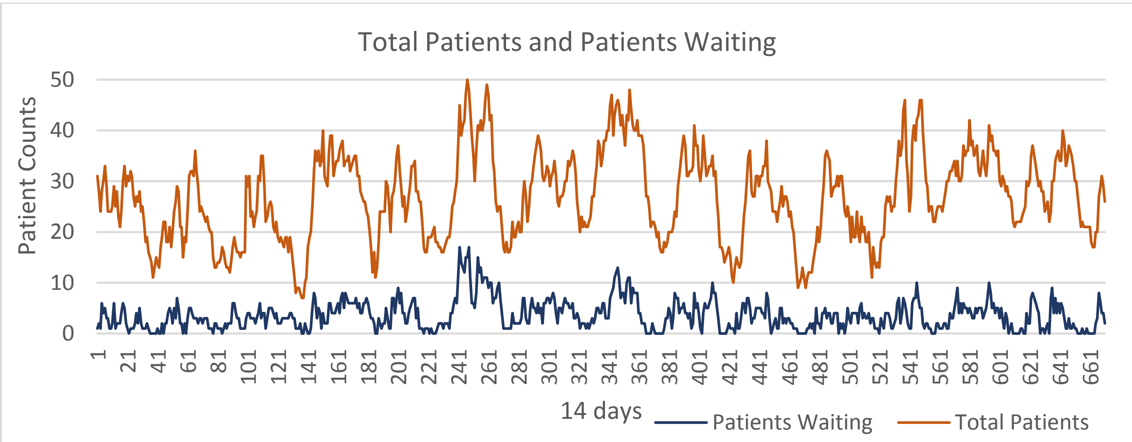


Figure 6-43 14-day plot of sample of Total Patients and Patients Waiting Data

This is helpful for visual validation of the simulation output with real-time data in the following plots.

Figure 6.44 maps simulated *Total Patients* (a single run of 7 days, initialised with a simulated 'real-time' data point) with average hourly *Total Patients* from real-time *NHSquicker* data. This 'long run' can be equated to 7 short (24-hour) replications for initial validation (Robinson, 2004), and appears to be sufficiently accurate.

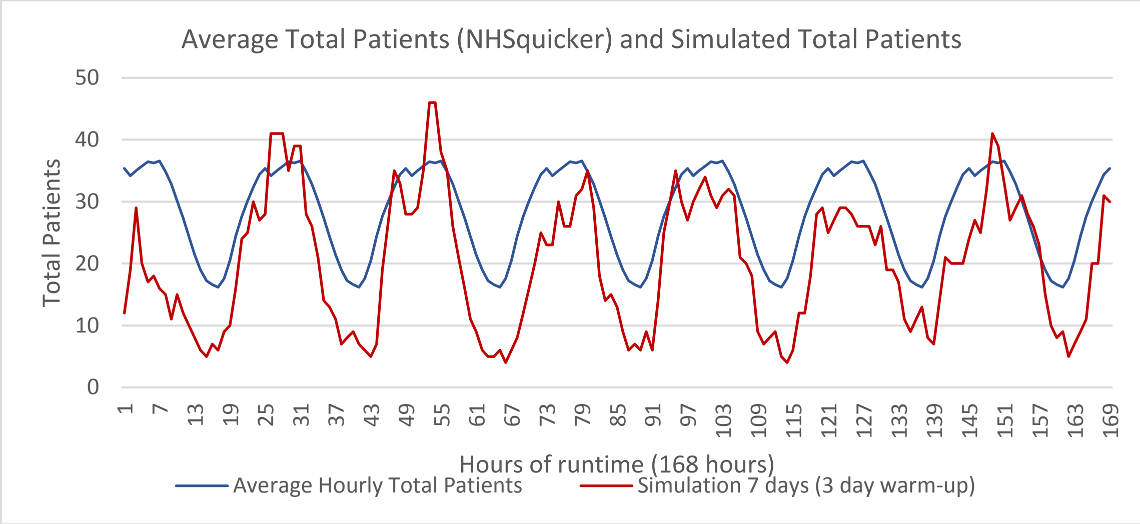


Figure 6-44 Average hourly Total Patients (2018) and simulated Total Patients

Figure 6.45 maps simulated *Patients Waiting* (a single run of 7 days, initialised with a simulated 'real-time' data point) with average hourly *Patients Waiting* from *NHSquicker* data. Again, this appears sufficiently accurate.

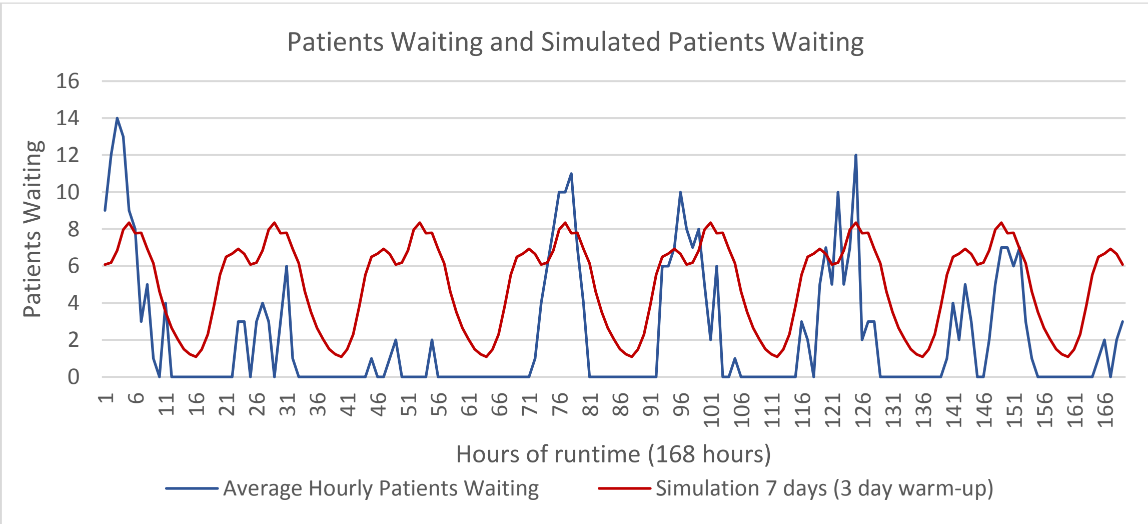


Figure 6-45 Average hourly Patients Waiting and simulated number of patients waiting for treatment

The simulation outputs *Patients Waiting* by triage category, and this is shown in Figure 6.46; the totals compare satisfactorily with *NHSquicker* data. No patients of Categories 1 waited for treatment, which is as expected, and very few for Category 2. Note that data is not displayed for the first 3 days (the warm-up period), however the warm-up period is retained in the plot as a visual reminder.

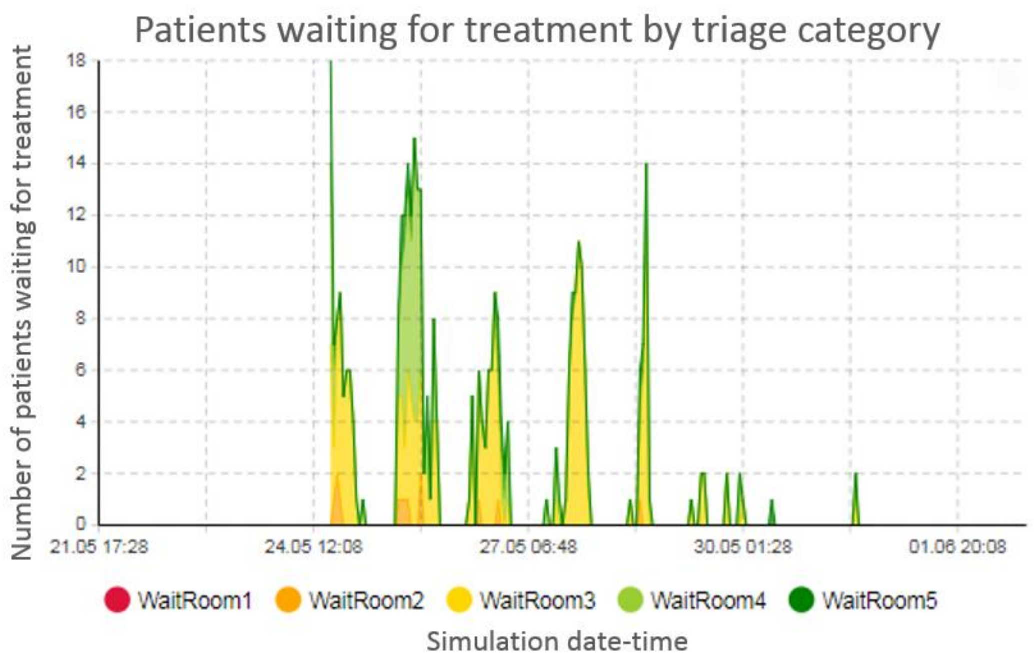


Figure 6-46 Patients waiting by triage category. y-axis = number of patients waiting, x-axis = simulation date/time

Table 6-11 summaries this information by minimum, maximum, average and standard deviation for both Total Patients and Patients Waiting over a 7-day period and demonstrates a sufficiently good fit for *Total Patients*, with some under-prediction of *Patients Waiting*.

Table 6-11 Summary statistics for NHSquicker and simulation output data for Total Patients and Patients Waiting (7 days)

Replications = 7 days	Total Patients NHSquicker	Simulated Total Patients	Patients Waiting NHSquicker	Simulated Patients Waiting
Minimum	3	3	0	0
Maximum	63	49	27	27
Average	28	21	4	3
Std Dev.	10.53	9.98	3.07	4.93

As simulation output is stochastic, i.e. it contains random events, MonteCarlo simulation was set up with 150 replications to confirm validation. 2D histograms were used to display the waiting room size (*Patients Waiting*) for each triage category, and the total number of people in the system (*Total Patients*). In this plot, each of 150 replications (different random seeds) are overlaid. Figure 6.47 shows the simulated *Total Patients*, with a 3-day warm-up period (no data collected) and a 7-day run period, and again demonstrates a good fit with the *NHSquicker* data (Figure 6.44).

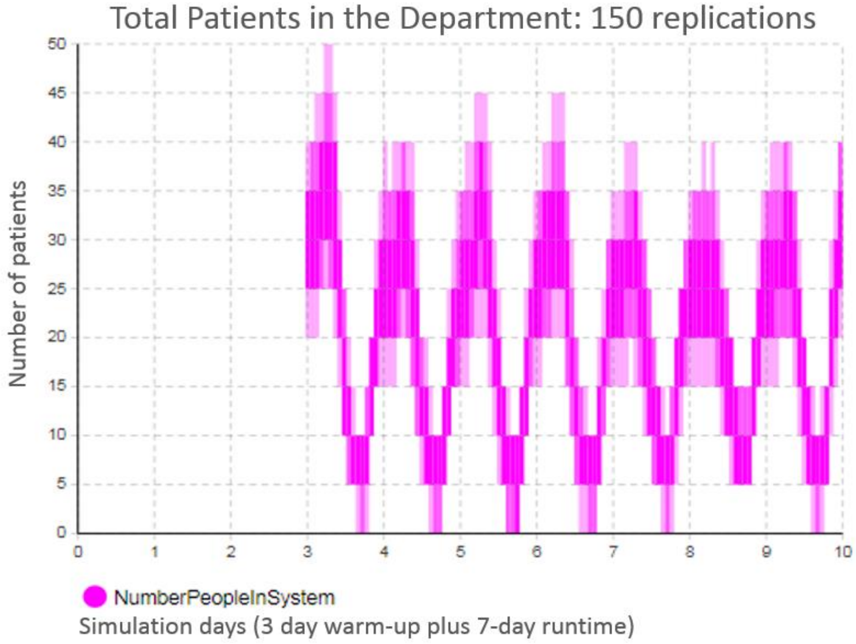


Figure 6-47 Simulated Total Patients, 150 replications of 7 days

The simulated *Patients Waiting* data has been plotted per triage category. Figure 6.48 illustrates with Category 4 waits in a 2D histogram, over 7 days.

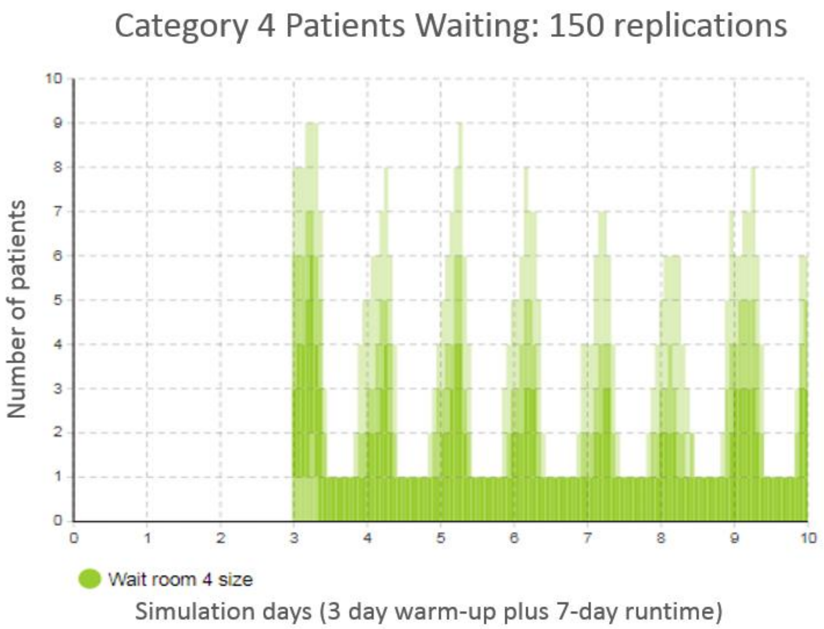


Figure 6-48 Category 4 simulated Patients Waiting, 150 replications of 7 days

This provides additional validation for model behaviour, as few Category 1 and 2 patients would be expected to wait, compared with categories 3, 4 and 5, as above. After 150 replications, a maximum of one Category 1 patient waits at any time, and a maximum of two Category 2 patients. This is to be expected, as all patients enter the 'queue' object before seizing a trolley and commencing treatment. Up to ten patients in Categories 3 and 4 are waiting at any one time, and up to 3 patients in Category 5, reflecting the relatively fewer numbers of these patients. The aggregated numbers align with summary statistics of *NHSquicker*.

To examine duration of *wait times*, a scatterplot for the minimum, average and maximum simulated *wait times* was produced (Figure 6.49), which plots the summary statistics from every simulation run.

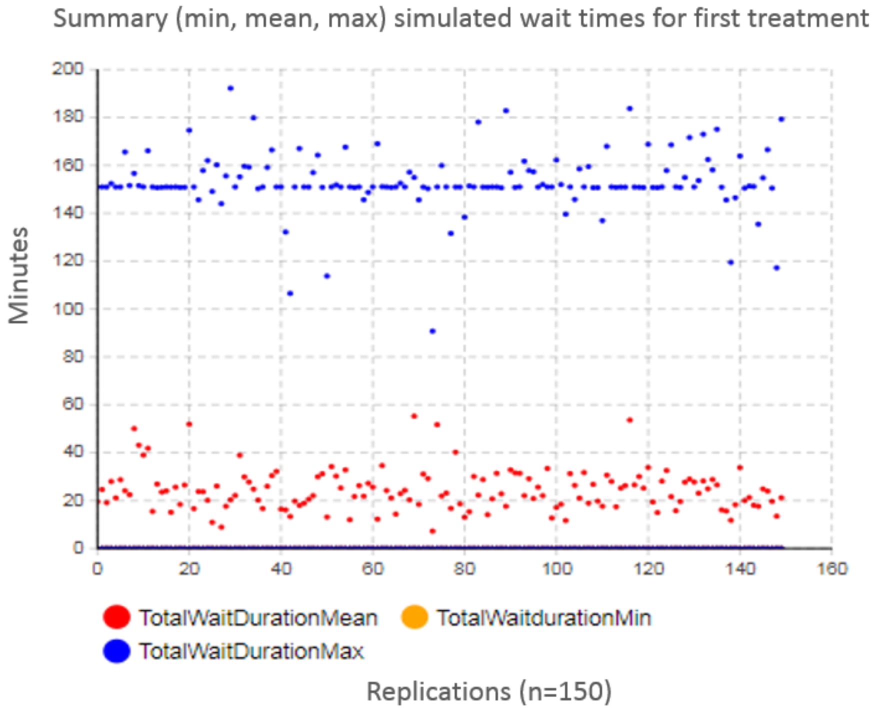


Figure 6-49 150 replications minimum, mean and average ED Wait Time for first assessment for each replication

This can be compared with Figure 6.50, which plots the wait times for first treatment of a sample of patients from the ED dataset. Wait times, while captured in the ED dataset as date-time stamps at key points, had significant missing data and errors (e.g. triage occurring after first treatment), hence the reduced subset plotted in Figure 6.50. It is worth noting however that *NHSquicker Maximum Wait* times are significantly higher than seen here, although as discussed in Section 6.3, recorded waits of up to 1222 minutes were concluded to be data errors. A sample of *NHSquicker Maximum Wait* time is plotted in Figure 6.51. Both can be compared with the simulated *maximum wait times* in Figure 6.49, and show sufficient accuracy.

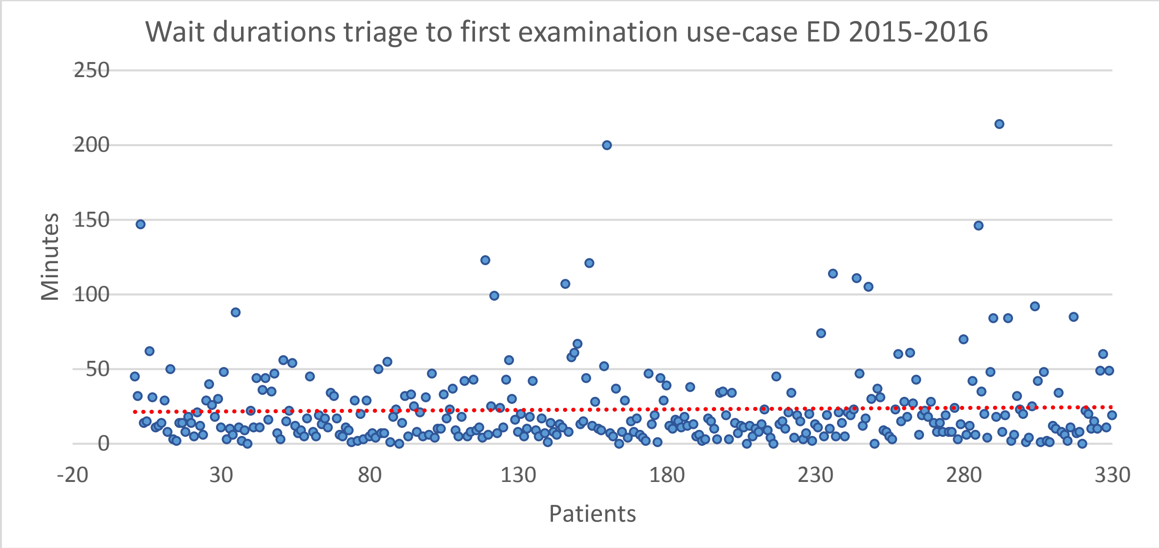


Figure 6-50 Sample 2015 Use-case data: Wait time for first assessment

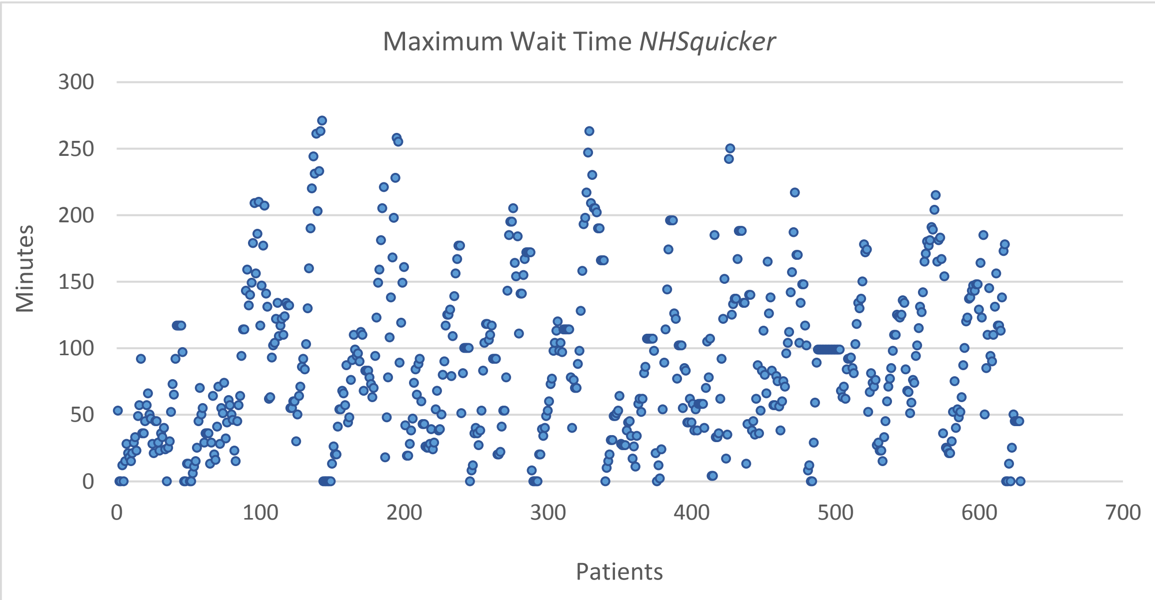


Figure 6-51 NHSquicker Maximum Wait time for first assessment

To examine length of stay (LoS), a scatterplot for the minimum, average and maximum simulated LoS was produced (Figure 6.52), plotting the summary statistics from every simulation run. This can be compared with Figure 6.53, the LoS of a sample of patients from the ED dataset. Here, the maximum recorded LoS is 800 minutes in the subset of 4000 patients plotted. While it is possible that patients might stay in ED for 13 hours or more, it is also possible that these are data errors, where a patient is entered into the system but not removed at the end of their stay. In the full dataset (70,000 patients) stays of up to 1683 minutes (28 hours) were recorded.

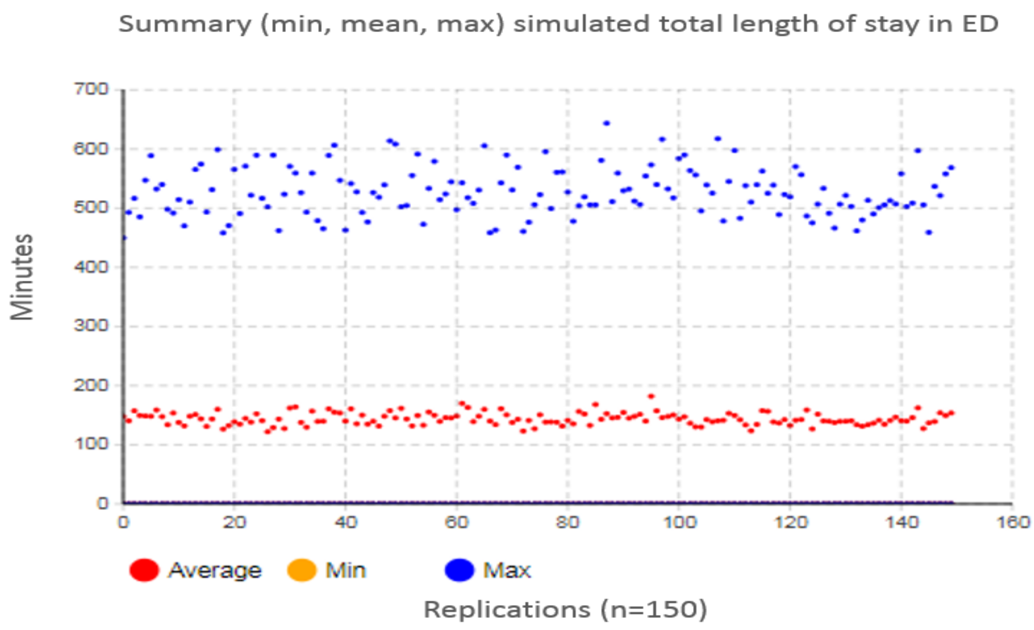


Figure 6-52 150 replications minimum, mean and average ED LoS for each replication.

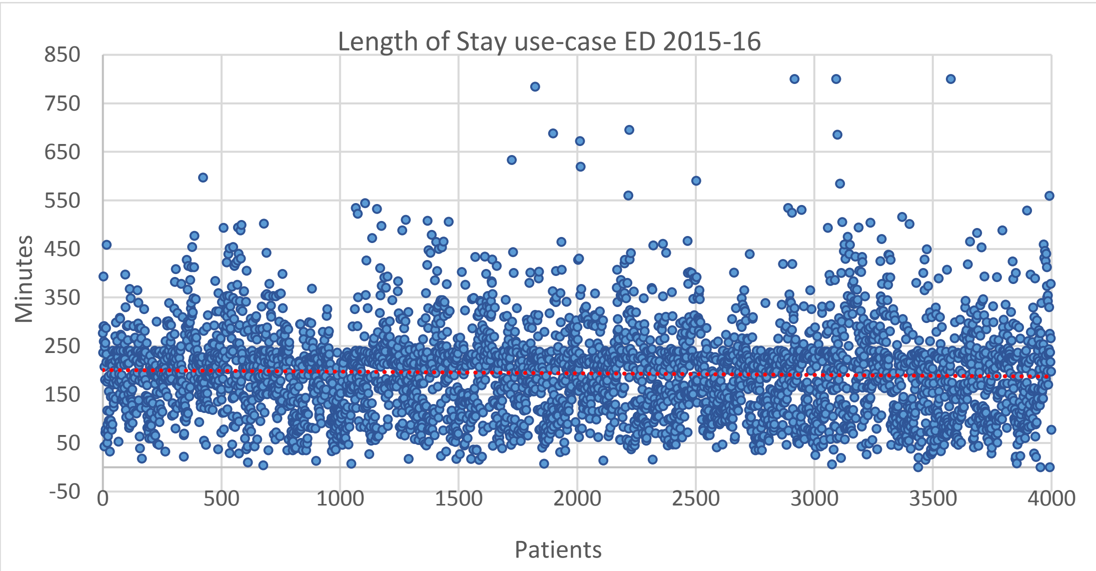


Figure 6-53 2015-2016 Use-case ED Length of Stay

Again, Figures 6.52 and 6.53 provide sufficient accuracy for LoS, apart from some uncertainty about the maximum waits. Table 6.12 summarises this information.

Table 6-12 Summary statistics Wait time for first treatment and total LoS

Replications = 150	Wait time (ED)	Simulated Wait Time	LoS (ED)	Simulated LoS
Minimum	0	0	0	0
Maximum	214	192	800	679
Average	23	24.3	194	141

The maximum 800 minutes LoS seen in the use-case data also increases the average LoS, however the simulated LoS in Figure 6.86 under-predicts by about 25%. This may be sufficiently accurate given the high possibility of errors in the ED use-case dataset, but is likely due to the estimated service times.

Patients who left the department without being seen (LWBS) were also plotted as histograms by triage category. As described in Step 5, patients who LWBS are given a maximum wait-limit using an interactive 'slider' for each triage category. The default slider settings are in Table 6.13 and were used for calibration. The average daily number of patients who LWBS are from 2018 ED use-case data. Simulated averages over 150 replications are in Table 6.13 and plotted in a combined histogram per triage category in Figure 6.54. Future work will investigate using real-time *Maximum Wait* times to predict LWBS.

Table 6-13 Patients who LWBS given a maximum wait limit calibrated to proportions found in ED data

	Category 1	Category 2	Category 3	Category 4	Category 5
Max. Wait limit minutes (calibrated)	131	193	150	145	100
Average LWBS per day (2018 ED data)	0.007	0.04	2.0	4.12	3.1
Average LWBS per day (simulation 150 replications)	0.07	0.03	7.14	5.12	5.04

.

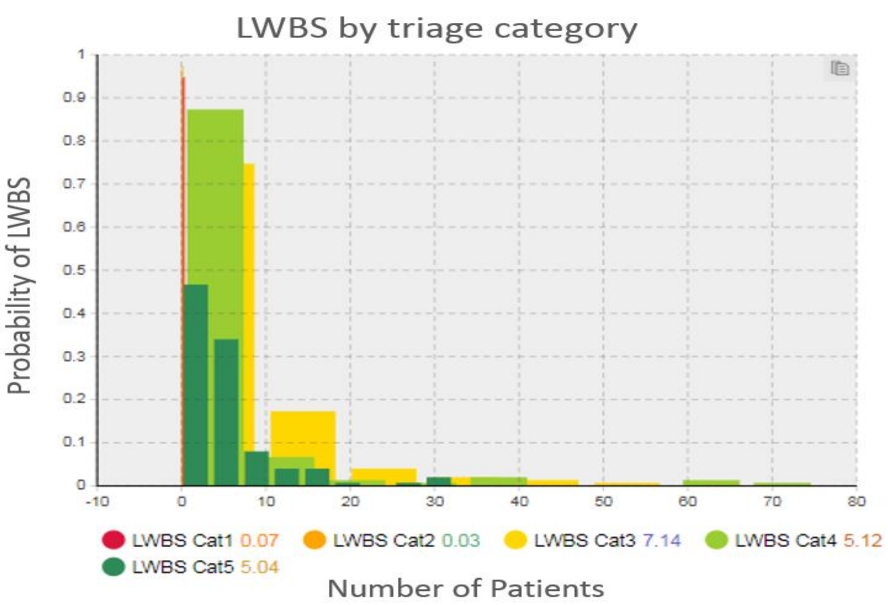


Figure 6-54 Patients who LWBS by triage category.

As abstractions of reality, simulation models cannot be described as absolutely accurate, however a valid model provides accurate outputs given a set of criteria. In this case, the model under-predicts the overall LoS by about 25%, possibly due to uncertainty of service distributions (Appendix 3, Table A3-6). The model may correctly predict crowding, or correctly predict no crowding, or it may represent errors which can lead to significantly reduced performance. These two types of errors are referred to as type I and type II error, respectively. In case of a type I error, the LoS will be over-estimated, and counter measures will be recommended to prevent the critical condition from occurring. In a type II error, the LoS will be under-estimated, leading to wrong assumptions. Depending on the application context, one type of error may be more serious than the other. For example, in some applications unnecessary interventions against a critical condition which does not occur might be considered better than failing to detect a critical condition which has considerable impact on performance of the physical system. In other applications, a critical condition may refer to sub-optimal performance which may interrupt a process flow. Therefore, the severity of each error type is context-dependent, and an appropriate error handling strategy should be chosen for each type of application. A subsequent model will be required to be better calibrated, for example using observation of processes and validating service distributions with a range of staff to lower the risk of type II error which is likely to reduce confidence in the model.

STEP 7 Model experimentation

Experimentation involves specifying scenarios and examining the output. The scenarios specified in Step 1 are indicative, but are not fixed solutions for crowding. Input, throughput and output factors can contribute to crowding. This implementation looks only at demand management (input factors), however a more flexible solution would be able to address throughput factors (e.g. adjusting staff rotas or improving other resource availability, or removing unnecessary process steps) and output factors (reducing delays, for example increasing bed base, changes to discharge protocols). The input scenarios tested are:

- Baseline – proportion of patients who LWBS (leave without being seen) per triage category, calibrated to 2018 data.

• Scenario 1 – Redirect all Category 4 and 5 patients when the number of patients in the department reaches hourly trigger (current scenario: reactive trigger). Simulation is initialised using simulated real-time data.

• Scenario 2 – Redirect a proportion of Category 3, 4 and 5 patients to MIU when the number of patients in the department is forecasted to reach the hourly trigger (in 2-4 hours' time, i.e. predictive trigger). Simulation is initialised using simulated real-time data.

• Scenario 3 – Redirect a proportion of Category 3, 4 and 5 patients to MIU when the number of patients in the department is forecasted to reach the hourly trigger (in 2-4 hours' time, i.e. predictive trigger), and given sufficient capacity in MIU. Simulation is initialised using simulated real-time data.

The baseline scenario has been described and validated in Step 6. The reactive (current scenario) and predictive triggers are conceptualised in Figure 6.55

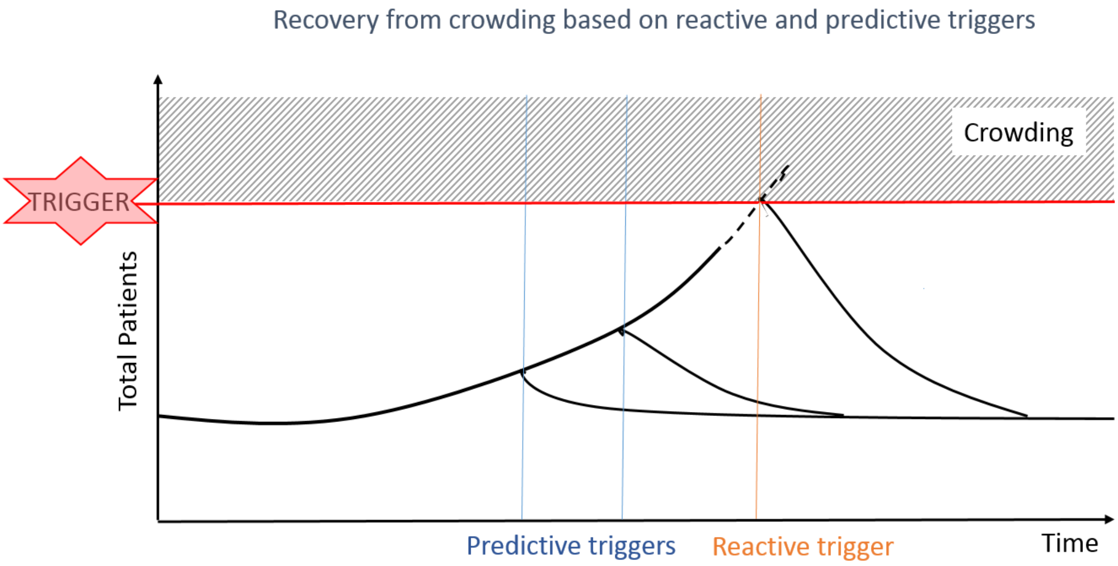


Figure 6-55 Conceptual mapping of reactive and predictive triggers for recovery from ED crowding

With an average of 205 patients per day presenting at the use-case ED, and given the triage category distribution specified in Step 3, the average numbers of patients per day per category are in Table 6.14.

Table 6-14 Average number of daily presentations per triage category

Category 1	Category 2	Category 3	Category 4	Category 5
1	8	99	83	14

Scenario 1

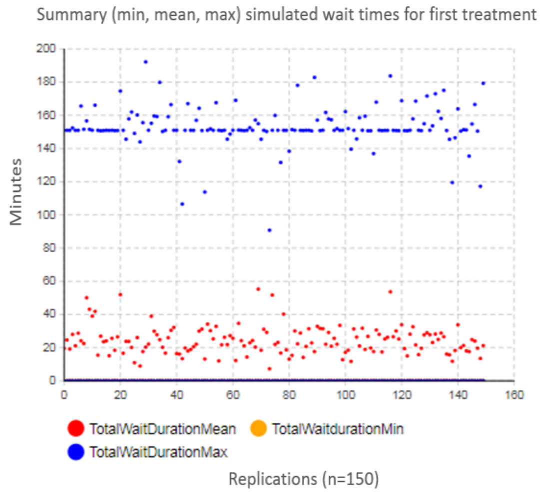
Scenario 1 is the current scenario, which is an ED escalation action which occurs when the operational state reaches maximum pressure (OPEL 4); this action is specified in the ED escalation policy. Upon reaching a given threshold of crowding, category 4 and 5 patients are redirected until the crowding situation has resolved. For scenario validation, this is assumed to be a 4-hour period.

The results show the average over 150 replications, and suggests that this is a successful strategy (Table 6-15). However 16 patients are redirected over a 4-hour period, to reduce the *Total Patients* to a maximum of 25, and the mean length of stay to 45 minutes. A proportion of Category 4 and 5 patients are likely to be best placed for treatment at ED, so while this strategy aims to control risk for the hospital overall, for some patients this may be potentially sub-optimal.

Table 6-15 KPIs for Scenario 1

KPIs for Scenario 1: Number of patients redirected = 16		
KPI	**Baseline**	**Redirect all Cat 4 and 5**
Total Patients	Max = 50 patients	Max = 25 patients
Length of Stay	Max = 550 min Mean = 150 min	Max = 420 min Mean = 45 min
Wait for initial treatment	Max = 150 min Mean = 20 min	Max = 140 min Mean = 10 min

The simulated wait time for the initial treatment is plotted in Figure 6.56 to provide an example. This shows the minimum, mean and average wait time for each of 150 replications. The minimum in both (a) the baseline; and (b) Scenario 1, is zero in each case, as some patients LWBS, or die in the department. The average wait (red dots) has dropped from approximately 20 minutes to approximately 10 minutes; and the maximum waits are lowered. As seen in the validation section (Step 6), the baseline scenario is performing accurately compared with *NHSquicker* data (*Maximum Wait*) and ED use-case data (wait times from triage to first examination), hence this scenario clearly improves waits. However the purpose of using a predictive trigger is to redirect fewer patients (Figure 6.55) as redirections occur prior to the onset of crowding. This is investigated in Scenario 2.



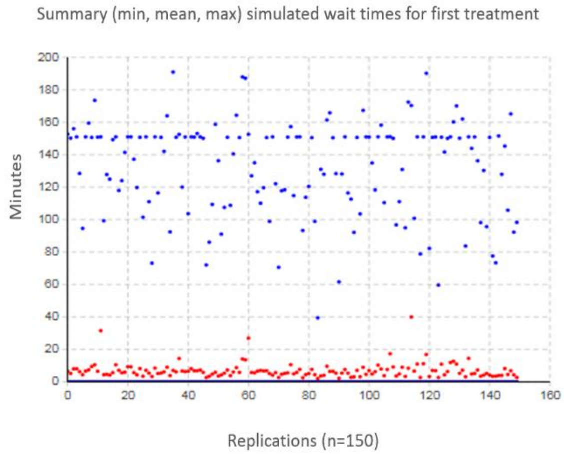


Figure 6-56 Summary statistics for 150 replications wait times for first assessment (a) baseline (b) Scenario 1

Scenario 2

Scenario 2 redirects 15% of Category 3, 30% of Category 4 and 50% of Category 5, given a forecast of *Total Patients* which reaches the hourly trigger in 4 hours' time. These figures are cautious estimates calculated in Step 3, and tabulated in Appendix 3, Table A3.9. This calculated the proportion of patients who 'could have gone to MIU' estimated by adding those who were coded as any of the following: 'Discharge – follow up treatment by GP', 'Discharge – no follow-up', 'Left department before being treated/Did not wait', and left department having refused treatment/self-discharged'.

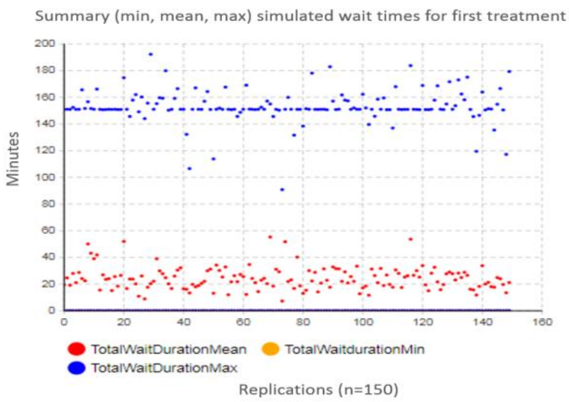
This policy redirects fewer patients, allows time to triage, such that the most appropriate patients are redirected, and using this proactive policy, aims to recover before the onset of crowding has actually occurred. *Total Patients* in the department is lowered to well below the 24 hour triggers identified in Section 6.4.3 (Table 6.16).

Table 6-16 KPIs for Scenario 2

KPIs for Scenario 2: Number of patients redirected = 8		
KPI	**Baseline**	**Redirect 15% Cat3; 30% Cat4; 50% Cat5**
Total Patients	Max = 50 patients	Max = 30 patients
Length of Stay	Max = 550 min Mean = 150 min	Max = 450 min Mean = 80 min
Wait for initial treatment	Max = 150 min Mean = 20 min	Max = 100 min Mean = 5 min

Lengths-of-stay reduce below baseline, and the wait for initial treatment has reduced markedly compared with Scenario 1. These are plotted in Figure 6.57 below. Again, the baseline for the waits for initial treatment as seen previously compares accurately with NHSquicker (*Maximum Wait*) and ED use-case data (wait for first treatment after triage), and Scenario 2 has reduced the mean to 5 minutes, and the (mean) maximum to 100 minutes. The variation for the maximum wait has increased markedly across 150 replications, but overall, very few people are waiting over 150 minutes, the baseline (mean) maximum wait across 150 replications. Additionally, only 8 people have been redirected in a 4-hour period – half as many as in the previous Scenario.

Note that these are average validations over the same time period (across 150 replications). Vastly different results would be expected during any 4-hour period in any given time-of-day or day-of-week. Additionally, the real-time initialisation is simulated (to average for time-of-day), and vastly different initialisation states can expect to change the results.



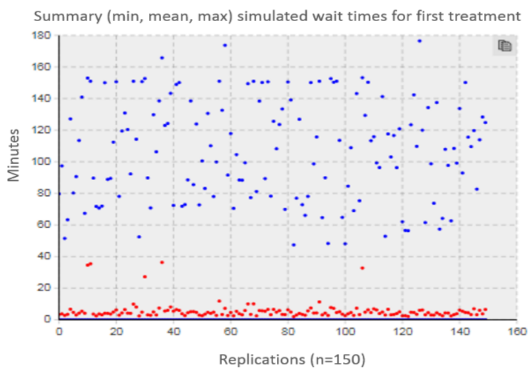


Figure 6-57 The change in patient wait times for a first assessment at (a) baseline, and (b) Scenario 2.

Scenario 3

Scenario 2 demonstrates significant reductions in the number of patients in the department, the length-of-stay, and the initial wait for treatment. However this scenario relies on redirecting low-acuity patients to the nearest MIU (Newton Abbott), which may already be at capacity.

Scenario 2 provides a predictive trigger, rather than a reactive trigger, where crowding has already occurred. Given 2-4 hours' notice, and given the real-time information provided by *NHSquicker* about the operational state across the UCN, the most appropriate MIU can be selected per patient, and the appropriate number of patients can be safely redirected. While MIU(1) (the current scenario) may be at capacity, it is still possible that the other two (equidistant) MIUs contain adequate capacity to accept patients over the ensuing 4 hours. Nonetheless, Scenario 2 results indicate that improvements can be seen with modest redirections over a short-time period.

It is anticipated that this may mean that fewer patients are able to be redirected (although feasibly more may be possible, as historical *NHSquicker* data indicates that MIUs frequently function below capacity). This means that Scenario 3 will perform at least as well as Scenario 2, at times better. This is future work, when the HM components are fully integrated, as discussed in the following section. In this scenario, the forecasts are generated, the simulation triggered and initialised with real-time data, and a decision-rule agreed for each of the MIUs to determine their real-time capacity. This takes a system-level view of the redirection scenario, allowing time to ensure the redirection policy is implemented safely. It involves patients both in the choice to wait or go elsewhere where they can be seen more quickly, and in the choice of facility. It also considers the available capacity in the MIUs so that patients aren't being redirected to a facility that is already at or near capacity.

Implications for IHAF implementation

A fully validated baseline DES model has been developed using historical data, and validated against real-time and historical data. Three scenarios have been investigated using proxy real-time initialisation (average for time-of-day and day-of-week) and executed using historical distributions. The next section outlines the current status and limitations of the integration of real-time data, the predictive trigger, and the real-time DES into a HM.

6.8 Integration component: simulation

As discussed in Section 6.6, the simulation model needs to be integrated with other components of the hybrid model for a seamless decision-support tool. The model is built using AnyLogic 8.5.2 PLE which has some limitations. The model can't be exported as a standalone application, nor can it be uploaded into AnyLogic Cloud to send/receive data from third-party applications. This is required for a Java application to call the AnyLogic model and pass it the real-time data parameters, execute the model, and receive the experiment results back to the Java application.

Setting the initial conditions requires some or all of the real-time data to be injected into the model.

- *Total Patients* can be distributed between each of the treatment/investigation sub-models by triage distribution and probability of treatment
- *Patients Waiting* can be injected into the wait to seize a trolley for first treatment by triage distribution
- *Maximum Wait time*, (and the maximum wait times in the historical ED dataset) contain inaccuracies, with patient waits of up to 28 hours recorded, and known errors with removing some patients from the system. However the average of maximum waits has a close linear relationship with LWBS as seen in Section 6.6, and these can be used to more accurately determine a wait time tolerance per triage category and time of day using a linear model. LWBS is an important indicator of ED safety and performance.

A 'workaround' method for integrating the HM components uses a download loop which is initialised on model start-up. The download loop acquires the real-time data from an *NHSquicker* URL every 30 minutes using a data download scheduler. These are written in Java and implemented in AnyLogic as classes, as groupings of data and methods. The data is parsed for the ED and all MIUs in the urgent care network into an excel file with worksheets for each facility. Finally, the model is initialised using the download loop, executed, and data from the file can then be injected into the model or read off the database.

All of the constituent parts are in place to integrate the model components into a single, automated hybrid model which updates every 30 minutes, creates forecasts and triggers the simulation when predicted thresholds are reached. This will be the subject of future work, which will be further informed by the subsequent evaluation in the next chapter. Future work also hopes to access additional real-time data which can strengthen the HM for short-term decision-making in ED.

This forms the first iteration of the HM component of this IHAF implementation, which can now be effectively demonstrated and evaluated in context.

6.9 Chapter Summary

This chapter has developed a hybrid model (HM) within the IHAF framework to address the second objective of the second research question, to apply the framework within a case study at an NHS Trust ED. It has done this by incrementing through the stages of the HM component of IHAF which are the *descriptive* component (identifying the data requirements and availability), the *diagnostic* component (identifying a trigger for the simulation model), the *predictive* component (developing a forecast model for a forecasted trigger) and the *prescriptive* component (a validated simulation model, using mixed real-time initial conditions and a warm-up period). All constituent parts are in place to integrate the real-time data, the forecasts and the simulation model into an automated HM. This is subject to future work, which will be further informed by the subsequent evaluation presented in the next chapter, to support the next iteration.

Chapter 7: Use-case evaluation of the application of IHAF

7.1 Introduction

Chapter 6 developed a hybrid model (HM) consisting of real-time predictions, a predictive trigger, and a real-time simulation model for supporting short-term decision-making at an NHS use-case ED, focusing on low-acuity patients and ED crowding. Chapter 6 concluded with all of the constituent parts, included a validated prediction model (SARIMA time-series forecasting) and a validated DES model with a set of example scenarios for balancing demand and capacity across the urgent care network by redirecting low-acuity patients. These scenarios have been tested in the HM components of the Integrated Hybrid Analytics Framework (IHAF), and the final component, evaluation is undertaken in this chapter. IHAF is illustrated in Figure 7.1, with the evaluation component highlighted.

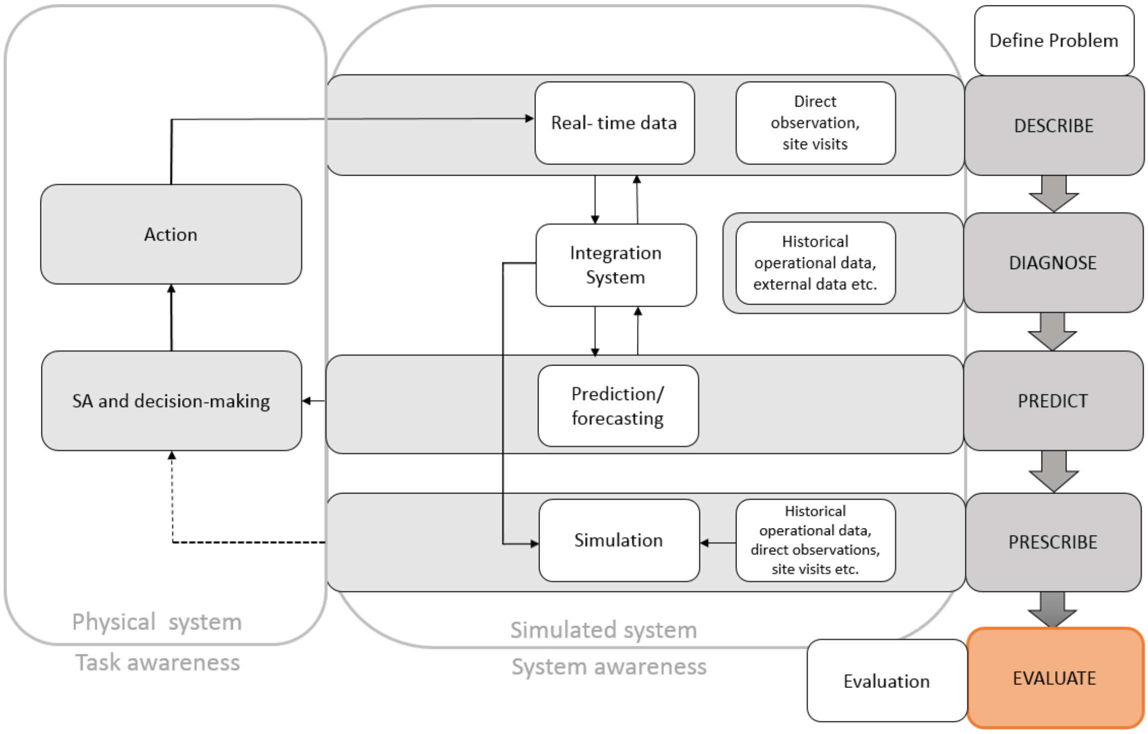


Figure 7-1 Integrated Hybrid Analytics Framework (IHAF)

This chapter addresses the second and third aims of Research Question 2, to evaluate the HM application in its context, and to evaluate the framework, IHAF, proposed in Chapter 4, as a conceptual framework for supporting the application of short-term decision-support tools in sociotechnical systems. It also addresses Research Question 3, to analyse the system-level impact of real-time data applications by both patients and staff to determine the implications and added value to the system.

The chapter is structured as follows: Section 7.2 implements the evaluation component of IHAF through staff interviews; Section 7.3 synthesises the interview findings with the patient questionnaires (Chapter 5) and the extant literature to draw conclusions about the use of real-time decision-support tools in healthcare and how these might be generalised to other sociotechnical systems. The following section outlines the justification for using semi-structured interviews, the development of the interview guide, and the results of the analysis of the interview data.

7.2 Evaluation component of IHAF: Staff interviews

While there is increasing interest in the use of real-time decision-support tools in healthcare, including real-time simulation, there is still a gap in understanding what actually does work in practice. To realise the potential of such tools using IHAF (Chapter 4), it is considered to be beneficial to evaluate iterations of implementations to inform future applications. ‘Adoption’ can be seen as a process rather than a discrete event, that comprises both ‘formal’ organisational decisions and a series of ‘informal’ decisions by individual users and teams, which intends to ultimately lead to the assimilation of the application into routine practice (Robert et al., 2010). The evaluation component aims to uncover aspects of these decision processes, as the value of such an application can only be realised through successive stages of implementation and utilisation.

A similar approach is common in the field of Information Systems (IS). However, models such as the *Technology Acceptance Model* (TAM) (Davis, 1993; Dixon, 1999) and the *Unified Theory of Acceptance and Use of Technology* (UTAUT) (Venkatesh et al., 2003) focus on individual perceptions of usefulness and usability, but fail to account for organisational and clinical environments which have been shown to influence implementation (Greenhalgh et al., 2017), or to account for a diverse set of users and tasks in healthcare (Ash et al., 2003; Callen et al., 2008). A number of researchers have promoted the use of qualitative research methods to evaluate health IS and explore uptake (e.g. Ash et al. 2003, 2005; Callen et al., 2008). These in-depth analyses of how clinicians and managers use and adapt clinical IS into their existing work practices reflects the complexity of implementing new digital interventions in healthcare. Studies have revealed the need to cover multiple issues in developing these applications, and highlight the inter-relationships between technology, people, and organisational issues. To address challenges, Shachak et al. (2019) recommend a range of strategies such as defining and achieving 'value-added use', evaluating implementation in context, and establishing common-ground amongst end-users. This perspective moves beyond the narrow scope of adoption or acceptance according to individual beliefs and opinions, toward increased attention to the role of human and organisational factors in a sociotechnical system.

Kukafka et al. (2003) proposed a multiple-factor approach based on prominent models of behaviour change, which determines that tools must not be simply functional, but must be compatible with the user population and aim to satisfy most user needs. Researchers have discussed the link between the implementation of new technology in health, and organisational change (e.g. Pfannstiel & Rasche, 2017). Much of this work incorporates information on why people resist change, and strategies for overcoming this resistance, for example matching technology to the correct level within the organisation, and understanding how digital innovations diffuse through healthcare. For example, Callen et al. (2008) used interviews and participant observation to propose an evaluation model which accounted for organisational, clinical unit, and individual contexts. At the organisation level, cultural- and analytics-maturity influence attitudes and support. At the clinical unit level, the needs and work practices, previous experience with IT, and support from management are factors, while at the individual level, knowledge, skills and experience are relevant. This approach takes a sociotechnical perspective, acknowledging the complexity and diversity of clinical and organisational environments at multiple organisational levels. However the specific value proposition of the application, at both the supply-side and the demand-side, requires attention throughout the process (Greenhalgh et al., 2017; Shachak et al., 2019), and efforts should also be directed at increasing the usability and flexibility of the intervention, adaptation to the needs of different users, and workflows (Liberati et al., 2017), its interoperability with existing systems, and effects on existing workflows and workloads (Ross et al., 2016). Informed by these studies, the approach taken for evaluation of the HM in this thesis addresses sociotechnical factors through the use of staff interviews. It addresses the generic criteria identified through the literature (Chapters 2 and 4), and specific use-case criteria identified through observation and patient questionnaires (Section 7.2.2). Section 7.2.3 outlines the data collection, Section 7.2.4 presents the analysis and results of the evaluation, and a summary is provided in Section 7.2.5. The next section briefly looks at measures of situation awareness (SA), to relate the chosen evaluation method with a measure of SA.

7.2.1 *Measuring SA within evaluation*

The evaluation centres on the potential for the IHAF method to support staff short-term decision-making by the development of a model to enhance SA. For this reason, some measure or determination of the effect of the HM on SA is required in the evaluation. SA has been measured using a variety of tools, including *Situation Awareness and Global Assessment Tool* (SAGAT) (Endsley, 1995) and *Situation Present Assessment Method* (SPAM) (Durso et al., 2004). These tools involve presenting probe questions to participants to measure accuracy and reaction time during a simulated work process. A well-known tool is NASA-TLX (Hart & Staveland, 1988), which is a validated measurement questionnaire commonly used for individual SA, capturing task-based perceptions of workload demand and performance. The above methods are used within field experiments where a level of control is required over the research design. Other methods of data collection for individual SA include recording conversations and communications during work activities (Rafferty, Stanton & Walker, 2013; Stanton, Salmon & Walker, 2015). These concurrent-style techniques bring risks associated with disruption to real-world operations.

In contrast to individual SA, distributed SA (DSA) is often measured in real problem settings (Fioratou et al., 2010). DSA views SA through a systems perspective, rather than a cognitive psychology lens, such that the whole system holds information, whether human, teams or technology. A DSA method described by Stanton et al. (2006) is the *Critical Decision Method* (CDM), which uses retrospective cognitive probes in semi-structured interviews to elicit information about how experts make decisions in sociotechnical systems (Klein et al., 2008). It sits within the Naturalistic Decision-Making paradigm (see Chapter 2, Section 2.4.1), and provides very efficient data collection compared with other methods such as grounded theory (Harenčárová, 2017). Naturalistic Decision-Making describes dynamic decision processes occurring in the field, which are often time-constrained and based on uncertain information, and where decisions have high stakes (Klein, 2008). In these situations, SA is an important part of decision-making, but situational information is often not presented optimally. CDM aims to uncover how and what information can be provided or optimised, rather than deliberating between alternative courses of action. However CDM isn't a suitable method for evaluating an artefact or information system unless it is already integrated and in use.

Jeffcott and Mackenzie (2008) described different methodologies used to capture team performance in healthcare, including surveys, direct observation, and video-based analyses performance. Gillespie et al. (2013) measured DSA using interviews and field notes, exploring the type of information staff perceived was needed to support decision making. A similar analysis was used by Casimiro et al. (2015) for analysing the factors that facilitate teamwork and effectively engage patients and families. Unlike the concurrent methods described in the previous paragraph these retrospective approaches don't disrupt workflow in an ED during critical times, and are therefore appropriate for evaluation in this application. The staff interviews will be informed by the CDM by focussing on a critical situation, when examining factors related to the criteria for evaluation. The next section is a reminder of the criteria for evaluation, identified in Chapter 4 (Section 4.3) as part of the 'criteria definition' stage of the chosen methodology.

7.2.2 *Criteria for evaluation of case study application of IHAF*

As discussed in Chapter 4, the methodology advanced by Blessing and Chakrabarti (2009) iterates through a series of stages. The first stage, *Criteria Definition*, identified the criteria for evaluating an artefact. Generic criteria for evaluation of a real-time decision support tool were categorised in Chapters 2 (literature review) and 4 (Phase (a) *Criteria Definition*).

The second stage, *Descriptive Stage I*, required identifying the influences on success, how these influences interact, and how they can be measured to improve the design process. Influencing factors are considered to be inter-related, creating a network of causes and effects connecting influencing factors with evaluation criteria. Those from the literature were identified, and these are listed in Chapter 4, Section 4.3 under '*Descriptive Stage 1'*. Additionally, specific criteria applicable to the use-case application were identified from the observational data (introduced in Chapter 5, and in Appendix 2) and the patient questionnaires (Chapter 5).

The criteria and influencing factors can be summarised as a series of themes to inform the development of the interview schedule, which forms *Descriptive Study II* in the Design Science Methodology, and the final stage of IHAF. The evaluation determines the degree to which the application has the expected effect on influencing factors, and whether these factors contribute to success, providing feedback for further development, and to enable conclusions to be made regarding the conditions under which the model was or was not successful. A reflective understanding of its limitations and how it is being used can ultimately increase the level of trust and confidence toward successful implementation. In this application, demonstration of the HM forms part of the evaluation phase, as described in Chapters 3 and 4. At the end of this activity the researchers can decide whether to iterate back to the previous activity to try to improve the effectiveness of the artefact or to continue on to communication and leave further improvement to subsequent projects.

The following criteria have been summarised, which informs the interview guide:

Table 7-1 Criteria for evaluation of application of IHAF

Broad criterion	**Specific criteria**	**Example**	**Reason**
1. Usefulness of the information provided by the HM.	• Patient Experience • Staff satisfaction • System efficiency	Joint demand-capacity planning across the urgent care network.	To take a system-level understanding of what matters in practice. A QI perspective can enhance the relevance of the study.
2. Confidence in all aspects of the HM for decision-support (data, trigger, predictions, DES).	• Task-level (adaptive behaviour) • System-level (escalation responses)	How real-time data and other decision-support tools are currently used by the ED to support task- and system-level actions.	To understand how the HM can support existing knowledge about what is happening, or is likely to happen.

3. Potential unintended effects.	• How the information is presented • Effects on workload • Fit with workflow • Usability and functionality • Effects on patient attendance decisions	A poorly designed tool can increase stress and workload, rather than reduce it.	As an integrated, recurrent-use support tool, it is important that potential unexpected effects or uses are understood.
4. Supports all three levels of SA.	• Perception • Comprehension • Projection	Automation and integration of components supports SA without interrupting workflow.	SA requires the perception of environmental information, the comprehension of its meaning, and a projection about the future based on this knowledge. This information should be comprehended by staff without interrupting workflow.
5. Barriers to use at all levels of the organisation.	• Time • Capacity • Politics • Resistance to change • Individual factors • Technology readiness • Sociotechnical context • Model maintenance	Barriers to the development of the tool will be encountered, but barriers to sustaining it also should be considered.	Potential barriers should be identified early and require understanding and managing toward implementation of the HM.

7.2.3 *Data collection*

For this thesis, semi-structured interviews were chosen, as this is an exploratory evaluation, and opening up the questions for detailed responses is a priority. Focus groups were not used, as a consensus wasn’t sought. It is important to gain a genuine understanding of the worldviews of participants, a combination of NHS clinical, managerial and information technology (IT) staff. This is a convenience sample, due to the need to work with NHS staff availability. While a range of staff viewpoints were sought, these are not considered to be representative (Liberati et al. 2017). Participants included end-users of the HM (healthcare staff) as well as the other staff that play an important role in shaping the structural and political underpinning of the HM adoption, such as IT staff and members of the hospital executive team. Including these staff members in the sample aimed to allow the exploration of clinicians' willingness and ability to use a new technology, as well as the impact of a broader initiative to support operational decisions using data analytics, making short-term decisions less discretional.

An initial sample of 12-15 participants was sought, aiming for approximately equal numbers of clinical, IT and senior management. However, consistent with the principle of "theoretical saturation" (Rowlands, Waddell & McKenna, 2016), the final number of participants was to be decided in the course of data collection, based on preliminary analysis of a sub-sample of interviews. Interviews took place in February and March 2020, hence were interrupted by the spread of COVID-19 in Devon. Six interviews were completed. A further six were cancelled or unable to be scheduled. Due to uncertainty surrounding COVID-19, clearly shifting priorities, and that the healthcare sector would undoubtedly be highly impacted for an unknown time period, it was decided to work with the data collected. Table 7.2 summarises this information.

Table 7-2 Interview participants: numbers completed and cancelled

Participant	Number completed	Time of interview (minutes)			Number planned
Doctors	2	59m	93m		4
Nurses and nurse practitioners	0				2
Executive Management	3	24m	58m	34m	4
IT staff	1	67m			2
TOTAL	6	5 hours, 35 minutes			12

While CDM is used for decision-analysis, providing rich data on demanding incidents, it is not a method used for evaluation. However in this case the interviews are informed by the CDM method to focus the interviews on specific crowding incidents when ED is likely to be at its most demanding. This is when the effectiveness of analytic support is likely to be at its most critical (Wong & Blandford, 2002).

The interview schedule, which guides the direction of the interview; its development, informed by the literature review, observations, and the patient questionnaires; and the analysis are presented in the following sections. The interview guide is in Appendix 4. It starts with focusing on a critical situation and decisions made during this situation, and centres on:

(a) Participants views and experiences with existing real-time decision support technologies in ED and their contribution to task and system-level decisions;
(b) Specific beliefs and experiences for staff with the currently available real-time data through *NHSquicker*;
(c) Confidence in, and perceptions of the value and usability of the components of the HM, including potential unexpected effects and barriers to use;
(d) Beliefs and experiences of the value to patients of the real-time data component, and its effects at the system level;
(e) Perceptions regarding the potential of real-time decision-support tools to provide information which can add value to supporting short-term decision-making and SA during critical periods in ED.

All interviews were started with the broad statement '*Think about the last time ED was under pressure, and you felt that there were potential risks to patient safety'*, followed by a series of questions, for example, 'What information is available at the time of the decision? This allows participants to produce accounts of incidents which can then be discussed in the context of existing real-time decision-aids, the demonstrated HM, and future iterations of the HM. Props were brought to the interviews for explanation and demonstration of the different elements of the HM for evaluation. These are included in Appendix 4, and include: plots and figures to demonstrate the hourly trigger (Chapter 6, Section 6.4.3); plots and sample output of the forecasts (Section 6.5); screenshots of the simulation model and example scenarios; outputs of the simulation scenarios (Section 6.7.1).

The interview guide was used flexibly and adapted to the different professional roles. For example, while clinicians were prompted to reflect on their first-hand experiences of using real-time tools (e.g. the ED Dashboard, *NHSquicker*), hospital managers were asked to discuss organisational strategies with respect to the tools. Managers and IT staff were also encouraged to reflect on potential ways to tackle resistance to real-time decision-support tools, as well as their own responses to the same critical situation. Audio-recorded interviews were 30-90 minutes (Table 7.2), and conducted on the hospital site. Signed, informed consent was obtained by all participants (Appendix 4). Recordings were subsequently transcribed for analysis.

7.2.4 *Analysis and Results*

Interview transcripts were analysed using NVivo 12.0. Thematic analysis was structured using the *a priori* criteria defined in Section 7.2.1 (Table 7.2) as themes, and the main sub-themes were obtained from the data, as outlined in Tables 7.3 – 7.7 and discussed in this section. Where new themes were identified from the data, they are included in the tables and discussion. Excerpts from the transcripts are categorised according to themes and subthemes and narratives are used to summarise and conceptualise the process. The output from NVivo is summarised in the hierarchy chart in Figure 7.2.

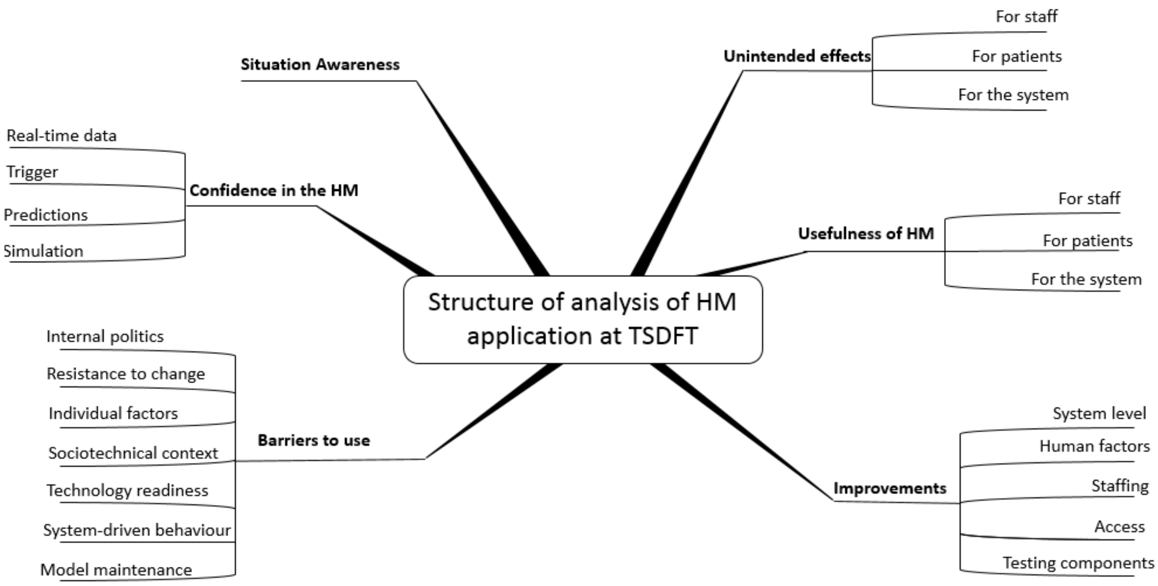


Figure 7-2 Tree hierarchy chart to identify prominent themes. Child nodes are nested in parent nodes

The results are presented below.

Criteria 1: This criteria considers the usefulness of the information provided by the HM: for patient experience, staff satisfaction, and efficiency/cost savings across the system. From the interviews, the usefulness of the model focused more on the benefits for staff, than for patients or the system as a whole. Table 7.3 provides a summary of themes and subthemes for Criteria 1.

Table 7-3 Criteria 1 themes and subthemes

1. Usefulness of the HM		
1.1 Patients	1.1.1	Patient experience
	1.1.2	Patient education
1.2 Staff	1.2.1	Support adaptive behaviour
	1.2.2	Support escalation actions
1.3 Efficiency/cost/system	1.3.1	Urgent care network

For patients, there were two foci: patient experience, and patient education. Patient experience (1.1.1) was interested in the effect on patient experience of interventions, such as the simulation scenarios investigated, for example*:*

> "*So if you're in the system, does it feel better here [*Scenario 1*] than here [*Baseline*]? Is that what you're saying?"*

Although safety concerns were not raised, this was an example of an attempt to maintain a focus on the experience of the users of the system. A second example focused on patient health-seeking behaviour, and the use of healthcare operational data for supporting attendance decisions, which was seen to improve the patient experience. Patient education (1.1.2) centred on behaviour, and the need to educate the public about alternative facilities to ED.

One purpose of the HM is to support adaptive behaviours (1.2.1) in ED, and staff were interested in investigating how adaptive behaviour might change given forecasts of crowding, rather than working reactively. While staff are required to adapt to keep pace with changing workloads and most participants recognise this, it was also recognised that human nature plays a role in how effectively this behaviour occurs, for example:

> "*They [consultants] don't do that. They won't. I think the nurses probably do that a lot, they move their band 5s and 6s around, and they respond to what they see going on. I don't know what information they use, probably only what they see."*
>
> "*That's something we struggle with. Persuading staff to ramp up is always a problem."*
>
> *"The ED clinicians will talk about how great they are in a crisis, which they are, but when we are in OPEL 2 or 3 [*Operational Pressure Escalation Level, 1=low; 4=high*], they should be staying on it, but they don't, they relax, and then of course it goes off."*

Generally the value of the HM for supporting escalation actions (1.2.2) was clearer, although low-acuity patients were of low concern, despite early conversations with staff about the effects of low-acuity patients on crowding. Re-directing minors patients is only one of a set of actions, many of which involve mobilising resources from other parts of the hospital, however this solution can be problematic, for example:

> "*We'd be told 'we need extra staff in ED' … we'll cancel the training… send 2-3 people up, and they would come back and say 'I literally haven't done anything…I haven't been made use of'*."

Here, the overwhelming emphasis was on supporting patient flow through the whole hospital system, where ED was seen as only one part of it. Discussions with ED were linked to OPEL status, which moves from 1 (lowest) to 4 (highest) and involves a range of measures keeping the status updated in real-time. However pressure has been high for some time, and there is a perception that being in OPEL 4 has lost its impact. Staff emphasised that long-term budget cuts were part of the problem, for example:

> "*There's a question about have we taken too many beds out of the system because there just isn't enough room to put people… you can't get them out of ED…"*
>
> *"I get emails. It's OPEL whatever…Everyone ignores it…Because they all know what the problem is. They know that this current problem is the making of the managers who decided to close the beds".*

This is a further issue with escalation actions, i.e. who drives them, and how they are communicated. Currently, escalation requirements are communicated via the intranet, which is accessible via desktop computers, inconvenient for many staff, for example:

> "*The whole organisation has access to the OPEL status, but it requires somebody being on a computer in a fixed place. The majority of people at any one time will be looking at clinical data, so why would they go and look on the intranet…people don't take much notice of it."*

At the system level (1.3.1), while participants indicated that there was interest in what was happening across the urgent care network, this was generally with a view to supporting their own internal performance, rather than optimising the performance of the entire network.

Criteria 2: Task level (adaptive behaviour) and system-level (escalation response) confidence that there is sufficient reliability and accuracy in all aspects of the model (data, trigger, predictions, DES). How other decision-support tools are currently used for decision support. The main themes are summarised in Table 7.4.

Table 7-4 Criteria 2 themes and subthemes

2. Confidence in model	
2.1 Real time data (*NHSquicker*)	2.1.1 Uncertainty of terms 2.1.2 Driving patient behaviour 2.1.3 Informing staff
2.2 Trigger	2.2.1 Total patients as a proxy for crowding 2.2.2 Hourly trigger
2.3 Forecasts	2.3.1 For ED 2.3.2 For the hospital 2.3.3 For patients
2.4 Simulation	2.4.1 Staff flexing 2.4.2 Staff rotas 2.4.3 Staff shortages 2.4.4 Scenarios

There is some concern, despite extensive engagement activities, that some of the terms used within the public-facing app were uncertain (2.1.1), for example

> "*[*NHSquicker*] needs clarity, there's nothing to indicate exactly what it is or isn't… we need to put more clarity around the definitions and what these numbers mean*".

This is clearly an important part of trust, and the decision to act based on the information in *NHSquicker*. Nonetheless, the real-time information is considered useful. The purpose of *NHSquicker* is to support patient behaviour (2.1.2), and all participants were confident that it has been a contributing factor in managing demand, with more patients attending MIUs and moving away from ED, for example:

> *"It has had an impact, has it had a huge impact… I think there are other factors that determine choice… We have seen… a significant shift toward the treatment centre in NA … and NHSquicker is part of it, it's not the only reason."*

Finally, all interviewed staff have used the real-time data themselves (2.1.3), as a shortcut to understanding performance in ED, and how it is comparing locally. The convenience of access to the data via a mobile phone app is cited, for example when off-duty or away from their desks. These indicate confidence in the real-time data for supporting decision-making and action, both for patients and for staff.

The trigger contained two important elements (Chapter 6, Section 6.4): the use of the *Total Crowding* data as a proxy for ED crowding (2.2.1), and the 24-hourly time-dependent trigger (2.2.2). Both were considered to be valid and useful, however contextual, which is important within a sociotechnical system. For example:

> *"I think the thing around number of patients in the department… that's a measure of crowding in ED. That's not necessarily around the level of activity, the flowing, in that hour, it's more likely to be lack of processing of patients building up into that hour… why couldn't we process those patients, was it outflows, was it inflows, was it just that we didn't have enough doctors on. So those things are really useful."*
>
> *"We know the thresholds around that, the trigger. At 40 patients it will instigate an escalation, but the problem with 40 is it creates crowding in the department and slows all our processes substantially."*

This validates the hourly trigger (which averages 39 patients across a 24-hour period, see Section 6.4.3 in Chapter 6), as their own trigger (Total Patients = 40) feeds into the escalation (OPEL) tool reactively.

The forecasts were discussed in terms of their usefulness in ED, and with regard to their current ED tools (2.3.1), which forecast daily arrivals and admissions:

> *"At the moment we have a forecast tool, how many will appear in ED and how many will be admitted, but it's not sensitive enough to really trigger a different response…it tells you the same thing the day before and the day before etc. - you are less likely to react to it."*

Interest in the 2-hourly *Total Patients* forecasts with prediction intervals was high, and the information was seen to add insight and value. At the hospital level (2.3.2), the predictions were considered to add potential value for patient flow, for example:

> *"..That would be very useful because all we are doing is forecasting what is physically arriving at the door and working from there, we aren't looking upstream… it would be particularly helpful for people in the control room... matching the potential for beds".*

The forecasts were seen as potentially contributing to adaptive planning, in particular flexing the workforce in ED. However, for patients (2.3.3) the predictions were seen to be problematic by managers, for example:

> *"…We want it so people make the decision to go the right place for them, because the wait will be less. One of the risks around [*predicted wait-times for patients*] is… 'I'll wait and go later when it gets quieter'."*

The IT participant noted the importance of this information being accurate, as patient satisfaction will reduce if their expectations are not met.

A significant focus by all participants when demonstrating and explaining the DES model was staffing. The difficulty encountered when developing the DES in accessing staff rotas were confirmed. Short-staffing is an issue, planning rotas is a priority action, and how to appropriately flex staff for short-term adaptive change was repeatedly raised. The challenges of staff flexing (2.4.1) was a significant theme, for example:

> *"it's really important - there's a point with overcrowding, where we need to put a second HCA in to triage the patients, so that to me is a critical point that you need to then be starting to get ahead, not to wait until its overheating but to get ahead of it."*
>
> *"Just getting one more doctor is not going to help if they're useless. It's not going to make any difference."*

Staff rotas (2.4.2) were raised, with a similar theme:

> *"The rotas are quite moveable. But anyway, that would be a simplistic way of doing it, a very linear way of doing it. You could get two very good, quick nurses, but we've all worked with people who are not very good, and we know how awful it is..."*

*"When we had that recent meltdown, somebody went back and looked at the consultant rotas, and over half of the ED consultants were off that day…So that level [*in the simulation] *would be really good, to use the complexity of the tool… you would then get people to use it."*

The difficulty of keeping rotas consistent due to staff shortages (2.4.3), is a problem. These issues were raised in particular when demonstrating the DES scenarios (2.4.4), where staffing/supply side issues were consistently of greater interest than demand management, for example where to put senior doctors at different times of the day, and optimising skill-mix throughout the day:

"The biggest problem for us is our demand side is actually less variable than our supply side. So if our supply side were more uniform, the small variations in demand that you get because of whatever is going on externally would be fine…If we could keep our supply more stable and consistent then we probably wouldn't have the problem".

The focus returned repeatedly to patient flow through the hospital system, and the impact of downstream blockages and staff shortages on ED crowding, while the effects on the urgent care network were of limited interest. Some concerns about who would use the HM were raised:

"But it would be – *who would use it? Perhaps planning. You could use it in real-time, but my question is would they? I think they'd use the forecasts and then plan, what do we think is coming in, the types, then what would they do to move the resources."*

This highlighted a broader issue, around where the model would sit and who would use and maintain it. Scenarios were considered useful:

"I think what's great about this, I can see the art of the possible. So if I were the COO, I would be thinking if I can move some of my demand around, in to different places, into different pools, which we sort of inherently know, but now I can physically see the impact."

Specifically, regarding scenario three, which was based on the availability of crowding data about MIUs in the urgent care network, staff were asked whether it was important to consider MIU crowding before redirecting patients there. Responses were brief:

"Yeah, my sense is that that is often forgotten"

"Yeah, I think there is a disconnect there"

Criteria 3: This focused on potential unintended effects for staff/ED, including how the information is presented, effects on workload, fit with workflow, model usability and functionality. Unintended effects at the system-level also need to be considered, including effects on patient attendance decisions, and potential effects of predictive data on attendance decisions. The main themes are summarised in Table 7.5.

Table 7-5 Criteria 3 themes and subthemes

3. Unintended effects	
3.1 For staff	3.1.1 Model complexity 3.1.2 Output information
3.2 For patients	3.2.1 Decision support
3.3 For the system	3.3.1 Providing patients with predicted wait times 3.3.2 Information/alert overload 3.3.3 Non-linear effects

The complexity of the model (3.1.1) was an important theme, as confidence in the outputs of a model is associated with understanding and trusting the model (Harper, Mustafee & Yearworth, 2020). This was definitely considered to be an important factor. For example while demonstrating the DES:

> *"Visually, how you could present that. It looks complicated, it is complicated. It would be, if you were trying to get people to use it in real-time, that would be a struggle."*
>
> *"They have to trust it. Trust is the most important thing because we all know we can do anything with data."*

This was also discussed in relation to the difficulties of modelling sociotechnical systems, including uncertainty around weather, unpredictable events, and human factors in decision-making. There was also interest in how the model could be improved, for example bringing in weather data. The presentation and usability of information (3.1.2) is similarly relevant, and related to the accuracy of the outputs and prediction intervals, which again was seen as contingent and not absolute, such that planning for a number of scenarios could take place based on a range of forecasts. Interestingly, effects on patient safety and other unexpected outcomes were largely not brought up, despite prompting. As

NHSquicker has now been available to the public for over two years, and no adverse incidents have been reported, all participants seemed assured that providing patients with a subset of hospital operational data is a satisfactory approach to supporting their health-seeking behaviours (3.2.1), and that if the department is crowded, the public should know about it. However participants were not in agreement with providing patients with predicted wait times (3.3.1), believing it to be more likely to support 'when to go' decisions, rather than 'where to go' decisions. While 'when to go' supports the shaping of demand across the urgent care system, participants were keen that those who could attend an MIU, did so, rather than shift the ED queues around, for example:

> *"For staff absolutely. For patients, I think the real-time is what is useful, because they make a decision in the now. And actually they might decide to come later rather than go to an MIU now, or I'll go earlier, and unintentionally join the queue,"*
>
> *"..if anyone is in a position where they can look at a map of information like that and think, well I'll tell you what, I'll pick my time to come because it's quieter then, for me that means that it's not an emergency."*

Information or 'alert' overload (3.3.2) was a significant theme. Linking with the usefulness of the HM for supporting escalation activities given OPEL status, and the forecasted daily arrivals (3-year moving averages) which changed little day-by-day, staff were seen to have stopped responding, such that alerts had lost their value. This is also linked with an issue of tension between managers and frontline workers, for example:

> *"It doesn't work if I'm honest. Because we spend all our time in OPEL 4, and you can walk along and you can see clinicians sitting in offices while we're on OPEL 4. I've been in OPEL 4 with 5-6 clinicians and they haven't moved, not moving."*

Several participants raised the possibility of non-linear effects, such that outcomes may be unexpected. For example, several provided an example of a non-intuitive outcome of providing additional ambulances over winter, which has unintentionally created an increased surge of arrivals earlier in the day.

Criteria 4: This criteria asks to what extent the HM might support all three levels of SA: perception, comprehension and projection. This is considered alongside existing decision-support tools. These are summarised in Table 7.6.

Table 7-6 Criteria 4 themes and subthemes

4. Situation Awareness	
4.1 Existing tools used to support SA	4.1.1 OPEL status 4.1.2 ED dashboard 4.1.3 Forecasts 4.1.4 Crowding tool

Participants were not asked specifically about SA, but how existing tools were used to support SA, or gaps in SA support were identified from the interview data. The main tool discussed was the OPEL status (4.1.1), which is an organisation-wide tool with a comprehensive set of triggers (including social care and community care) which can impact on patient flow, and progresses through a series of stages. While this provides perception, comprehension and projection information, a wide range of behavioural factors come into play that can impact on action. As previously discussed, it has become increasingly ignored when pressure has been high for prolonged periods ('alert overload'). The tool contains a mixture of quantitative data and 'guess work', and in general there was consensus that data could be better utilised for escalation.

The ED real-time dashboard (4.1.2) is used both in ED and in the control room, to show patient numbers and acuity in ED, and in which areas, and how long patients have been in the department. Similar themes arose here as in the previous discussions. However one of the biggest issues about this tool, which provides useful real-time information to support perception and comprehension, and from which experienced staff could make projections given the day and time of day, is its availability, for example:

> *"*[We] *haven't got tools that are available on the go… you have to always go to a board and look at something. Whether it's the… ambulances coming in, anything, you've got to go to a board and look."*
>
> *"That might be ok for the bed manager. But the clinicians will be sitting looking at their stack of people, so again, they go to a board, to look at it. It's not handy."*

While some didn't identify this as a problem, it clearly impacts on workflow.

Forecasts (4.1.3) are available of daily ED arrivals, admissions and discharges which are moving averages of the same day for the last three years. While accepted as unreliable, they are considered to provide information of value. A theme raised in general about data for decision-support is interpreting it in light of its context. For example:

> *“..The data only gives part of the picture, you do need to have that extra layer on top, the qualitative layer.”*
>
> *“Lots of things impact on activity, so there’s only so much that data can tell you… we had Storm Dennis, we knew it would hugely impact on people’s behaviour, we know with the rugby, if there’s an FA cup match…it would be fantastic if somebody someday could write us a model that could do that, but right now that’s down to our experience and our knowledge of how those days generally pan out.”*

A ‘crowding tool’ was introduced into ED at some point, which took a crude average of a range of measures. Anecdotally, the output was difficult to interpret and the tool fell into disuse. Some participants weren’t aware of it at all. This is an example of a tool which provided ambiguous information, which likely had a negative effect on SA, and ultimately was discarded.

Criteria 5: This looks at potential barriers to use at all levels of the organisation, e.g. time, capacity, politics, resistance to change, individual factors, technology-readiness, sociotechnical context, maintaining the model. Each of these factors proved to be of interest, in particular political barriers and individual factors. Time and capacity issues were only raised from the perspective of maintaining the model. From the data, an eighth factor was identified, which was system-driven behaviour. These are behaviours that are considered to emerge from structural components of the system. These are summarised in Table 7.7.

Table 7-7 Criteria 5 themes and subthemes

5. Barriers to use	
5.1 Individual factors	5.1.1 Staff quality 5.1.2 Individual behaviour
5.2 Internal politics	5.2.1 Clinical-management tension
5.3 Resistance to change	5.3.1 Tradition 5.3.2 Data access 5.3.3 Innovative data solutions
5.4 Sociotechnical context	5.4.1 Information access

	5.4.2 Data input quality
5.5 Technology readiness	5.5.1 Data-driven decision-making
5.6 System-driven behaviour	5.6.1 ‘Gaming’
5.7 Maintaining the model	5.7.1 Innovation sustainability

The ‘quality’ of staff (5.1.1) has been previously discussed, in relation both to those on the shift, and extra staff brought in to assist during escalation. However individual human factors (5.1.2) are also relevant, for example staff will reportedly slow their pace if they feel they are working harder than other staff members. One way to address this was discussed in terms of decision-analysis, for example:

> *“Will people act on what they find out? … they need to do something, and they need to do it earlier than they are currently doing it. But you need to work out – what are those things that the more experienced people are currently doing? Why are some doctors, some nurses, really good at what they’re doing to move people through the system, what is that they do? There’s a huge amount of work to do around that.”*

Clinician-manager tension (5.2.1) is the most significant internal political factor, and was observed by all participants. For example, from the clinical perspective:

> *“Who is going to make the change? …the most irritating thing is when you have someone who turns up with clipboards, telling us how busy it is, we know its busy, and then telling us to work harder.”*
>
> *“The CEO came along one day and said… ‘What can I do to help?’… ‘the best thing you can do is to resign, and we can use the money to employ some more nurses’.”*

Overcoming this is seen as a negotiation between individuals, with experience and a clinical background considered to help.

> *“We allow clinicians to be so wonderful and righteous… we can’t go on like this, because as much as we are really modern, we are still hugely traditional, hierarchical.”*
>
> *“Clinicians are just like ‘I’m not doing that’.”*
>
> *“So I just think, because there is so much data quality problems around NHS data, that’s always an issue with clinicians. If it says what they want it to say, fantastic, but if it doesn’t, they don’t trust it.”*

Resistance to change was not a significant theme, presumably because NHS Trusts have been in flux for some time. Some issues covered the theme of manager-clinician tension and traditional behaviours (5.3.1), often described as ‘how it has always been done’. Access to data (5.3.2) was seen as a significant barrier, with governance rules seen to stifle new ideas. This could present barriers to future work. Innovation (5.3.3) was similarly seen to be difficult, for example integrating NHS data with external data sources, or having new analytic tools accepted, which is likely to be a barrier to implementation of the HM.

The sociotechnical context is an important theme. Much of this has been previously discussed, for example how current information is presented and accessed (5.4.1) for supporting short-term decision-support. Data quality (5.4.2) is also important, and has been previously discussed. While data used in the model wasn’t seen to have quality issues, as reporting on it is mandatory, data in other parts of the hospital is problematic, for example patient discharge data is manually entered, and often not a priority for clinical staff. This is important to consider for expanding the HM, which currently has a ‘delay’ to leave ED (Chapter 6), but is not specific to bed capacity.

Technology readiness (5.5.1) was a significant theme. In general, participants expressed high interest in new approaches to data-driven decision-making. Responses indicated that there is still progress to be made, which is seen as taking a risk, however all participants saw the value in the HM, for example:

> *“Where we’re saying ‘we recognise we’ve got a problem, we need to shift our resources’, we need to understand the consequences. Because we are always moving things before we understand the consequence of it… this is absolutely why we should be using simulation.”*
>
> *“This* [the HM] *is going in the right direction. I’m very aware that we aren’t as data or information driven as we could be. Having data is one thing, having data which informs decision-making is something else.”*

System-driven behaviour (5.6.1) was a constant theme in the interview data. This was discussed in general terms, and is behaviour which is seen to arise as a result of interaction with structural elements of the organisation (Mullins, 2007). ‘Structures’ are defined as the order and systems put in place by managers to direct the efforts of the organisation into goal-seeking activities. Target-driven behaviour is well-recognised, and can affect data quality, for example, in ED patients can be removed from the data-system despite still being in the department to meet the 4-hour target. This means they are still contributing to crowding, but will not be reported in the data, if they are ready for discharge but delayed. The problem of entering ward discharge data into the system was also reported as system-driven behaviour. The data helps the bed managers to manage beds, but creates additional work for the nurses, both in entering the data and by adding workload when the bed is immediately filled. This is relevant for future work, which hopes to include real-time bed capacity as a significant contributing factor to ED crowding. An additional issue with transparency about available beds is 'gaming' in ED. Capturing and predicting these behavioural elements is difficult, and again is relevant for future implementations of IHAF, as well as affecting historic data, for example:

> *"I think the 4-hour target forces that* [admission to a hospital ward]*. It's harder work to send people home, and it takes more time, and you've got to get it done in four hours. If you have a whole stack of people here waiting, you've got to work hard to get them out… then we breach* [the target]*, so we get punished, so we admit."* (Doctor)
>
> *"Well if we could, what they* [ED] *would love to know is what the bed position is. But you also have human behaviour that happens in ED. If ED know there are beds, they will use beds. So there is a really strange way that doctors work".* (Manager)

Maintaining the HM (5.7.1) was considered by one participant only, who raised some interesting propositions which present significant practical barriers to implementation. For short-term decision-support, the HM is built for recurrent-use and needs to be embedded in the organisation. Even once a model is validated and integrated into the system, it still needs to be kept updated, structurally and in terms of the data:

> *"So for me, we should have somebody permanently… I don't think it's a one-off, you need someone who knows how to drive it* [the HM]*, but then it's also a case of, ok if we make that change, and it does make an improvement, awesome, so we now have a new baseline, so what else can we do."*

This requires understanding the system, which is as important as understanding the model. It was proposed that the information is fed to operations managers to make decisions, but it can be seen that communicating and enacting its output may not always be straightforward.

Outcome 6: Potential improvements: From the interview data, a number of potential improvements for the HM have been identified, some of which have been mentioned in the above sections. These are outlined in Table 7.8. These can be used to inform future iterations of this work, which requires a collaborative, co-creative approach with the health service involved in order to progress toward potential implementation.

Table 7-8 Criteria 6 themes and subthemes

6. Potential improvements	
6.1 System level	6.1.1 Flexibility 6.1.2 Urgent Care Network
6.2 Human factors	6.2.1 Decision analysis
6.3 Staffing	6.3.1 Planning 6.3.2 Flexing
6.4 Access	6.4.1 Flexibility of access to information
6.5 Testing components	6.51 Unexpected effects

A significant finding is that a model of ED alone, without specific consideration of downstream processes is of limited value, as flexibility (6.1.1) is an important aspect of usability. In this application, while crowding could be the result of unexpectedly high demand, it is clear that considering input and throughput, without specific consideration of output, is a limitation. While the HM simulation has a 'delay' component which represents bed delays (extending the length of stay for a specified proportion of patients, Chapter 6, Section 6.7.1), delays waiting for specialist reviews, transport delays, or delays for results, the bed delays are the most significant system-level problem and need to be incorporated. For example:

> *"It might be the demand. We might need… to tell the GPs not to send anyone in, or to send them somewhere else, just tell them."*

> “[If the HM could tell us] *of all the things that could be going wrong, what could make it better…if the problem is exit block, to focus on that, and not on sending people to the wards to discharge patients if there’s nowhere to put them…”*

This has wider implications for IHAF applications. As a recurrent-use tool, it needs to be flexible with a range of possible scenarios. It is also important to consider the boundaries of the model within these scenarios, so that the functionality of the model is not too restricted, and is able to adjust and evolve as problems shift.

The boundaries should consider both upstream and downstream elements (for example, with this application, as illustrated in the above quotes), but they may also need to consider wider implications. From this application it is apparent that while the hospital sees itself as a single unit, it doesn’t necessarily consider the needs or capacity-constraints of the wider urgent care network (6.1.2), yet they are aware that this wider network impacts on both themselves, and are impacted by themselves. Nonetheless, there is an awareness that the system as a whole could be used more efficiently, and there is still work to be done taking a higher level view of the urgent care network, for example:

> “[The real-time data is] *creating value from a wider point of view. So it would be value for patients but it will also be value for the system because it’s about allocating people into a more efficient place to be.”*

Human factors were a major theme, and in a sociotechnical system where the decision-makers are people, and the activities are performed on people, it is clear this area might benefit from further investigation toward ongoing iterations of the HM for short-term decision-support. This can, for example, investigate aspects of the usability of an embedded HM, including usability testing; or evaluation of the tool in situ to determine whether it is used as intended. Another example is how experienced staff make decisions under time pressure and uncertainty (6.2.1), and what can be learned from this. This type of research may be a prelude to new applications of IHAF as part of the problem-structuring stage, and may be important for considering the HM output, how it is presented and communicated, and maximising the value that can be delivered for decision-support. This type of analysis is common for clinical decision-making, but significantly less common for operational decision-making. For example:

> *"Why are some* [clinical staff]*…really good at what they're doing to move people through the system, what is it they do… let's talk about why you're doing that… 'what were you thinking when you made that movement'. And they'll tell you why, and then you can have a discussion about it."*

Staffing, as previously discussed, was a significant finding, with regard to planning rotas (6.3.1) and flexing staff (6.3.2). The relevance of this is related to the flexibility of the model to address new questions, and potentially to make planning decisions as well as short-term decision-support.

Many staff mentioned the current issues with accessing necessary data or information (6.4.1) to maintain an updated awareness of the current system state. Examples have been previously given.

A further theme, was raised but not explored, and this is the need to test individual components in situ (6.5.1). The potential for non-linear and unexpected effects was previously mentioned, and prior to the introduction of *NHSquicker*, there was significant concern about the potential effects of patients having access to subsets of hospital operational data. This has now been tested, and it appears to be having the desired effect on patient behaviour, with no undesirable consequences reported. However one unexpected benefit has been for staff, for example:

> "[NHSquicker is] *something that isn't aimed at clinicians, but actually they're using it, because they can on the go, and they can see what's happening with their system... it doesn't require them to have a formal phone call to see if the system is overheating…"*

7.2.5 *Summary of analysis*

Despite the cancelled interviews and small sample size, there was significant consensus of findings across participants, which supported analysis and conclusions, and can be used to inform future work.

By focusing the interviews from the start on a 'critical incident', the data converged on the issue of OPEL status and escalation, where ED is one part of the hospital system, and maintaining patient flow through the system is the goal. Although escalation actions implicitly consider patient safety, even participants who have no direct involvement with patients spoke of negative patient experiences during crowding situations. The need to avoid crowding is seen as a hospital-system problem, requiring a range of hospital-level solutions, where demand management is only one of many possible interventions. The focus on staffing suggested that this is considered to be a particular priority currently, and the HM needs to be flexible enough to account for a wide range of potential scenarios, some of which are downstream of ED. The input-throughput-output model (Asplin et al. 2003) is one way of conceptualising the range of factors which can contribute to ED crowding. While all participants were positive about the potential value of the HM, improvements are needed. For example, the model needs to be more flexible. Currently, the model is constrained by the real-time data available, however accounting for real-time acuity in the department can be an important future aim, alongside arrivals, admissions, and bed capacity. One consideration is the data quality issues highlighted on the wards at the use-case hospital, which are subject to internal politics and conflicting priorities.

Human factors were a significant theme from the interviews. A repeated issue was that of reactive behaviour, both in terms of adaptive, task-level behaviours and system-level escalation behaviours. Environmental information supports both perception and comprehension of current conditions. It appears that the simple forecasting methods in current use provide adequate environmental information to support a mental projection of the system state into the near future. Nonetheless, proactive adaptive behaviours in the ED were reported to be unsatisfactory, for a range of individual and group-related factors. The tension between managers, who drive goal-directed behaviour and maintain performance, and clinical staff, who provide care, is well-recognised worldwide (e.g. Ranawat et al., 2009), and has been a pressing consideration in the NHS for some time. Managers and doctors are distinct groups of people, who share some common goals, but also often have different ways of working, different incentives, diverging objectives, and different tribal loyalties. In the UK, these tensions have been becoming starker in the face of ongoing financial constraints and complex organisational challenges (Davies, 2015). Unprompted, this issue was raised by all participants in the interviews, and clearly is a significant challenge, causing frustration and suspicion between these two key groups. While the simulation can potentially provide specific, best-case solutions, internal politics might result in a failure of action. This could even extend to a failure of acceptance of the HM in the first place, as autonomy of decision-making could be seen to be undermined. This is one reason for testing in situ the components of the model separately, as was done with *NHSquicker*, prior to integration.

A key focus of all participants is how to maximise the value of available resources in planning rotas, for short-term escalation actions, and for flexing staff and skill-mix, so the model needs to be validated for this purpose. Nonetheless, significant barriers exist when considering its future implementation. Firstly, its mode of implementation, as the majority of participants indicated that a mobile device would significantly increase the usability of the outputs, in particular for the forecasts, which provide decision-support even without the DES component. Secondly, the question over who would 'own' the HM, as resistance is possible if the implementation is seen to be a management initiative. Thirdly, and significantly, early consideration would need to be given to who would manage and maintain the model, and it was suggested by one participant that this would require a dedicated internal staff member. These barriers are over and above the usual barriers of engaging with staff collaboratively to validate each aspect of the model in practice, develop scenarios, and communicate results. However, there was found to be an openness to using enhanced data applications for supporting decision-making, and maximising the value that can be obtained from the hospital operational data to improve service provision. The use of real-time data, forecasts, and even simulation are not new in healthcare, and were recognised by all participants as valuable. However as a researcher, developing and finally implementing and embedding a real-time decision-support tool that is useful, usable and tested in practice, is a substantial challenge.

The next section integrates the analysis of the interviews in light of the patient questionnaire from Chapter 5, to address Research Question 3. This asks how real-time data can add value at the level of the urgent care system, and considers what IHAF can learn from this application.

7.3 System-level value and challenges for real-time data applications

The aim of Research Question 3 is to analyse the system level impact of the use of the real-time decision-support tool, both for patients and staff, and to synthesise this with the existing literature. The value proposition of the HM on the supply-side (staff) and demand-side (patients) is addressed, its efficacy and safety, and evidence of benefit to patients and to the system as a whole.

Involving patients in healthcare OR research is not common, but aims to start with an understanding of what is important to patients, to ensure that modelling efforts focus on measures that patients view as important as well as improving an in-depth understanding of the problem situation (Pearson et al., 2013). A summary of the data used in this thesis is illustrated in Figure 7.3. The 'users' are patients and the public, and have access to real-time data (descriptive analytics) and proposed access to wait-time forecasts (predictive analytics), both with the aim of supporting demand management. This has been evaluated formatively using patient questionnaires (Chapter 5). From the NHS side, it is proposed that staff will have access to the forecasts and the simulation model, the prescriptive component. This was evaluated using semi-structured interviews (Section 7.2).

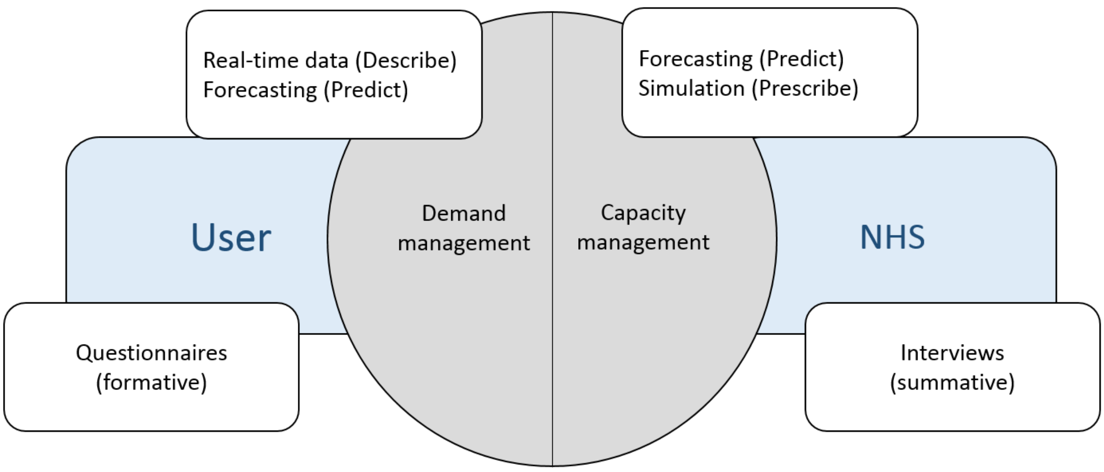


Figure 7-3 System-level data used for analysis

By involving end-users, who are part of the system under investigation, an understanding of the current requirements and perceived value for patients can be considered, as well as ensuring that potential unintended consequences of interventions are considered early. This wider view should also consider other services within an NHS healthcare network. When all services are under strain, the effects of actions on other services (e.g. GPs, MIUs) is often not considered. Five system-level themes have been identified from the formative (questionnaires) and summative (interviews) evaluations. These are: (i) Building resilience across the urgent care network; (ii) Managing manager-clinician tension toward HM implementation; (iii) Patient health-seeking behaviour and its impact on demand; (iv) Situation awareness and decision-making; and (v) Model sustainability over time. The next sections will consider these in more detail, and indicate what IHAF has learned from this use-case application.

7.3.1 *The urgent care network*

A particular challenge for acute hospital-based care is maintaining patient flow through the system. The ED open-door policy within the NHS system, which is free at the point of delivery, enables access for those who might otherwise have considered self-treatment or chosen alternative facilities. When hospitals are working near to full capacity, high attendance rates reduce patient flow and lead to ED crowding, as admission or discharge is unable to keep pace with new arrivals. The resultant queuing puts pressure on staff and resources, and impacts negatively on patient experience and safety. Current policies of closing ED to low-acuity patients as an escalation action when the hospital system is under pressure takes little account of the amount of pressure the MIUs are under, as patients are reactively redirected there. The interviews, which were focused on periods where ED is under pressure, found little evidence of accounting for pressure in other services. In fact the opposite was found to be indicatively true, as GPs are alerted by the hospital to refrain from sending patients in, and MIUs are expected to adapt to the sudden influx of low-acuity patients. When decisions are made in real-time to reduce operational pressure in a hospital, policies of demand management will necessarily be implemented that will inevitably have negative consequences on other parts of the wider system.

Patients were able to see two main sources of value in real-time analytics – firstly for optimising resources including staff. Planning rotas and flexing staff was also a major focus by interview participants. Unprompted, a large number of patients indicated that a second use for real-time analytics using *NHSquicker* data is to balance demand and capacity across the system. Patients indicated that EDs, GPs, pharmacies and MIUs might effectively direct patients to quieter services and share information between services. They saw the value in both staff utilising system-level information to shape demand, and in supporting the public to make more appropriate attendance decisions. The patient participants identified the need to take a system-wide view. They consider re-direction decisions to be acceptable if there is a clear advantage in it for both themselves, and for the care providers. However being moved from one queue to another queue, at the directive of the governing organisation is unlikely to be a satisfactory outcome for patients. Yet from the interviews, there is little evidence that a whole system approach is currently being utilised to tackle short-term demand management in urgent and emergency care. The Royal College of Emergency Medicine (2015) has urged urgent care systems to work together to improve efficiency and deliver more equitable and appropriate care. From the interviews *NHSquicker* is currently providing the most convenient and immediate source of information about the state of the urgent care system in real-time. The HM developed in Chapter 6 extends the value of this information by using predictions to provide a short-term window to safely enact patient redirections, while accounting for the current state of the MIUs in the system. Staff have previously specified an interest in working at this wider system-level. Working together at a regional level helps build system resilience, defined as the ability to *anticipate*, to *react* and to *mobilise resources* for rebuilding and recovering after a degraded or critical state (Hollnagel, 2009, 2011b). This involves spreading the risk across an entire urgent and emergency care system (Higginson & Boyle, 2018). While more work needs to be done, effective crowding management requires a whole system view to address the balance of workload across all sectors of the urgent care network. It also requires collaborative leadership, as discussed in Section 7.3.3.

Key implications for IHAF

As IHAF supports the development of a HM which is for recurrent use, scenario flexibility is important as priorities shift with time. It can be difficult to articulate future decision requirements, however this maximises the value and utility of the HM and its real-time data, and its sustainability as a decision-aid. One element of simulation model flexibility is its boundaries, and while the current DES can be improved in terms of capturing admission delays, it does look more widely at the urgent care network. One of the goals of IHAF is to support system resilience, and an important part of this is spreading risk. As part of the problem definition phase, the boundaries of the model and the wider implications of scenario interventions should be considered from the perspective of resilience.

7.3.2 *Patient decision-making*

IHAF explicitly suggests considering all relevant stakeholder groups in the problem definition stage (Chapter 4, Section 4.4). A complex interaction of physical, psychological, social, and demographic factors influence a patient's choice in healthcare utilisation. The literature review in Chapter 5 categorised these factors, and the questionnaires investigated how real-time knowledge of wait-times might influence attendance decisions. While the majority of patients overall considered themselves to be in the most appropriate place, significantly more patients who would have found the real-time data useful indicated that alternative facilities such as MIU or GP could have been appropriate for their visit. There is a general consensus that people are using ED services inappropriately, however there is a significant amount of debate about how to define ‘inappropriate’ and the proportion of inappropriate attenders.

The boundaries between urgent care provision, including general practice, MIUs, urgent treatment centres, walk-in centres, ED, pharmacies and other service providers are confusing and unclear, with health-seeking behaviours informed by a number of factors, as identified in Chapter 5. Nonetheless, the questionnaire found that 79% of patients reported that they were ‘certain’ or ‘very certain’ that they were in the most appropriate place for their care. Pope et al. (2019) also found a ‘moral positioning’ where the health-seeking behaviours of others are judged, while patients’ own are viewed as legitimate. Attempts are underway to clarify the confusion over fragmented healthcare service provision for urgent and emergency care to support the public in making attendance decisions (NHSEngland, 2019), however Pope et al. (2019) argued that patients are not deliberately making ‘wrong’ attendance choices but that their choices are socially constructed and informed by past experience and beliefs. Other research findings agree (e.g. McGuigan & Watson, 2010; Beache & Guell, 2015; Sancton et al., 2018). Within the wider urgent care network, it may be relevant to take the focus away from whether decisions are ‘appropriate’ or not. The need for urgent and emergency care is contingent and subject to multiple definitions (Durand et al. 2011), and can be determined by service providers, by users, or both. Quan et al. (2013) found that professional assessment of urgency was based around timeframe and contextual subjectivity, such as whether the patient or their family were distressed, rather than clinical features alone. They also reported that definitions of urgency varied between physicians and nurses, with nurses more likely to take in the wider context of the patient experience.

Despite patient certainty about ED as their attendance choice, it seems the real-time wait-time information can form a contributory factor toward ED attendance decisions. When framed according to Andersen's Behavioural Model of Health Service Use (Andersen et al. 2014), it is clear that there is the potential for the real-time data analytics to impact on individual health-seeking characteristics, in two distinct ways. Firstly, estimates of wait-times can be classified as a predisposing characteristic representing a patient's beliefs, attitudes or knowledge about health services wait-times. The provision of *actual wait-times* will potentially change this to an individual enabling resource, as knowledge of wait-times then enables a more informed choice. Changing a *belief* to an *enabler* is important if the health behaviour being enabled can be considered best for both the patient and the NHS. This supports the urgent care network by assisting the most appropriate distribution of patients. However, an additional, unexpected benefit of the real-time data analytics for patients who have been referred or who consider their condition to be serious is using it for planning, managing expectations and reducing anxiety, hence for this group of patients, the real-time information is bypassing the 'Health Behaviour' Component of Andersen's model and influencing 'Outcomes' directly, in particular patient satisfaction.

Additionally, information from patients about the potential benefits to the system indicate that they support levelling demand across the system, both through their own (and others) attendance behaviour, and through the NHS using the information to manage demand and resources. In this case, the real-time analytics has the potential to influence contextual enabling resources. These are conditions that facilitate or impede use of services, such as the number and distribution of services, staffing, and structure in the community, resources, opening hours, and facilities. While not directly influencing any of these factors, where demand is managed across an urgent care network by both patient behaviour and NHS processes, a reduction in crowding, and subsequently wait times in ED, acts as a contextual enabling condition for those individuals whose attendances are appropriate and necessary. However managing NHS processes relies on staff behaviours, as discussed in the following section.

Key implications for IHAF

The analysis of patient health-seeking behaviour and real-time descriptive data emphasises the need both to evaluate components separately, and to specifically look for unexpected outcomes, which can be negative or positive. This is an important element of the 'evaluation' component, and the questionnaire research has emphasised that outcomes can be different to those intended in complex sociotechnical systems.

7.3.3 *Manager-clinician tension*

As previously acknowledged, crowding can have wide-reaching impacts, and is associated with poor clinical and operational outcomes, and perceived quality of service. While the potential negative effects on patients were discussed in the interviews, including recent anecdotal examples, the ongoing issue of tension between managers and clinicians was raised repeatedly by all participants, and frustration on both sides was evident. This issue particularly arose in discussions about initiating both adaptive behaviours and escalation actions based on current tools, and it is evident that it stems from differing priorities and beliefs. Persistent tensions between cost, quality, and access are irreconcilable, and the decisions and trade-offs that attempt to resolve these issues are central to all NHS Trusts (Davies, 2015). From the interviews, it seems probable that staff who could act sooner to prevent a crowding situation escalating, sometimes choose not to do so, potentially compromising the care of patients. This is a problem felt across the urgent care network. For example, nurses in MIUs working autonomously sometimes have to manage acutely unwell patients, who self-present due to ED being under pressure (Bowen, 2019). Additionally, while GP services are not addressed in this work, indications from the interviews are that workload from GPs is considered to be a source of demand which needs to be attenuated. Yet they are a significant element of the urgent care network, and the impact of a chronic decline in the GP workforce is being felt in ED (NHS England 2020a). The interface between primary and secondary care clinical activity recognises the same behaviour of 'resisting' work as that recognised between managers and clinicians, leading to loss of goodwill and a sense that professional responsibilities are not being fulfilled (Sampson et al., 2016). In hospitals, the use of clinical managers is an established approach to increasing engagement of clinicians;

unfortunately due to COVID-19 a planned interview with a Clinical Director was cancelled on the day, which may have provided additional insight into the complexity of the situation in this specific context. Nonetheless, empirical evidence continues to show frustrations and tension between doctors and managers and a lack of optimism that relationships will improve in the future (Powell & Davies, 2016). It is generally accepted that although there have been some shifts in the power balance between doctors and managers, overall doctors continue to be a powerful group who retain considerable autonomy, which they seek to uphold, despite management initiatives. The comprehensive longitudinal study by Powell and Davies (2016) found three forms of subtle resistance by doctors: eroding aspects of the managerial system (e.g. by not using guidelines or protocols); co-opting managerial tools into professional work and adapting them in ways that maintain clinical autonomy; and critiquing managerial initiatives (e.g. by arguing that available data are flawed). For this case study, it is evident that these behaviours could potentially impact on HM adoption at multiple levels: (i) testing the HM components in practice; (ii) implementing the HM into existing workflow; (iii) responding to the outputs of the model, both at the task-level (adaptive behaviour) and at the system-level (escalation actions).

While gaining a contextual understanding of management culture in the NHS is beyond the scope of most OR studies, addressing this specific challenge is likely to be key to the acceptance or otherwise of a decision-support tool, even in the early stages of testing in situ. Stakeholder theory has been applied within the field of OR for understanding how to best identify and manage those stakeholders who are important for ensuring outcomes. For example, Ackermann and Eden (2011) found that the challenge of managing stakeholders becomes clearer when their interests are separated from their power to influence outcomes. Additionally, while formal relationships are well-understood, stakeholders' informal networks are invariably more complex, such that some stakeholders are more or less powerful than initially anticipated. Problem-structuring research has had at its forefront an interest in supporting decision-makers who are engaged with complex problems. One of the most significant of these types of problems is collaborative working. This requires participatory processes, resulting in an increased qualitative understanding of the problem, and lowering the risk of misrepresenting the goals and values of stakeholders.

Functional barriers across many different groups of professions, each with different knowledge, skills and perspectives, can be a barrier to adoption and integration of a new technology (Greenhalgh et al., 2005). From the interviews it is clear that there are individual as well as group processes to consider, as senior clinicians are key decision-makers, and alongside the political nature of the processes of adopting a new technology, this individualistic factor requires provision. In order to realise the potential benefits to patient care of technological innovations it is important to see 'adoption' as a process rather than as a discrete event, which comprises both 'formal' organisational decisions and a series of 'informal' decisions by individual users and teams, which ultimately leads to the integration, or not, of the technology into routine practice (Robert et al., 2010). While a participatory approach is more likely to support acceptance and implementation, de Gooyert et al. (2017) reported a lack of attention toward implementation and results in OR studies. However the effects of crowding are seen in both patient outcomes and staff morale (Morley et al. 2018). Post-COVID, renewed financial constraints and complex organisational challenges will potentially increase pressure and reduce staff morale in NHS hospitals, the tension between clinicians and managers is likely to intensify (Davies, 2015). The implications of this are that as crowding becomes an increasing issue, the commitment of clinical staff to action management solutions may reduce, putting patients at potential risk of harm. Early in this study, observational data found that clinicians demonstrated concern about the safety of individual patient decision-making and non-optimal attendances, while managers focused on understanding patient behaviours and the factors that drive decision-making. Evidence shows that doctors are more likely to consider the fairness of interventions related to individual patients while managers are more likely to consider populations of patients (Fitzgerald et al., 2006; Powell & Davies, 2016). For example, regarding the quality and safety of health care (Degeling et al., 2006; Klaber et al., 2012), managers are more likely to favour changes that move services towards a more systematised approach, that increase team-working and that balance clinical autonomy with greater transparency and accountability. Engagement with real-time decision-support is a priority, as it provides an overview of the system state, enhancing SA, and the ability to enact both adaptive behaviours and escalation actions. However in this application, without the support of both clinical and management staff, it is likely to fail early. The next section considers the role of real-time tools in enhancing SA and subsequent decision-making.

Key implications for IHAF

The challenges of collaborative working in an M&S study are heightened in a real-time study, when engaging in test-develop cycles toward implementation and recurrent use. Stakeholder groups may need to be managed throughout development, and a further consideration is how different groups respond to the outputs. Additionally, engaging staff in the implementation of a real-time decision-support tool can strengthen the link between the process of using data analytical methods for decision-support, and outcomes. This supports transfer of learning and makes it easier to demonstrate the relevance of the OR methods in specific domains.

7.3.4 *Situation awareness*

Positioning IHAF in Situation Awareness (SA) theory (Chapter 4) deliberately focuses attention on what information is needed where, and when, to support a continuous understanding of the current system state in a complex, dynamic system (Endsley & Garland, 2000). A key feature of the framework is the purpose of real-time decision-support tools in enhancing SA, an important constituent in decision-making processes (Chapter 4, Section 4.4.1). The choice to act on the information belongs to the decision-maker, and formed part of the discussion in Section 7.3.2, and later in Section 7.3.5.

For this reason, the design and use of the HM required specific consideration with regard to its contribution to SA. From the interview data, the output of the predictions provides acceptable information, however the complexity of the simulation model was challenging for some participants. The use of AnyLogic presents a further barrier, as it may not be usable on an NHS machine, and doesn't support integration. It is likely to be necessary to code the model using open source software, which also supports research collaboration, and model review. Staff who will not be interacting with the model itself (five participants) reported that they were satisfied with outputs that provide indicators for action, in particular clinicians articulated a lack of interest in the workings of the model. Additionally, aspects of the usability were highlighted, including the convenience of having the data outputs available 'on the go' in a mobile device, which were considered more useful than shared workstations or displays. These are examples of individual and environmental factors which influence SA, decision-making and action, and subsequent performance. Environmental factors are relevant for the design element of the HM, as unanticipated effects can result from the type or presentation of information, for example, technology-induced errors (McGeorge et al., 2015). Similarly, IT systems that provide ambiguous information or with poor usability can actually reduce human decision quality and speed (Endsley, 2016).

One interesting issue which was raised in the interviews was decision analysis, the study of how experienced doctors and nurses *actually* make decisions that move people through the system (rather than how they ought to make decisions in approximation to a rational standard). One way of investigating this is the study of Naturalistic Decision-Making. Naturalistic Decision-Making is concerned with how people, particularly experts, make decisions in complex, real-world, uncertain contexts that can require real-time decisions in urgent situations with significant implications for errors (Zsambok & Klein, 2014). Beach (1997), in the context of organisational decision-making, stated that values and beliefs, specific organisational and individual goals, and operational plans for reaching the goals, will guide and limit decision–making. This merges goal-orientated individual behaviour with the decisions and goals of other organisational stakeholders. The fit with the problem of management-clinician tension is clear. Organisational decision-making is often challenged by shifting or competing goals and uncertain, dynamic environments. Other factors relevant to ED include ambiguity or incompleteness of information, a longitudinal context, incentives, repeated decisions and conflict (Gore et al., 2006). It has become increasingly accepted that in order to build information systems that can support complex decision-making it will be necessary to more fully understand human decision-making processes (Zsambok & Klein, 2014).

In clinical settings, studies have gained an understanding of how clinicians make decisions in dynamic environments amidst interruptions, distractions, and uncertainty (Falzer, 2018), however little work has investigated how staff make dynamic operational decisions under the same circumstances. Using the Critical Decision Method (CDM) as a form of retrospective interview of decision-making processes during a critical event is one method of approaching this. In the evaluation phase of this application, the interviews were focused on a critical event, as used in the CDM, however decision processes, key points, and the main information sources were not explored in this context. Investigating SA to improve operational decision-making and decision-support applications in healthcare is an area that is likely to benefit from significant further work as part of the process of development of real-time decision-support tools. The final section looks at ownership, maintenance and sustainability of the HM once embedded in operational processes. This is important to consider early.

Key implications for IHAF

SA is an explicit element of IHAF, focusing the purpose of the real-time HM. The application in an NHS ED evaluated SA in general terms, which may be appropriate for this early stage of development, however consideration should be given to examining SA in the problem definition phase (for example using Naturalistic Decision-Making analysis), and the use of a specific measurement of SA may be valuable in the evaluation stage.

7.3.5 *Model ownership, sustainability and long-term evaluation*

Interview data in the use-case highlighted the importance of considering early how to embed the HM into a complex sociotechnical system, and who would maintain and interact with the model. However, significant work is needed to build a shared vision, identify and engage the right staff, and monitor the impact, which is difficult to capture in the long-term. Stewart and Williams (2005) emphasised that this work is often underestimated, and can be hidden and extensive. The co-evolution of the technology presents challenges, for example having the right staff to support adaptation and flexibility of the model over time, in a rapidly changing policy context, with technological evolution and organisational responses to these factors. Data from the interviews and questionnaires has shown that adopting and embedding new technologies into a dynamic sociotechnical system requires understanding and navigating its multiple interacting facets, and these cannot be overlooked or ignored. Greenhalgh and Abimbola (2019) summarised strategies for accepting complexity when engaged in healthcare technology programmes. These include strengthening programme leadership; maintaining a clear, co-developed vision; identifying and talking about uncertainty; supporting adaptivity and flexibility; and accepting that unintended consequences will occur.

Poor uptake of technology is often explained in terms of barriers and facilitators, however Greenhalgh et al. (2017) reasoned that individual factors do not make or break the implementation of technology in healthcare, but the dynamic interaction between them. This includes all of the factors discussed above in Sections 7.3.1 – 7.3.4. The more complex the domain of implementation, the less likely the innovation will be adopted. Greenhalgh et al. (2017) developed and applied a framework for predicting and evaluating the success of technology-supported health and social care interventions. Causes of non-adoption include a failure to acknowledge the complexity of issues, i.e. that issues are unpredictable, emergent and dynamic; insufficient prototyping, testing, and awareness of human factors issues; low technology maturity; and lack of consideration to sustainability of the technology. IHAF supports the consideration of all of these factors. Its value lies in the insights gained through iterative build-evaluate activities. This enhances understanding of the problem situation and the interactions of its subcomponents. The flexibility to change and evolve the model until it is useful and effective toward addressing the problem can simultaneously adapt and align with organisation capacity and readiness to innovate. In healthcare, as in other sociotechnical systems, it is important to develop an approach to modelling and analysis that abstracts away from the specifics of particular algorithms and obtains systems-level understanding. For example, the problem of 'alert overload' was raised in the interviews. Future iterations at the use-case hospital will need to consider to how to manage this complex human-technology issue as an example of data analytics components that interact with each other and with people (de Weck et al., 2011). A further issue raised in the interviews is that of trusting the outputs of the HM to act on them, and as the approach develops it should provide insight into how improvements in accuracy translate into gains that matter in terms of reduced costs, lives saved, time conserved, effort reduced, and quality of care increased (Wagstaff, 2012). Although data analytics is increasingly a key part of sociotechnical systems, the academic literature does not typically focus on the system-level impact of data analytics. Consequently typical measures of performance that are optimised and reported do not always align with domain experts' assessment of performance (Wagstaff, 2012; Rudin and Wagstaff, 2014). In the use-case, these included patients and staff. When data analytics are used in real-world applications, success is often not due to small differences in performance between models or algorithms, but by how well the solution fits the unique aspects of the domain and its evaluation measures, and these issues should be considered early and often.

7.4 Chapter Summary

This chapter addressed the third aim of the second research question, to demonstrate and evaluate the HM application in its context (Section 7.2). It also addressed the third research question, to analyse the system-level impact of real-time data applications by both patients and staff to determine the implications, barriers, and added value to the system (Section 7.3). Table 7.9 reiterates Research Questions 2 and 3, with the areas addressed in this chapter in bold.

Table 7-9 Research Questions 2 and 3

2. *How can an integrated hybrid approach using real-time simulation and predictive analytics support short-term operational decision-making?*	To test and evaluate the potential of an integrated hybrid approach for short-term decision-support in healthcare combining real-time simulation with other analytics approaches.	**1.** To propose a generic framework supporting an integrated hybrid approach for short-term decision-making in healthcare. **2.** To apply the framework within the case study in a hospital ED. **3. To evaluate the application of the framework.**
3. What are the implications and the added value to the system of using real-time data applications for both patients and for NHS decision-support?	To analyse the system level impact of the use of real-time data for both patient and staff decision-support.	**1. To critically evaluate the perceptions that patients and NHS staff have regarding the value that real-time applications provide at the system level.** **2.** To synthesise previous findings and evaluate the framework in light of the application.

The value proposition of the HM developed in Chapter 6 for the use-case is at the system level, aiming to improve efficiency, deliver more equitable and appropriate care, and support system resilience. The technology, its usability, aesthetics, dependability and accuracy, and the extent to which the information generated is accepted, trusted and considered sufficient for decision-support are all relevant. The information may empower and inform, but it may also be misinterpreted and cause unintended or indirect effects.

The real-time components have been evaluated through patient questionnaires and staff interviews, and indicate that descriptive analytics appropriately supports patient attendance behaviour, however predictive information may facilitate ‘when to go’ decisions, rather than ‘where to go’. Patients support demand management actions at the urgent care network level, however from the staff interviews, staff are consistently more focused on improving patient flow through their own system, and their interest in the wider system is mostly its impact on their own demand. While it is clear that a simulation model of ED must incorporate downstream hospital processes, the wider network is also an important consideration when managing patient demand.

Additionally, despite the focus on hospital system activity, hospital–level challenges exist. Stakeholder engagement and management is essential for most M&S studies in a sociotechnical system, but specific challenges were identified from staff interviews. The conflicting goals and behaviours of managers and doctors are likely to be a significant challenge, as the predictive and prescriptive components will require both clinical and management support. There is often no single adoption decision, and inter-professional relationships, power and politics are important. Additionally, while the output might support SA, whether it changes behaviour is uncertain, for example it might be seen to reduce clinicians’ professional autonomy in support of higher-level management decisions. Future research might use stakeholder theory for managing this situation toward the integration of a real-time decision-support tool into practice. Additionally, there is a need for decision-analytic research which works to understand how experienced or skilled staff make operational decisions that support patient flow. This can be used to inform real-time adaptive behaviours, escalation decisions, and the information or tools needed to support these behaviours. Finally, a further hurdle is sustaining the tool in practice with consideration of the political, policy, regulatory and legal contexts, the sociocultural environment, and how the organisation adapts to these rapidly changing contexts, alongside technology evolution over time. From the interviews, it is clear that these factors, alongside the short-term ownership of the model, how it is used, maintained and updated day-to-day, require early consideration and are not insignificant concerns.

The next chapter revisits the framework – IHAF - developed in Chapter 4, and applied in a use-case in an NHS Trust ED in Chapters 5, 6 and 7, for revision.

Chapter 8: Revisiting the IHAF framework

8.1 Evaluation of the Integrated Hybrid Analytics Framework (IHAF)

Chapter 4 proposed a generic framework for the development and testing of a hybrid model (HM) for real-time decision-support in sociotechnical systems. The previous chapter completed the application of the Integrated Hybrid Analytics Framework (IHAF) in a case study at an Emergency Department (ED), through its *evaluation* component. This chapter will revisit IHAF in light of its application in Chapters 5, 6 and 7. IHAF was motivated by the increasing need and opportunities to use real-time data to support quick and effective decision-making (Bumblauskas et al., 2017). Concepts derived from the Human Factors literature take account of sociotechnical system precursors of decision-making, including individual and team-level situation awareness (SA), and Quality Improvement (QI) theory was suggested as a means to bring together, in a generic framework, the concepts from data analytics, simulation and sociotechnical theory toward supporting short-term decision-making. The framework proposed in Chapter 4 is in Figure 8.1.

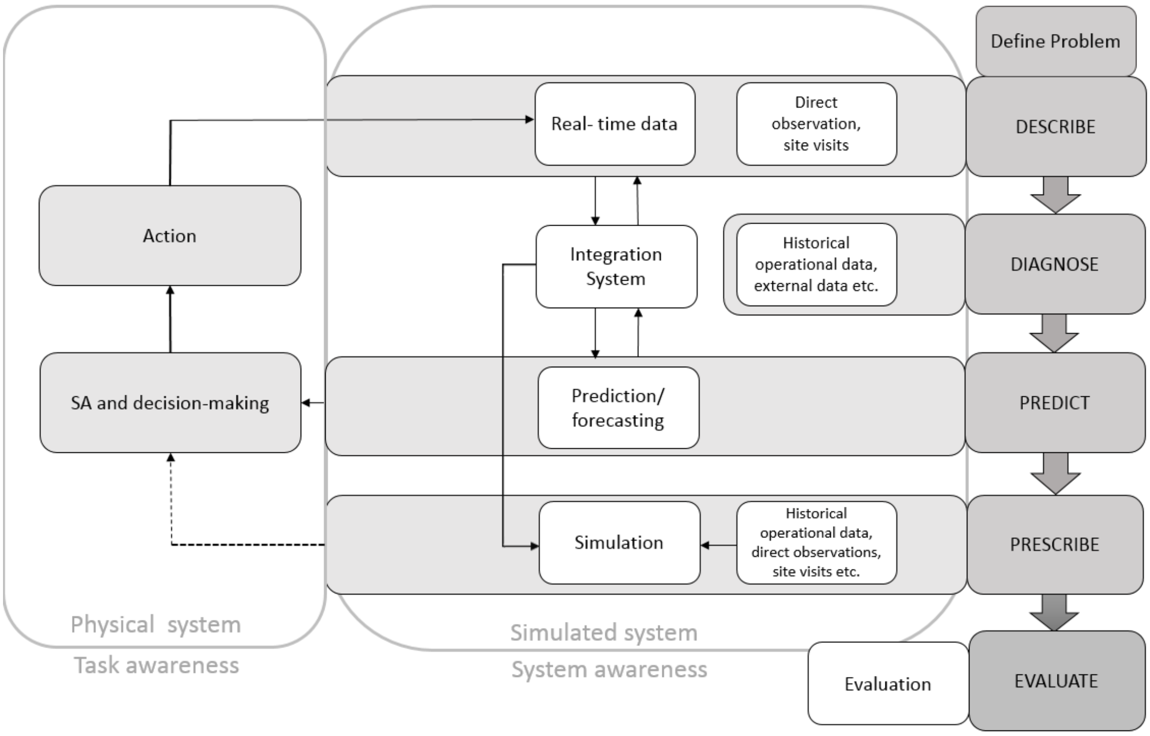


Figure 8-1 Integrated Hybrid Analytics Framework (IHAF)

The development of the framework was done in two ways. The first development was through an examination of the stages of a Design Science methodology (Blessing and Charkrabarti, 2009). The second was derived from insights from the literature review in Chapter 2, examining decision-making in dynamic, sociotechnical systems; data analytics and simulation for decision-support; and real-time simulation as a HM. The framework was developed to be generic, and tested in practice, with transferable knowledge aimed at supporting similar future work. The framework was evaluated with a use-case at an NHS ED using real-time data made available from *NHSquicker*. The following sections will evaluate each component of IHAF as structured in Chapter 4, starting with the use of the Design Science methodology.

8.1.1 *Revisiting the Design Science methodology*

Design Science is used extensively in computer science and Information Systems research, but rarely in OR. O'Keefe (2014) argued that Design Science is one way of achieving design-oriented OR, taking OR values and approaches back toward early OR practice. This means that OR concerns itself with the larger system—its context, its data and where it is placed within (and beyond) the organisation. Basing the IHAF framework in Design Science methodology arguably makes the research more relevant in its applied setting, while providing methodological rigour. A key feature of Design Science is the evaluation component, which leads to further design, demonstration, and evaluation. While a formative evaluation may be incorporated in the problem definition phase, the final evaluation informs future work. This means there is a close relationship between problem definition and evaluation.

The evaluation phase has another important distinction: it is not based on the value of the underlying method or algorithm, but upon the utility or usefulness of the artefact in practice (Hevner et al., 2004). A technically more 'correct' model may not have improved utility if it is not demonstrable in terms of gains that matter to stakeholders. OR as Design Science supports work that crosses functional boundaries, and O'Keefe (2014) argued that hybrid models are likely to be required, and that system integration should be the aim. This means that model development is a collaborative design problem, and its usefulness and usability are considered from the beginning. As discussed in Chapter 3, Design Science supports mixing methods and provides clarity about the added benefits to be gained from this approach (Gregor & Jones, 2007; Ågerfalk, 2013), which include enabling different stages of the project, dealing with complexity, allowing iteration and flexibility, and considering the wider system. Combining the development of a contextual quantitative tool, and a qualitative case study of users' interaction in the applied setting provides insight beyond the methods used in isolation (Ågerfalk, 2017). The model requires a staged, evolutionary implementation, which may not be necessary for single-use models, but for embedded real-time decision-support tools, the potential value is clear. This approach operates at the intersection of knowledge about physical artefacts, and human behaviour. For example, Design Science addresses the limitations and issues that arise during data collection, modelling processes and users' concerns (Blessing & Chakrabarti, 2009). Additionally, the methodology supports a critical approach (Hodgkinson & Starkey, 2012), such that the study process and application of the HM is contingent and situated (Zachariadis et al., 2013). It contributes to knowledge in a cumulative way, by supporting rigor in design research and extends its external validity through generalisability of results to other contexts that exhibit similar characteristics (Offermann et al., 2011).

One important criteria for Design Science is that it requires a careful definition of the artefact, and IHAF provides this. It is applicable to a particular set of problems where real-time decision-making is required in a sociotechnical context. One limitation is that 'implementation' is not an explicit component, where development, testing, evaluating and communicating results forms the process steps of the study, and the limits of the researcher role. The process of implementation may be a separate activity, which could be supported, for example, by Implementation Science, researchers-in-residence, or an internal team within the organisation. However Hodgkinson & Starkey (2011) offer a caution to remain sensitive to the danger of distortion by practitioners and policy-makers in the search for evidence-based management in the implementation stages. Nonetheless, Design Science provides a suitable foundational methodology for the IHAF framework. It can operate at the interfaces of academia-practice, and rigour-relevance, and support a design approach toward the development of a recurrent-use decision-support tool in a sociotechnical system such as healthcare. Each of the stages are discussed in turn in the next section, including changes made to the framework in light of its application. The changes are summarised in Table 8.1.

Table 8-1Summary of changes to IHAF in light of application

IHAF component revisited	**Summary of review of component**	**Change made**
Design Science methodology	The build-evaluate cycles of Design Science research supports evaluation in context, which includes examination of social and technical components (Chapters 3, Section 3.4 and Chapter 4, Section 4.3).	None
Problem definition	Stakeholder management assumes greater importance when developing a model for recurrent-use (Chapter 7, Section 7.3.3). Stakeholder groups are relevant throughout model build and evaluation.	Problem definition phase is expanded to be called 'Define problem and identify stakeholders'
	The wider implications of scenario interventions should be considered from the perspective of system resilience (Chapter 7, Section 7.3.1).	None
	A formative evaluation is optional but Chapter 5 and Chapter 7 (Section 7.3.2) demonstrated the value of this phase.	'Formative evaluation' is added to the problem definition phase as optional.
Hybrid Modelling stages	For an embedded solution, stakeholder engagement and the potential need for qualitative evaluation methods emphasises the need for participatory approaches.	The data collection box in IHAF within the 'Describe' component has 'workshops' and 'PSMs (Problem Structuring Methods) added as example methodologies.
	Consider scenario flexibility to maximise the value and utility of the HM, its real-time data, and its sustainability as a decision-aid (Chapter 7, Section 7.3.1; Chapter 6, Section 6.7).	None
	Triggers may be reactive, proactive or predictive, and simulation may form the predictive phase with no prescriptive phase (Chapter 6, Section 6.7).	None
	Decision-makers retain autonomy and may choose not to act. Any actions taken as a result of the	The arrow between prediction/forecasting

	decision may or may not have been informed by components of the HM.	to SA and decision-making is converted to a dashed arrow.
Evaluation	Consider investigating SA as a component of formative and/or summative evaluation (Chapter 7, Section 7.3.4).	The 'evaluation' stage remains as the last stage of the framework, but is extended to subsume all previous stages to indicate that evaluation activities occur throughout the development cycle.
	Evaluation of each component in situ can identify both expected and unexpected outcomes. They may be positive or negative (Chapter 5, and Section 7.3.2).	
	Consider early issues regarding model ownership, maintenance, and sustainability (Chapter 7, Section 7.3.5).	
	Evaluation of the model should include demonstration (Chapter 7, Section 7.2.4)	Evaluation component is changed to 'Demonstrate and Evaluate'

8.1.2 *Revisiting the problem definition stage*

IHAF provides a conceptual model for the HM which defines its purpose, namely its contribution to task-level and system-level situation awareness (SA). Design Science suggests that researchers address relevant problems in contexts that require a designed system to provide a 'solution' (O'Keefe, 2014), in this case, enhancing SA for short-term decisions. This problem must exist in the natural environment, although part of it might be extracted, for example to build the model. The problem is repeating or regularly occurring, rather than one-off, and thus requires an embedded, recurrent-use solution. Within IHAF, the *problem definition* stage contains each of these requirements, as well as the criteria for evaluation, and the possibility of a formative evaluation, and has acknowledged applicable prior theory as criteria for evaluation. A reflective understanding of the model's limitations and how it is being used can ultimately increase the level of trust and confidence toward successful implementation. The formative evaluation can form a component of the problem definition stage (Venable et al., 2016). It enables the possibility of reducing risk by evaluating early, before building the model, while the summative evaluation forms the final component of the methodology. This has been added to the problem definition stage of IHAF (Figure 8.2, Table 8.1).

While identifying and engaging stakeholders is implicit to this stage, the *evaluation* component in Chapter 7 identified stakeholder discord that is likely to present a significant barrier to future testing and development iterations of the HM, and is likely to require explicit consideration in any similar application in a sociotechnical system. For this reason, the problem definition phase has been modified to explicitly incorporate 'Identify Stakeholders' (Figure 8.2, Table 8.1). This requires identifying all stakeholder groups at whom the value proposition is aimed, on both the supply-side and the demand-side, actions needed to foster key stakeholder support, and what actions might be needed to work toward implementation and sustainability. The use of problem structuring methods or other participatory processes can be considered. Stakeholders, users and the potential to implement change can all present as constraints to the design process.

8.1.3 *Revisiting the hybrid model stages*

Identifying the necessary data is a key early part of an M&S study, and forms part of the problem definition stage, hence the close proximity of these stages in IHAF. When creating a solution for supporting short-term decision-making, the quality and availability of data can present a further potential constraint, particularly where real-time data is required, and where data is considered to be of a sensitive nature, such as in healthcare. Significant stakeholder engagement and co-creation processes were required for the data to be made available to *NHSquicker,* as indicated in the *descriptive* stage in Chapter 4. To emphasise this, the data collection box in IHAF within the 'Describe' component has 'workshops' and 'PSMs' (Problem Structuring Methods) added as example methodologies to emphasise the need for participatory methods (Figure 8.2). Furthermore, additional data requirements may be needed to monitor for unanticipated, particularly negative, effects. These may be at different levels of the system, for example demand management within the hospital was shown to impact demand in other parts of the urgent care system.

The data needed to make an accurate *diagnosis* of the system state will also require consideration, and may present a compromise. Nonetheless, IHAF is intended to be used iteratively, and significant learning can come from early cycles. This might include the use of near real-time data which is updating daily or weekly, or manual data updates for example staff absence or shifting rotas.

The constraints presented by data quality and availability (Espinoza et al., 2014) can point toward future improvements.

The *predictive* element has been presented as optional, where a predictive trigger is required and where the trigger can be forecasted. Alternatives are reactive triggers (the simulation triggers when the critical threshold is reached in real-time) and proactive triggers (the simulation is triggered at regular intervals) (Table 8.2).

Table 8-2 Examples of triggers in previous research, which can be implemented in IHAF

Type of Trigger	Description	Reference
No trigger	The predictive component informs action. Simulation may be used for prediction rather than prescription.	Hoot et al. (2008)
Reactive Trigger	The simulation triggers when a critical threshold is reached in real-time.	Marmor et al. (2009) Augusto et al. (2018)
Proactive Trigger	The simulation is triggered at regular intervals.	Bahrani et al. (2013) Oakley et al. (2020)
Predictive Trigger	The simulation triggers when the critical threshold is forecasted.	Aydt et al. (2009b) Harper & Mustafee (2019)

Different approaches can be used, including time-series forecasting and machine learning. However a limitation to these approaches is that while a critical situation can be forecasted, the underlying causes which may suggest action, are not necessarily apparent (Menke et al., 2014). Simulation can also be used for prediction, and the evaluation found significant interest in the use of better forecasts for supporting adaptive behaviours. Simulation can provide additional information, for example, Hoot et al. (2008) used DES to model an ED to forecast near-future operating conditions, outputting a range of input, throughput and output measures. This allowed it to distinguish between causes of crowding, which can point to both adaptive behaviours and escalation actions. Depending upon the problem under investigation, the output may stop here, for example Oakley et al. (2020) developed a real-time simulation model for predicting bed capacity as an early warning system. IHAF can support a flexible use for simulation, including a *prescriptive* component, as used in the case study in Chapter 6. Where process changes are not a priority, but predictions can support decision-making, the purpose of the simulation may be *prediction*. In contrast, where optimal or near-optimal solutions need to be generated, simulation may be combined with optimisation to identify the best solution in real-time as a *prescriptive* method (Onggo et al. 2018). Therefore the *prescriptive* stage, like the *predictive* stage, is an optional component. Within IHAF, the arrow between prediction/forecasting to SA and decision-making is converted to a dashed arrow. This is because decision-makers retain autonomy and may choose not to act. Any actions taken as a result of the decision may or may not have been informed by components of the HM (Figure 8.2, Table 8.1)

8.1.4 *Revisiting the evaluation phase*

To evaluate, any appropriate method or range of methods can be used on any iteration of the HM. In line with the methodology outlined by Peffers et al. (2007) demonstration has been made explicit in IHAF, to enable users to more clearly understand the potential value of the HM in the final evaluation stage.

A fundamental component of IHAF is that criteria for evaluation and influencing factors are identified early (at the problem definition stage) from the literature, and from the other sources such as workshops or questionnaires, and Section 8.1.2 has shown that it is valuable to identify stakeholders for building the model, but also for testing and determining how it is used in practice. The evaluation aims to determine whether the model has the expected effect on decision-support suggested from the literature and from existing studies. However, testing the framework has found that evaluation is continuous between these two stages, for example collecting and validating data, simulation building, verification and validation activities, and stakeholder engagement. Some engagement will be informal, although appropriate methods are needed to capture the outputs from this engagement. For this reason, in the framework, the evaluation stage is extended to subsume all of the IHAF stages, culminating in the final evaluation stage. This also means that further iterations or new applications of the framework start, by definition, with evaluation. Furthermore, as iterations progress, consideration of long-term evaluation, and of generic applications for use in similar settings will become a focus early in the process. The revised framework is in Figure 8.2, with the revisions highlighted in orange, and can be compared with Figure 8.1.

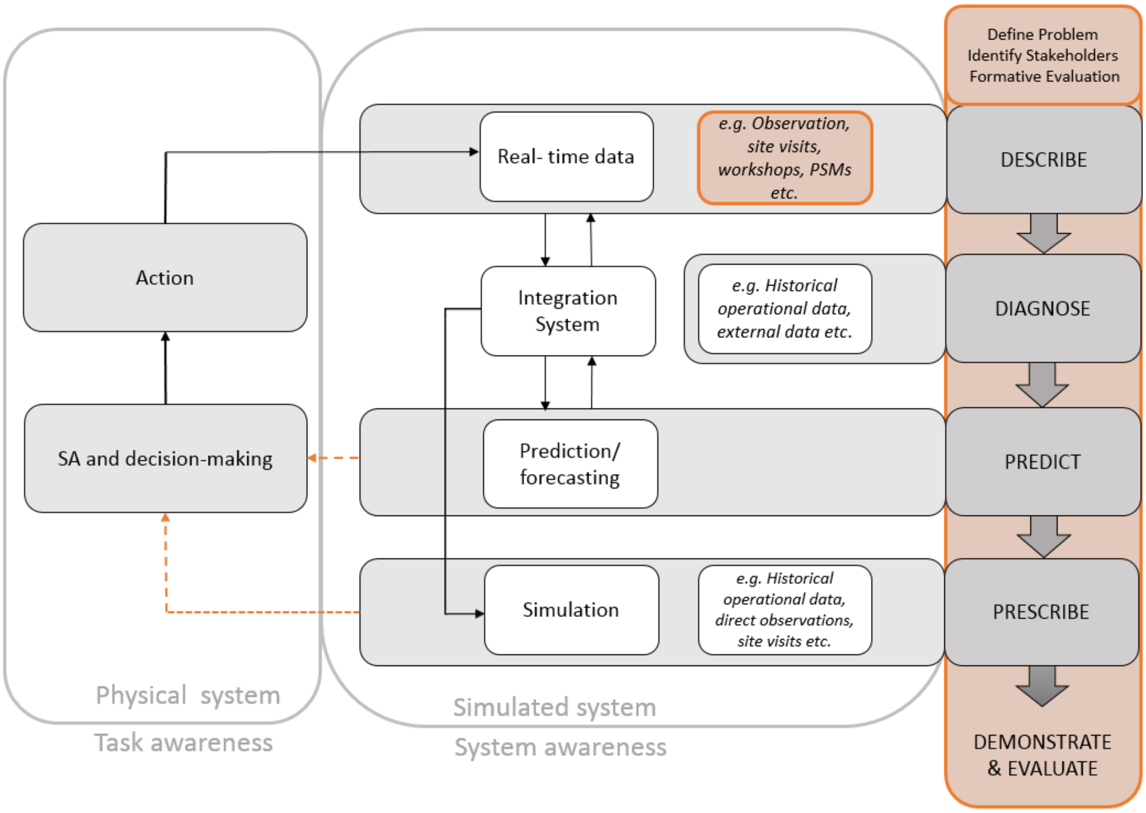


Figure 8-2 IHAF following evaluation and modification

8.2 The transferability of IHAF

IHAF was developed as a generic framework for supporting the development of real-time decision-support tools, motivated by ED as a fast-paced, dynamic, sociotechnical system, but with a clear ambition of its use in other application domains. Design Science was found to provide a useful and applicable methodology with which to base the framework. Within this methodology, mixed-analytic approaches have been demonstrated toward enhancing SA, a precursor to decision-making and action in sociotechnical systems, which focuses attention on providing the right information at the right time to support a continuous understanding of the system state. This emphasises the role of choice and organisational design in the interaction between people and technology, in any domain.

A core value of the sociotechnical system approach is that, given the right choices, social and technical systems can be synchronised such that productivity, worker satisfaction and safety can be optimised in parallel (Clegg, 2000;

Waterson et al., 2015). While simulation modelling in sociotechnical systems is commonplace, the challenges and barriers to integrating a real-time decision-support tool represent different issues which are widely applicable. This means that the application of IHAF is relevant to other sociotechnical systems. Issues of socio-organisational context, organisational culture and behaviour, the external environment, interaction with tools and technologies, and evaluating, maintaining and sustaining an intervention are fundamental challenges associated with new technologies in complex systems (Carayon et al., 2015). To increase the robustness of the intervention, the more situations a design has been shown to work, the more likely it is considered to work for similar new problems. This presents a limitation of this formative research, as the intervention is tested in one use-case.

Nonetheless, IHAF supports a focus on emerging or latent risks, and on progressive, iterative development that builds on a cumulative knowledge base toward real-time analytics tools that are useful and usable in practice. Additionally, the framework is flexible, supporting multiple data sources and methods, and conceptually extendable, such that, for example, the prescriptive component may extend to optimisation, design-of-experiments or machine learning to identify the best solution based on predefined objective functions (Onggo, 2019). As complexity increases, for example through multiple sensor feeds, or dynamic changes in the physical system, new issues may arise, however the iterative approach, which attends simultaneously to development and implementation, can support these challenges.

8.2.1 *Examples of other applications of IHAF*

The application of IHAF is not domain-specific. For example, in the transport sector, sustainable mobility is more than a question of technology. Cities are complex systems, where mobility is only one element, hence the challenges and barriers to new models of mobility must be examined in the context of these sociotechnical interactions. The concept of multiple passenger ride-sharing allows passenger to choose their pickup and drop-off time and locations, and allows multiple passengers to share the route (Ma et al., 2015; Linares et al., 2016). The fleet of vehicles and passenger requests can be tracked in real-time (*descriptive* stage), and using a reactive trigger (*diagnostic* stage), simulation (*prescriptive* stage) can determine the best route for both the passengers' wait-time tolerance, arrival destination, and arrival time flexibility. *Evaluation* of this approach in the early phases would examine the technological reliability and validity of the dynamic route calculations. Passenger satisfaction with the concept of transitioning from vehicle ownership to vehicle 'usage' and the inherent loss of flexibility associated with this is also an important part of evaluation and sustainability of the approach.

A second example is in the field of water management (e.g. Wu et al., 2011). The system of water distribution is required to ensure a safe, reliable and efficient delivery of water supply to consumers. However with ageing infrastructure and expanding populations, proactive management is required to maintain system resilience by providing detailed analysis of the system condition in real-time. Real-time monitoring systems (*descriptive* stage) update the data acquisition database, and a short-horizon forecasting model (*predictive* stage) with a predictive trigger *(diagnostic* stage) triggers the simulation (*prescriptive* stage), providing results including decisions, alarms and parameter settings to perform preventative maintenance before an event occurs. These faster response times are economically efficient, as well as improving customer experience. An alternative approach is to use a proactive trigger, running the simulation at a predetermined frequency, and using the simulation to predict events to support operators to find remedial actions. This means that network failures can be detected at an early stage such that operators can react quickly to minimise damaging effects. A range of *evaluation* measures can capture the impact of the application.

A final example of a potential application of IHAF is in police routing (Dunnett et al., 2019) ensuring the most efficient resources are allocated in the case of incident response. In this application, the *descriptive* stage consists of incident reports, response unit availability and demand coverage (reactive trigger, *diagnostic* stage); the *predictive* stage forecasts traffic conditions; and the *prescriptive* stage can indicate the best unit response to the incident. These examples emphasise the flexibility and utility of IHAF.

8.3 Chapter Summary

As a result of applying IHAF, some modifications were made to the framework. This included extending the evaluation across the development stages, recognising that evaluation is a process. A reflective understanding of the model's limitations during development can ultimately increase the level of confidence toward successful implementation. A discrete demonstration and evaluation stage is retained at the end of the process, to determine whether the model has the expected effect on decision-support. This also means that further iterations or new applications of the framework start, by definition, with evaluation.

Cumulative knowledge creation that can be generalised beyond individual solutions to individual problems can occur where settings are similar, in particular where research involves social dimensions, and insights might be transferred from one to the other. The lessons learned from the application of IHAF can be extended to other sociotechnical systems, aiming for alignment of social and technical systems in decision-support. Additionally, the framework itself is high-level, flexible and extendable. IHAF supports a focus on managing risks by attending simultaneously to development and implementation, and on progressive, iterative development that builds on a cumulative knowledge base toward the implementation of real-time decision-support tools that are useful and usable in practice.

The final concluding chapter summarises the main contributions, limitations and areas for future work.

Chapter 9: Conclusion

This chapter presents a summary of the thesis and its contributions. This research has taken a critical realist perspective, which highlights diverse forms and types of knowledge of practical value (Mingers, 2008; Archer, 2013; Syed & Mingers, 2018). This approach seeks to narrow the research-practice gap, by enabling recognition of problems of relevance to organisations, including ethical dimensions, toward solutions that are useful in their application domain. The research identified the potential value of a real-time hybrid model (HM) for short-term decision-support with a particular focus on healthcare. It proposed a generic conceptual framework – IHAF - for real-time HMs used to support short-term decision-making in sociotechnical systems. The framework was tested in practice using a use-case in an emergency department (ED) in the UK, and evaluated at the system level using patient questionnaires and staff interviews. When an intervention is tested in a sociotechnical system, it can be difficult to characterise how the system will act, measured with relevant real-world criteria. The findings indicate that there is a need for short-term decision-support in healthcare, but that developing M&S decision-support tools that aim to be embedded into the system brings particular challenges. These can benefit from an approach that integrates evaluation as development and implementation co-progress through iterative study cycles. Acknowledging the complexity of sociotechnical systems, and that people are at the centre of these systems, supports a focus on quality and safety, as well as efficiency and performance. The remainder of this chapter will address the findings, limitations, and opportunities for improvement and future work. The contribution of the research is addressed in Section 9.1. Section 9.2 presents a summary of the research. The research limitations are explained in Section 9.3. Finally, Section 9.4 discusses future directions of research in the area of real-time hybrid modelling in sociotechnical systems.

9.1 Research Findings

The increasing opportunity to use real-time data to support quick and effective decision-making, and the need for short-term decision support in ED motivated this work. The contributions to knowledge of this research are specified below.

RQ1: The need and opportunity for short-term decision-support in healthcare

A review of the literature identified the need for short-term decision-support in healthcare, particularly in ED, and the value in the use of real-time simulation and analytics for healthcare decision-support. It also found that viewing healthcare as a sociotechnical system enables both social and technical elements of the system to be taken into account when developing a recurrent-use HM for decision-support. The purpose is to support the progression of the HM toward sustained implementation in practice.

M&S studies are used widely to gain insights into existing or proposed systems of interest, with hybrid simulation and HM used to better represent the system of interest. Powell and Mustafee (2016) made a distinction between hybrid simulation (HS), where two or more simulation methods are combined, and HM where simulation is combined with other distinct methods at specific stages of a simulation study. The majority of these studies are single-use models, while real-time simulation – in its infancy in healthcare – can provide short-term decision-support in a recurrent-use tool. However the articulation between real-time decision-support tools and a sociotechnical approach to their development and implementation was found to be lacking. This means attending to both social and technical elements, as people in the system will interact with the tool and make decisions based on its output. *The primary finding from RQ1 is that there is a need for a generic conceptual framework that supports the development of a real-time HM in sociotechnical systems such as ED.* Considering decision-making as a consequence of situation awareness (SA) focuses the HM on what information is needed where, when, how, and by whom, and centres the problem definition and evaluation on aspects of the model that can enhance or impede SA in practice.

RQ2: A conceptual HM framework for supporting short-term decision-making in sociotechnical systems

The literature review found that a real-time decision-support system which combines real-time data, predictions, and simulation has the potential to support short-term ED decision-making. Having identified the need for a generic conceptual framework for short-term decision-support in sociotechnical systems (RQ1), the main finding of RQ2 was establishing the sub-components of the framework through a review of the literature and the chosen methodology. There are currently a lack of studies which have adopted this approach for this purpose. *The Integrated Hybrid Analytics Framework (IHAF) is the main contribution of this thesis.* While the framework is intended for developing HMs for short-term decision-support in sociotechnical systems, engaging with Design Science methodology has identified that the first stage of the framework, the *Problem Definition* phase, should also determine a set of criteria for later evaluation of the HM. A proposed set of generic criteria were identified from a review of the literature to support the development and evaluation of the HM in sociotechnical systems. An ideal model should demonstrate the fundamental application of these criteria if it is to be useful in the real-world, while additional criteria applicable to individual cases forms part of the problem definition stage of IHAF.

RQ3: The system-level implications of a real-time HM

Having applied the IHAF framework, and evaluated the application according to the generic criteria identified (RQ2), RQ3 analysed the system-level impact of the case study. While from a systems perspective it is important to consider the impact a HM will have as a result of stakeholder decision-making, for a recurrent-use tool the short- and long-term impacts require closer consideration. *Consequently, five system-level outcomes were identified from the formative (questionnaires) and summative (interviews) evaluations, which can be generalised to customers or users in other sociotechnical systems.* These are: (i) Building resilience; (ii) Managing stakeholders; (iii) Unintended consequences; (iv) Situation awareness and decision-making; and (v) Model embedding and sustainability. *These findings can act as evaluation criteria for future real-time simulation or real-time HM studies in sociotechnical systems*, in particular in healthcare, building on the criteria identified in RQ2. *Building resilience* refers to the need to consider the wider system when planning scenarios which support system recovery from critical events. The wider context and its contribution to system resilience need to be considered if the HM is to be useful in the long-term. *Managing stakeholders* refers to the need to view sociotechnical systems as political entities, as this context can impact the implementation process and sustainability of the solution. *Unintended consequences*, both positive and negative should be investigated in all relevant stakeholder groups, particularly in systems where safety is a relevant outcome, and where evidence of safety is a requirement. Whilst a key feature of the IHAF framework is to make explicit the purpose of a real-time HM for *enhancing SA*, it was found that to build information systems that can support complex decisions it may be necessary to more fully understand human decision-making processes and how these decisions can be supported in practice. Decision-analysis and methods for measuring SA may usefully be integrated into the Problem Definition and Evaluation stages. This can subsequently support the HM development stages of the IHAF application. These findings strengthen future applications of IHAF by identifying challenges that modellers may face when developing a viable real-time HM toward *implementation and sustainability* in a sociotechnical system, and ensuring that the Evaluation stage adequately addresses these challenges.

9.2 Summary of the thesis

The research questions (RQs), aims and objectives, and methods employed to realise the RQs and objectives during this thesis are outlined in Table 9.1, and summarised in the following subsections. The implications for practice are emphasised.

Table 9-1 Research questions, aims and objectives, and methods

Research Questions	Aim	Objectives	Method
1. *How can simulation approaches support short-term operational decision-making in healthcare?*	To determine the need for short-term decision-support in healthcare, and to examine how simulation, real-time simulation, and hybrid modelling approaches have been used for short-term operational decision-support in the healthcare context, in particular emergency care.	**1.** To explore the need for short-term decision-support in healthcare, in particular emergency care.	**1**. Literature Review (Chapter 2)
		2. To explore how analytics methods can be used for short-term decision-support.	**2**. Literature Review (Chapter 2)
		3. To critically evaluate simulation approaches used in healthcare for decision-support and to identify how simulation is used for short-term decision-support.	**3**. Literature Review (Chapter 2)

		4. To determine criteria for evaluation of a hybrid simulation approach for short-term decision-support in healthcare.	**4.** Literature Review (Chapter 2)
2. *How can an integrated hybrid approach using real-time simulation and data analytics support short-term operational decision-making?*	To test and evaluate the potential of an integrated hybrid approach for short-term decision-support in healthcare combining real-time simulation with other analytics approaches.	**1.** To propose a generic framework supporting an integrated hybrid approach for short-term decision-making in healthcare.	**1**. Literature Review and Design Science approach (Chapters 2, 3 and 4)
		2. To apply the framework within the case study in a hospital ED.	**2.** Direct observation, patient questionnaires, secondary data analysis, time-series forecasting, real-time simulation (Chapters 5 and 6)
		3. To evaluate the application in this context.	**3.** Semi-structured staff interviews (Chapter 7)
3. What are the implications and the added value to the system of using real-time data applications for both patient and NHS decision-support?	To analyse the system level impact of the use of real-time data by both patients and staff.	**1.** To critically evaluate the value that real-time applications provide at the system level.	**1.** Patient questionnaires and semi-structured staff interviews (Chapters 5 and 7)
		2. To synthesise previous findings and evaluate the framework in light of the application.	**2.**Synthesis of findings (Chapters 7 and 8)

9.2.1 *The need and opportunity for short-term decision-support in healthcare*

Over forty years ago, Bostom and Heinen (1977) argued that OR needed to be reframed within a sociotechnical systems design approach. A sociotechnical system contains both social and technical elements. The technical elements are concerned with the processes, tasks and technology needed to transform inputs to outputs. The social system is concerned with peoples' attitudes, skills, values, the individual and group relationships, the authority structures and reward systems. Outputs of the system occur as a result of the interactions between these, and Bostom and Heinen (1977) argued that any OR design intervention with a view to improving system functioning must contend with both system elements. Healthcare is an example of a sociotechnical system, but very few healthcare M&S studies frame their work with an equal focus on technical elements and the people in the system, both staff and patients.

As NHS healthcare services are likely to become progressively more constrained financially, and the medium- to long-term impact of COVID-19 remains unknown, system resilience is likely to be impacted. Data Analytics plays an important role in improving the delivery of healthcare services (de la Torre Diez et al., 2016), but despite the volume, velocity and variety of data being produced, it is arguably not yet being fully exploited for enhanced effectiveness and efficiency of delivery (Wang et al., 2017, 2019; Mehta et al., 2019). To gain impact from data requires a focus on intervention and change (Rasmussen & Ulrich, 2015). Healthcare analytics publications have proliferated in the last five years as the value in data is increasingly realised (Günther et al., 2017; Galetsi & Katsaliaki, 2019a), and data-driven decision-making is gaining traction in healthcare. The technology needed to design and create real-time simulation models to support system recovery is not new. Nearly thirty years ago, Annan and Banks (1992) described the architecture of a supervisory control system via online simulation in a shop floor, focusing on interfacing sensory information and operator knowledge with real-time knowledge bases and simulation. In healthcare, ED has been the focus of multiple studies using real-time simulation (e.g. Tavakoli et al., 2008; Marmor et al., 2009; Espinoza et al., 2014; Augusto et al., 2018), while Oakley et al. (2020) expanded the area of reach to the management of inpatient beds.

These studies focused on specific technical challenges associated with real-time simulation such as model performance in the absence of adequate data, and validation of real-time models. However, as the volume of healthcare data continues to multiply, the benefits and value created by real-time DA in healthcare still remains relatively unexplored. For example, despite increasing interest in the use of real-time decision-support tools in healthcare, including real-time simulation, there was found to be a gap in understanding what works in practice. RQ1 focused on the opportunities and barriers to the development and use of a real-time HM application to support SA and enhance system resilience, by taking a sociotechnical perspective. Additionally, positioning the application in a QI approach ensures that all relevant stakeholders are considered, and that quality and safety are as important as efficiency and productivity, which has clear implications for M&S practice.

9.2.2 *A framework for supporting the development of a real-time hybrid model for short-term decision-support in healthcare*

In response to the first research question, Research Question 2 proposed and tested IHAF, a generic conceptual framework for the development and application of a HM for short-term decision-support in sociotechnical systems, which starts from the assumption that the HM will be useful in practice. IHAF was tested in a use-case which combined real-time data, forecasts, a predictive trigger and discrete-even simulation in a HM, with a view to supporting decision-making aimed at the reduction of ED crowding. A clear advantage to having a predictive stage in the HM is that it addresses all levels of SA: perception of elements in the environment; comprehension of their meaning; and projection of the future state. Anticipation is a key element of SA and adaptive action, by detecting as early as possible that a critical event is imminent, however a reactive or proactive trigger can mitigate the effects of an event.

As a result of the application and evaluation, IHAF was modified, in particular subsuming all stages into evaluation, to reflect the fact that this is a process, and not an event. However the final stage remains an *evaluation* stage, which includes demonstration of the model, its components, or outputs, and its contribution to short-term decision-support. The conceptual framework is high-level, flexible and extendable, supporting multiple methods and interactions. IHAF implicitly points to a relevant, user-centred design, aligning both social and technical factors, and to considering the unintended consequences of the approach in a continuous way. It enables a focus on managing risks by attending simultaneously to development and implementation, and on progressive, iterative development that builds on a cumulative knowledge base toward the implementation of real-time decision-support tools that are useful and usable in practice. This is the primary contribution of the thesis.

9.2.3 *The system-level implications of a real-time hybrid model in sociotechnical systems*

The case study found that implementing a real-time decision-support tool in healthcare is inherently challenging. Positioning the framework in Quality Improvement (QI) takes a multi-dimensional approach to interventions in healthcare, concerned with efficiency, equity, effectiveness, safety, and taking a patient-centred approach (IoM, 2001). QI therefore synergises with a sociotechnical approach by addressing both social and technical factors. The increasing complexity of healthcare plays a significant role in its vulnerabilities and risk of error, and where the objectives of people or subsystems within the system are not aligned, inefficiencies and other quality problems can arise (Carayon et al., 2011). The case study found multiple examples of these issues, which are characteristic of complex adaptive sociotechnical systems.

The specific contribution of the framework is at the level of 'nascent design theory', described by Gregor & Hevner (2013) as "knowledge as operational principles or architecture". In addition to a knowledge contribution (IHAF), Gregor and Hevner (2013) argued that Design Science should also make a clear contribution to the real-world application environment from which the research problem or opportunity is drawn. Based on this principle, the following summarises the implications for practice arising from the empirical findings.

Synthesis of the literature, the formative evaluation (patient questionnaires) and the final evaluation (staff interviews) found five significant themes:

(i) Building resilience across wider networks

The real-time HM aims to support decisions made in real-time to reduce operational pressure in a hospital, however these may have negative consequences on other parts of the wider system. For example, in healthcare one principle of good patient flow is to make sure that there is sufficient capacity in all parts of the system. From the evaluation, all stakeholder groups identified the need to take a system-wide view, yet there is little evidence that a whole system approach is currently being utilised to tackle short-term demand management in urgent and emergency care. Working together at a regional level helps build system resilience, and the HM extends the value of the available real-time data by accounting for the current state of the MIUs in the system to enhance system resilience by spreading the risk across an entire urgent and emergency care system (Higginson & Boyle, 2018). Different sociotechnical networks will manage their service delivery differently, and interactions between and across systems do not necessarily have to be resolved. However, if long-term sustainability and transferability (spread) are to be achieved beyond a successful demonstration project, then the wider context and its contribution to system resilience need to be considered.

(ii) Managing stakeholder tension

The evaluation found that tension between managers and clinicians is an organisational issue which can present as a significant barrier to progression of the HM. The extent of the issue will depend heavily on the sociocultural context, but is a recognised phenomenon worldwide in healthcare (Powell & Davies, 2016). Organisational politics describes the systematic use of power and influence by employees to resolve conflicts and meet personal or organisational goals (Drory & Vigoda-Gadot, 2010, Kapoutsis et al., 2016). In any sociotechnical system, the impact of organisational politics and trust across hierarchies can present a challenge (Lampaki & Papadakis, 2018). Viewing sociotechnical systems as political entities, where actors have different needs and aspirations, accepts that context can impact the implementation process. Professional resistance to technology may be due to lack of knowledge or skills, or may occur when the roles and practices assumed by the technology threaten values and norms. One solution to these issues may lie in the use of problem-structuring methods or facilitated modelling. Long et al. (2019) found empirically that internal politics, stakeholder commitment and involvement, and stakeholder-researcher communication, were key to implementation of simulation results in healthcare, and that flexibility and reflection on opportunities and barriers were necessary throughout the study lifecycle. Issues of internal politics affect all sociotechnical systems (Lampaki & Papadakis, 2018), and IHAF supports continuous appraisal of these issues and their interactions throughout the study process.

(iii) Unintended consequences

In the use-case application, two main stakeholder groups were identified: patients (users) and staff (beneficiaries). From the observational data, clinicians expressed concern that real-time wait-time information might support patients to make sub-optimal attendance decisions. However the evaluation found that a subset of higher-risk patients are more likely to use the information to improve their own experience of attendance, without changing their attendance decision. This is an example of a positive unintended consequence of providing additional information for patients to make attendance choices. A further example of a positive unintended consequence is staff using the real-time information as a quick and convenient check whilst on-the-go. This is an unintended consequence of the information being made available via a mobile phone application, while most current information for supporting operational decisions requires access to a computer. The evaluation also identified a possible negative consequence. Both evaluation stages with staff and patients found that predicted wait-times are more likely to support attending at a different time of day, than attending a different facility. The impact of this is on the NHS, as patients are better spread across the system than utilising quieter times of day where staff resources are reduced. These findings emphasise the importance of examining the consequences of a HM and its components, both intended and unintended, on all relevant stakeholder groups. This is relevant for any sociotechnical system, as unintended consequences can present unpredictably. Where there are risks to safety, this issue is essential.

(iv) Situation awareness and decision-making

A key feature of the IHAF framework is in making explicit the purpose of real-time hybrid decision-support tools for enhancing SA, an important constituent in decision-making processes. This focuses attention on what information is needed to support a continuous understanding of the current system state in a complex, dynamic system. This is relevant, for example, when considering the design of the tool, the information outputted, and how the information will be accessed. For example, a black box approach (Varshney, 2016) may improve interpretability and reduce cognitive load.

One way of determining what information is needed is through decision analysis. From the evaluation, it was found that there are differences in how staff make operational decisions, yet there is a gap in the literature toward understanding how staff make dynamic operational decisions amidst interruptions, distractions, and uncertainty. In order to build information systems that can support complex decision making it may be necessary to more fully understand human decision-making processes and how these decisions can be supported in practice (Zsambok & Klein, 2014; Catchpole & Alfred, 2018). This is an area that is likely to benefit from significant further work, in particular in sociotechnical systems where there are high-risk consequences to operational decisions such as the police, healthcare or emergency services. Within IHAF, decision-analysis may be useful in the formative evaluation stage, and specific measures of SA, such as the Critical Decision Method (CDM) can be used in the final evaluation stage.

(v) Model adoption and sustainability over time

The final finding is the issue of HM adoption, sustainability and long-term evaluation. The problem of low implementation of the results of simulation models in healthcare has been under discussion for decades and the challenges of engagement with stakeholders is often considered to be a primary reason. For recurrent-use tools, this issue is likely to be amplified. Further, adoption decisions can be influenced by lack of evidence of the benefit of the HM, but it is a challenge to predict the use, impact, and amount of investment needed to adopt, maintain, and sustain the tool. The evaluation stage of the application suggested that the HM could best be adopted and maintained internally by a dedicated modeller, however without firm evidence of its benefit, this is likely to be insurmountable.
Using IHAF to support HM development, the process of development toward implementation against a dynamically evolving context is done incrementally. Accepting complexity is likely to be key to the adoption and embedding of new OR technologies into healthcare practice. Long et al. (2019) examined factors related to implementation of simulation in healthcare and found there is no way to successfully pre-empt or plan for changes in the implementation context. This is relevant for any sociotechnical system, and reflective consideration of the interaction of different local implementation factors is therefore required to allow researchers to intentionally and effectively respond to these challenges, remaining mindful of emergent opportunities, outcomes and threats throughout the study process.

9.3 Limitations

There are several limitations of this research. These are summarised below, discussed as limitations of the IHAF framework, limitations of its application in the case study, and an overarching limitation:

- Limitations of the framework:

(i) The IHAF framework was tested on one use-case only, which limits its transferability to similar healthcare situations. However due to the PhD timeline, it was only possible to test it in one ED. Using a 'typical case' (Chapter 5, Section 5.2) allows generalisations to be made from case to similar case, in particular when triangulated with multiple data sources and with existing literature. Stake (1995) called this *naturalistic generalisation.*

(ii) The transferability to other sociotechnical systems is untested. Nonetheless, for the research design, one case was considered sufficient to answer the research questions. Testing the framework in other domains is planned as future work.

(iii) Design Science has become a well-accepted research methodology in computer science (CS) and information systems (IS) (Peffers et al., 2018), although the methodology has faced significant criticism over the decades, in particular in IS where much of the criticism is focused on lack of attention to the knowledge contribution of the approach (Gregor & Hevner, 2013). This research explicitly states that IHAF is the major contribution to knowledge, as a real-time HM framework for short-term decision-support in sociotechnical systems. The approach is new, interesting, and makes a genuine contribution to knowledge (Gregor & Hevner, 2013). However these authors also state that Design Science should make a contribution to praxis, that is, a clear contribution to the real-world application environment from which the research problem or opportunity is drawn.

- Limitations of the application:

(i) The questionnaires used for the formative evaluation were a relatively small sample size but provided indicative outcomes of how the real-time and predictive time data may be used by patients to support attendance decisions. However further evaluation may be needed to determine how the real-time information is actually being used in practice (*Formative Evaluation/Problem Definition Stage,* Chapter 5).

(ii) Currently *NHSquicker* data is only available in the South-West of England. This means that further applications of the same HM in other sites can only be tested in this region (*Descriptive Stage,* Chapter 5).

(iii) The forecasting methods were tested in one ED only, and may perform differently using different datasets. Additionally, time-series forecasting was applied in this research however other methods such as neural networks or ensemble methods may perform better but have not been investigated. The advantages to these approaches are that they can account for other factors that may influence forecasts, such as incorporating ambulance arrivals data, the weather, or special events (*Predictive Stage,* Chapter 6).

(iv) While designed to be generic, the DES model was developed for a single ED and hasn't been tested in other EDs. It is reliant on the ED reporting data across hospitals being presented in the same format. While EDs in the UK collect data for performance reporting, information used to build the model (such as distribution of treatments and investigations per triage category) may not be standard in databases from other EDs (*Prescriptive Stage,* Chapter 6).

(v) The use of open-source software requires consideration, in particular due to the limitations of the Personal Learning Edition (PLE) of AnyLogic. A tool developed using open-source software can be more easily shared, scrutinised, tested, and adapted in different environments. AnyLogic PLE 8.5.2 was used for the development of the DES model. It does not support exporting a model as a standalone application, nor can it be uploaded into AnyLogic Cloud to send/receive data from third-party applications. This is required for a Java application to call the AnyLogic model and pass it the real-time data parameters, execute the model, and receive the experiment results back to the Java application (*Integration,* Chapter 6).

(vi) The staff interviews used for the evaluation stage provided a useful summary of factors which matter to staff, however due to COVID-19, the number of interviews completed were limited. It is anticipated that these can be completed at a later date, and knowledge gained from the preliminary analysis in this research can be used to adapt the interview schedules to inform the next stage of evaluation, and increase the validity of the findings more generally. Nonetheless, there was significant consensus across interviews, supporting analysis (*Evaluation,* Chapter 7). Baker & Edwards (2012) emphasised that more interviews is not necessarily better, and for this application the aim was to gain the information needed to move toward the next phase of development.

(vii) The research application was researcher-led, and motivated partly by maximising the potential value that can be gained by accessible real-time healthcare operational data, as well as the identified need for short-term decision-support in ED. For this reason, there were barriers and challenges to accessing and maintaining the interest of relevant stakeholders. It was observed that throughout the duration of the work, interest-levels, priorities and key staff were in constant flux, which acted to limit the progression of the work at times.

- Limitations to findings

(i) The general findings discussed in Section 9.2.3 which address RQ3 are derived from the data (observational data and questionnaire data, Chapter 5; and interview data, Chapter 7). They are also synthesised with the literature. However it is acknowledged that these are not definitive, and do not necessarily form a complete set. Nonetheless, one of the principles of Design Science is the reuse of extant contributions and the accumulation and evolution of design knowledge (Vom Brocke et al., 2020). This is seen in the Design Science methodology (Peffers et al. 2007) in Chapter 3. The implications of this are that the general findings in RQ3, which are contributions of this research, can act as antecedents or evaluation criteria for future real-time simulation or real-time HM studies in sociotechnical systems, in particular in healthcare. For this study, the evaluation criteria were identified in Chapter 4 as generic criteria from the literature.

9.4 Future Work

9.4.1 *Future work on IHAF*

The IHAF framework aims to contribute to the accumulation of knowledge through empirical studies, which can be generalised beyond individual solutions to individual problems. IHAF is specific in terms of its purpose (short-term decision-support using a real-time HM) but is general in its application (sociotechnical systems in need of short-term decision-support) such that it is able to create designs that are relevant to practice whilst at the same time contributing to the knowledge base. Offermann, Blom and Bub (2011) suggested that generalisability or transferability of findings occurs where settings are similar, especially when research involves social dimensions, and insights might be transferred from one to the other for similar new problems. A wide range of methods can be employed within IHAF, and application of the framework to other domains will enhance its utility and applicability. The development in simulation has shifted from purely analytical and optimisation-focused models to those which integrate simulation into decision-support tools for recurrent use. This shifts the focus from models operated by simulation experts, to tools connected to data sources and controlled or modified using user-friendly front-ends or other applications (Rodiĉ, 2017). While implementing this new simulation modelling paradigm can present challenges for researchers and organisations, particularly in sociotechnical systems with strict data governance, hierarchical social and organisational structures, and a historical resistance to change, Burger et al. (2019) positioned this paradigm shift as a challenge to the further relevance of OR as a discipline, and characterised the solution as 'Smart OR'. This emphasises the interaction and tension between humans and technology, and the tension between technical efficiency and social desirability of self-regulating systems. IHAF is able to reconcile these tensions by attending to collaborative practice, the interaction of technological and social issues, and relevant contextual information. It is expected that future research in both the healthcare and non-healthcare sociotechnical domains will use this framework for supporting the development of real-time HMs.

With a focus on implementation of simulation, Long et al. (2019) showed that the interaction of sociotechnical factors change over time. The process of development is explicated in IHAF, and opening up these issues to scrutiny can support further knowledge about barriers and opportunities for future work. As the researcher's understanding of the problem situation, contextual factors and the stakeholders deepens, the modeller may become aware of indicators signalling that appropriate shifts in the process need to be managed to maintain credibility throughout the study lifecycle. Of particular interest would be the exploration of the roles of key individuals within the research and stakeholder teams, and their interactions with the HM. The relative relationships and influences across the project process could be explored from the perspective of trust toward implementation (Harper et al., 2020). Future work could also focus on individual, team and distributed SA in complex, collaborative environments, and how the application of IHAF can support these processes. The importance of cross-disciplinary research in the application of OR data-driven methods in sociotechnical systems has been emphasised (Burger et al., 2019) and the Human Factors and OR disciplines can offer much to each other in this regard (e.g. Holman et al., 2020). For example, SA within complex collaborative systems should be viewed in its entirety, as tasks are rarely performed entirely independently of others, especially in complex situations and when critical decision-making is required (Falzer, 2018). With the increasing importance of data- and technology-driven applications, these issues are likely to become increasingly prominent.

Finally, IHAF contributes to discussions about the application of hybrid modelling, as distinct from hybrid simulation. To maximize the value that M&S can contribute toward real-world innovation, methodologies which combine methods and theories are often required. In Chapter 2, it was recognised that hybrid M&S studies combine simulation with hard or soft methods and techniques (Mustafee & Powell, 2018), not only in the model development/implementation stage of a M&S study, but to other stages in the lifecycle, for example, conceptual modelling, input and output data analysis, model verification and validation, scenario development and experimentation, and engaging with stakeholders in the implementation of the results. IHAF supports all of these applications, and future research using IHAF can contribute to the evidence for the value of HM in practice. However it also supports future work which investigates cross-disciplinary M&S studies, combining simulation with methods from other disciplines. By accessing interdisciplinary knowledge, as with the example given in the previous paragraph, cross-disciplinary HM offers unique opportunities to address challenges by leveraging the diverse body of knowledge, and individual expertise and skill-sets towards common end goals. Through innovative use of HM in practice, cross-disciplinary HM also advances M&S practice.

9.4.2 *Future work on the use-case*

IHAF proposes a methodology which can be applied to any stage of the development of a HM toward implementation. While this work is exploratory in real-time hybrid modelling, it has also provided a solid foundation for the next stage of the HM application in the use-case. Information gained from the questionnaires, application of the HM, and interviews have provided next directions. There is also scope for further analysis of the questionnaire data which may be of interest for those who study low-acuity health-seeking behaviour. Additionally, the questionnaire is exploratory, as no participants actually used real-time information to inform their attendance decisions, hence how 'perceived need' for a service translates into health behaviours and outcomes needs to be addressed through future work.

Future iterations of the HM would benefit from additional real-time data, including arrival, discharges, admissions and triage category, which can support the progression of the model from a 'static' real time model which uses a fixed set-up with real-time data, providing a snapshot of reality, to a dynamic real-time model which can adapt the structure and logic of the model as well as the data (Rodiĉ, 2017; Kritzinger et al., 2018; Onggo et al., 2020). The *prediction* stage could be improved by investigating other methods, for example accounting for other environmental information such as the weather, sporting or calendar events, ambulance data, and bed capacity. While some degree of error is expected, short-term planning may be better assisted with more reliable models. NHS England (2020b) recommend a range of advanced forecasting techniques including ARIMA/SARIMA as used in the test case. Prophet, developed by Facebook, uses Bayesian forecasting. It can account for multiple seasonalities, special events and bank holidays, can allow for missing values and outliers, and changes in historical trends. NHS England (2020b) also recommend artificial neural networks to model complex non-linear relationships between inputs and outputs. Each of these are worth investigating, as the value from short-term forecasting remains relatively unexplored for making both escalation decisions and for adaptive behaviours to improve patient flow, though simple (daily) forecasts are currently heavily relied upon for task and system-level decisions in the use-case Trust.

The DES can benefit from improvements, such as validating it for a range of scenarios and future decisions. Additionally, its interface may need attention. Using an approach similar to that proposed by Varshney (2016), which uses black box abstraction of data analytics in sociotechnical systems, can potentially reduce data and information overload, which impacts on SA and attention. Ultimately, translating the DES into open-source software such as R or Python will support its integration, adoption, testing, sustainability and spread. It will also support integration of the components, which are currently hampered by limitations in AnyLogic PLE. This is an early consideration for future work.

9.4.3 *Future work on real-time simulation for short-term decision-making in healthcare and other sociotechnical systems*

With the proliferation and availability of data, and better methods for capturing and storing data, the opportunities for real-time simulation and HM are continuing to increase. In healthcare, with demographic change, socioeconomic shifts, and an unsustainable increase in global healthcare spending, the need has been demonstrated. Internet of Things, sensors and wearables, the Cloud, 5G mobile communication, digitalisation of health records, smart mobile devices, and systems integration, are existing technologies needed for simulation to realise the vision of Industry 4.0 in healthcare, as in manufacturing with 'smart factories' and more widely. Significant challenges remain, not least storing and protecting sensitive data, and accessing data in real-time, or near real-time. The protection of the critical functionality of healthcare infrastructure and the privacy of personal data is of principal importance, compared with the manufacturing domain, where economic or structural losses cannot be compared to the massive liability of healthcare breaches (Thuemmler & Bai, 2017). Nonetheless, healthcare 4.0, a recently-emerged, collective term for data-driven digital health technologies, is expanding rapidly (Thuemmler & Bai, 2017; Jayaraman et al., 2019). Blockchain technology can provide security and transparency for simulation (Kumar et al., 2020); challenges to validation of real-time simulation are being addressed, and adjusting the structure or logic of the simulation model, and optimisation and model fidelity are being investigated (Onggo & Karatas, 2016; Oakley et al., 2020). All offer important directions for future work. Yet, despite the rapid evolution of technology and technical capability, and our increasingly reliance on it, the interaction between technology, stakeholders and the environment alongside social uncertainty and complexity, results in a continuous shift in the application of methods for generating new knowledge and new conceptualisations of the technical applications of M&S (White et al., 2016). In practice, advanced methods, increased availability and quality of data, and new challenges do not necessarily require more complex solutions. As an applied discipline focusing on real-world problems, OR distinguishes itself as a discipline in which people work with technology to gain insight and understanding. Viewing a HM study from a sociotechnical perspective opens up opportunities for knowledge production, a deeper reflection and integration of organisational needs and a clearer focus on the interface between science and practice. This requires a shift from expert practice towards a shared learning culture, where methods, role understandings, competences, interpersonal relationships and contextual factors are all directly relevant to the outcomes of the study.

This thesis argues that a systems approach is necessary for the development of real-time decision-support tools which identifies system elements, their interactions, their impacts on quality of care, and the adaptive role of people in the system. While innovating and improving the use of real-time M&S, there is a need in parallel to manage risk to ensure that the safety or satisfaction of users is not compromised. The IHAF framework provides a suitable conceptual framework for supporting these studies, alongside criteria to identify, understand, and address the interacting challenges to achieving adoption, embedding, scale-up and sustainability of HMs for short-term decision-support. Although developed for use in healthcare with a focus on ED, these factors are consistent across most sociotechnical systems. Both the framework and the application offer the possibility of being applied to other parts of the healthcare system, such as general practice or the ambulance service, or to other sociotechnical systems such as the police, transport, or social care sectors. Wider impacts, consideration of multiple stakeholders in the wider system, internal politics, how people make decisions and how they can best be supported, and how to support the adoption, implementation, and sustainability of the intervention are sociotechnical issues. Any M&S design intervention with a view to improving system functioning must contend with both system elements.

Appendix 1 Ethics Committee

UEBS
University of Exeter Business School
The University of Exeter
Streatham Court
Rennes Drive
Exeter
EX4 4PU

+44 (0)1392 722523
Web: www.exeter.ac.uk

UEBS Research Ethics Committee

Dear Alison Harper

Ethics application - eUEBS000905
A questionnaire to investigate patient decision-making regarding urgent care attendance behaviour and the impact of providing additional i

Your project has been reviewed by the UEBS Research Ethics Committee and has received a Favourable opinion.

The Committee has made the following comments about your application:
Nav Mustafee commented, Yes, I confirm

If you have received a Favourable with Conditions, Provisional or Unfavourable outcome you are required to re-submit for full review and/or confirm that committee comments have been addressed before you begin your research.

If you have any further queries, please contact your Ethics Officer.
Yours sincerely

Dr. Adrian R. Bailey

Date: 24/04/2018
UEBS Research Ethics Committee

UEBS
University of Exeter Business School
The University of Exeter
Streatham Court
Rennes Drive
Exeter
EX4 4PU

+44 (0)1392 722523
Web: www.exeter.ac.uk

UEBS Research Ethics Committee

Dear Alison Harper

Ethics application - eUEBS000905
A questionnaire to investigate patient decision-making regarding urgent care attendance behaviour and the impact of providing additional in

Your project has been reviewed by the UEBS Research Ethics Committee and has received a Favourable opinion.

The Committee has made the following comments about your application:
Dear Alison, Thanks you for the notes and files for the reapproval. I am satisfied that no f urther research ethics scrutiny is required here - your information sheet makes clear the non-sensitive nature of the proposed interviews. In future, however, it would be better to provide an interview schedul e when making a submission like this. Some recommendations/suggestions are to: A [info sheet]: "The transcriptions are stored on an encrypted laptop, and will be held for twelve months." It
- Please view your application at https://eethics.exeter.ac.uk/UEBS/ to see comments in full.

If you have received a Favourable with Conditions, Provisional or Unfavourable outcome you are required to re-submit for full review and/or confirm that committee comments have been addressed before you begin your research.
ıse contact your Ethics Officer.

A. R Bailey

Dr. Adrian R. Bailey

Date: 03/09/2020
UEBS Research Ethics Committee

Appendix 2a: Patient Questionnaire

CONSENT FOR SURVEY RESEARCH

SURVEY ON NHSquicker AND URGENT HEALTH CARE TREATMENT

NHSquicker is a free app available for both Google and Apple. It aims to assist you to make a decision about the best place to attend for urgent medical care. It contains information about the nearest healthcare facilities, including GPs, dentists and pharmacies, with real-time wait-time data for the nearest Emergency Department and Minor Injury Units. It also includes travel times.

Many factors will influence your decision about where to go for treatment. We are also interested in evaluating the effect of providing you with wait-time information on your decisions about where to attend for treatment. We greatly appreciate your decision to help us with this study which supports the delivery of urgent healthcare.

This survey is entirely anonymous and all of your answers will be treated confidentially. We will use this survey to learn more about the impact of NHSquicker on waiting times. You need not have used NHSquicker in order to participate.

For further information please contact:

Alison Harper

ah596@exeter.ac.uk

+44(0) 7922 109 779

Consent

I have been fully informed about the aims and purposes of the project.
I understand that:

- There is no compulsion for me to participate in this survey and, if I do choose to participate, I may withdraw at any stage;
- Refusal to participate will have no impact on my medical treatment;
- I have the right to refuse permission for the publication of any information about me;
- Any information which I give will be used solely for the purposes of this project, which may include publications or academic conference or seminar presentations;
- If applicable, the information, which I give may be shared between any of the other researcher(s) participating in this project in an anonymised form;
- All information I give will be treated as confidential;
- The researcher(s) will make every effort to preserve my anonymity.

....
....
(Signature of participant) (Date)

……………………………………………….
(Printed name of participant)

NHSquicker QUESTIONNAIRE

PART 1: Tell us about yourself

1. **Are you the patient, or answering on behalf of a patient?**
- ☐ Patient
- ☐ Answering on behalf of the patient

2. **What is the age of the person seeking treatment?**
- ☐ Under 5
- ☐ 6-18
- ☐ 18-35
- ☐ 35-50
- ☐ 50-70
- ☐ 70+

3. **Gender**
- ☐ Male
- ☐ Female
- ☐ Self-describe_____

4. **How would you rate your usual health?**
☐ Poor 1 2 3 4 5 Very good

5. **How well do you know the local area?**
Poor 1 2 3 4 5 Very good

6. **Where were you when you decided to seek medical help?**
First three digits of postcode_______________ OR Location ___________________

7. **How did you get here today?**
- ☐ Personal vehicle
- ☐ Public transport or taxi
- ☐ Lift with friend
- ☐ Ambulance/non-emergency patient transport
- ☐ Walk/other

8. **Did you rely on navigation support to get here (e.g. paper map, digital maps on smartphones, car satnav, directions from friends)?**
- ☐ Yes
- ☐ No

9. **Was somebody with you when you decided to seek treatment?**
- ☐ Yes
- ☐ No

Part 2: Tell us about your visit today

10. **How urgent do you feel your condition may be? (I.e. how quickly do you need to be seen)**
Not urgent 1 2 3 4 5 Very urgent

11. **How serious do you think your condition is? (I.e. the severity of your condition)**
Not serious 1 2 3 4 5 Very serious

12. **Please tick ALL urgent care services that you are aware of:**
- ☐ Emergency department (A&E)
- ☐ Urgent care centre
- ☐ Minor injury units
- ☐ Walk-in centres

- ☐ GP and/or GP out-of-hours
- ☐ NHS 111
- ☐ Pharmacy
- ☐ Other e.g. dentist, optician

13. Did any of the above services advise you to attend here today?

☐ Yes ______________________ ☐ No

Please specify

14. We are interested in <u>why you came here today</u>, instead of other places which might help (depending upon your condition). Please tick ALL the boxes that apply to you:

OTHER SERVICES:

- ☐ This is the closest health service for me today
- ☐ I'm not sure where else I could have gone
- ☐ I don't know where to find other services
- ☐ I'm not sure if other services are right for me today
- ☐ I don't know what other services are open
- ☐ I'm more likely to have to wait at other services
- ☐ Nowhere else has 24 hour access
- ☐ Another service would have referred me here anyway
- ☐ This service is the nearest to me
- ☐ I think I'll be seen more quickly here
- ☐ It's easier to get here than to another service

GP:

- ☐ My GP wasn't available
- ☐ I don't know if my GP was available
- ☐ I didn't want to see my GP
- ☐ I would have to wait for a GP appointment
- ☐ I don't have a GP

THIS SERVICE:

- ☐ This is the most appropriate health service for me today
- ☐ I've used this service before and was satisfied
- ☐ I'm confident in this service

IF IN A&E:

- ☐ I wanted to see a specialist
- ☐ My condition is an emergency
- ☐ I want to see a doctor as soon as possible
- ☐ I might need to go to hospital
- ☐ I need reassuring that it's not serious

15. Which of the following helped you to choose where to go for treatment today? Please tick all the boxes that apply

- ☐ The NHSquicker app/website
- ☐ Advice from health care professional
- ☐ Advice from friends/family/other
- ☐ NHS Choices/internet re my condition
- ☐ Internet search about services nearby
- ☐ Personal knowledge
- ☐ Personal instinct
- ☐ Previous experience
- ☐ Nearness to my location
- ☐ Concern about parking

- ☐ Ease of accessibility by public transport
- ☐ The weather
- ☐ Transport options available to me
- ☐ Other_______________

It can be very confusing knowing where to go for treatment. Please consider the following information from NHS England:

Call NHS 111
If you urgently need medical help or advice but it is not a life-threatening situation

Call 999
If someone is seriously ill or injured and their life is at risk

Visit A&E
For serious injuries or a genuine life-threatening emergency

Walk-in Centre or Urgent Care Centre
For a minor illness or injury at times you can't wait to see your GP

Ask your pharmacist
For advice for many common minor illnesses

See your GP
Make an appointment if you are feeling unwell and it is not an emergency

With reference to this information, please answer the following two questions:

16. You have the right to decide which service to use. However, given this information, please consider how certain you are that this is the most appropriate place for you today:

Not certain 1 2 3 4 5 Very certain

17. What other services do you consider might be appropriate for your treatment today?
Please tick all the boxes that apply:

- ☐ Emergency department (A&E)
- ☐ Urgent care centre
- ☐ Minor injury units
- ☐ Walk-in centres
- ☐ GP and/or GP Out-of-Hours
- ☐ NHS 111
- ☐ Pharmacy
- ☐ Other_______________

18. Did you use the NHSquicker app before coming here today?

- ☐ Yes – please complete **Qu. 19-20**
- ☐ No – please complete **Qu. 21- 23**

If YES

19. How was it useful? Please tick all the boxes that apply:

- ☐ I found healthcare services near to me
- ☐ Gained knowledge about my condition
- ☐ Became aware of opening hours/facilities

- ☐ Found out how busy services are near to me
- ☐ I found a more appropriate service
- ☐ How to travel here today
- ☐ Other ______________________

If YES

20. In what way did your decision change? Please tick all the boxes that apply:

- ☐ Would have gone to a different location
- ☐ Would have gone at a different time of day
- ☐ Would have used a different form of transport
- ☐ My decision would not have been different
- ☐ Other ____________________

If NO

21. Have you seen or used the NHSquicker app on another occasion?

- ☐ Yes
- ☐ No

If NO

22. Do you think you would have found NHSquicker useful today?

- ☐ Yes ______________ (Please specify)
- ☐ No
- ☐ Unsure

If NO

23. Do you think NHSquicker may be useful for you at a later time?

- ☐ Yes ______________ (Please specify)
- ☐ No
- ☐ Unsure

Part 3: What do you think of NHSquicker?

24. Why might it be useful for you or your friends/family to know about current waiting times at A&E and other urgent care services near to you?

e.g.

Facility Name	Current waiting time – 10:00am
Torbay A&E	2 hour
Newton Abbott	1 hour 30 minutes

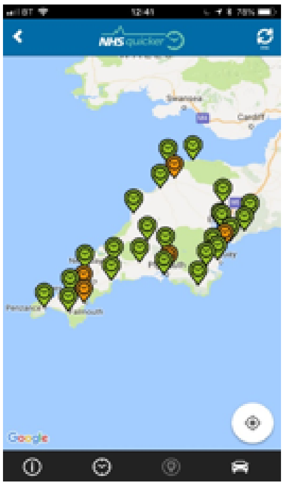
25. How useful would it be to you or your friends/family if you could have predicted waiting times for the next few hours?

e.g.

Facility Name	Current waiting time – 10:00am	Predicted waiting time - 11 am	Predicted waiting time - 12 noon
Torbay A&E	2 hour	2 hours 20 minutes	3 hours
Newton Abbott	1 hour 30 minutes	1 hour	45 minutes

Please answer here:

26. **NHSquicker is meant for <u>those seeking urgent care</u>. However can you think how the <u>NHS</u> could make use of this information in Devon & Cornwall?**



Use of current wait times:

Use of predicted wait times:

Other uses of this information:

27. **Would you recommend NHSquicker to your friends/family?**

☐ Yes ☐ No ________________ ☐ Unsure

Please specify

-----------THANK YOU FOR YOUR VALUED PARTICIPATION -----------

RESEARCHER USE ONLY

Facility [____________________] Time [______________________]

Date [______________________] Researcher [_________________]

Appendix 2b: Field Notes (observational data)

A2.1 3rd IMPACT network event 21 June 2016, UEBS

A2.1.1 Introduction

This collaborative workshop was delivered by the IMPACT Network, a collaboration between the UEBS and NHS Trusts in the south-west of England, with a focus on urgent and emergency care. The purpose of the workshop was toward the co-production of the *NHSquicker* project between UEBS and NHS Trusts in Devon and Cornwall.

Participants were subdivided into the following categories:

(i) NHS Manager (n=11): (Mx)
(ii) NHS IT and information staff (n=15): (IT)
(iii) NHS communications/marketing (n=4) (Comms)
(iv) NHS/SWAS clinical staff (n=4) (Clin)
(v) Patient representatives (n=3) (Pt)
(vi) UEBS academic staff (n=9) (Acad)
(vii) Other eg developers (n=2) (R15)

Handwritten notes were taken throughout the day by myself. These were categorised thematically and summarised by participant group. The majority of sessions were also video and/or audio-recorded, however a subset of audio data showed that the handwritten notes were sufficiently comprehensive. All participants signed consent forms for data recording throughout the day. The notes are summarised below. Points of relevance to this thesis are highlighted in yellow. Raw data for all analysis is available upon request.

A2.1.2 Thematic Summary of field notes

Issues identified from workshops () = category of stakeholder

- **Use of MIU:**

-MIU use looks like increasing in ED and MIU but ?cause as multivariate (IT, Mx)
-Care interpreting – are increased numbers coming from ED or other parts of urgent care system? Eg MIU closures, GP capacity (Mx)
-Clarity of MIU offering re appropriate attendance (Clinician, Mx, Mx)
-Xray opening times not necessarily easy to find (IT)
-Risk of inappropriate attendance if focus on low waits only eg in labour (Mx, clinician)

- **Use of ED:**

-Knowledge needed about how people (non-urgent) use ED (Mx, Mx, Mx)
-Won't reuse app if 'sent' to wrong place (Clin)
-Misconceptions re better care in ED (Mx)
-What basing decision on? Closest, fastest, or perceived care? Order is critical (Mx, Mx, Clin)
-Clarity of purpose re how choice is made (Clinical)
-'Hospital at home' use – if need urgent care, may need it more urgently (patient)
-Engage with patients – what matters?:

Whether to go, when to go, where to go (Mx, Mx)
50% Cat 4; average age 36 years (Mx, Mx)
Are they 'repeat offenders?' (Mx) If so, why and what do they want? (Mx)
Engage with MIU users directly – redirected from ED? (Clinician)
Patient focus groups, different categories, what is value (academic, comms)
Students; parents; holidaymakers; sports; schools; GP/pharmacists; staff groups; migrants; seasonal workers; homeless; care homes/N Homes; mental health; frail/elderly; carers
30% don't need to be there – who determines? ?assumptions (clinical)
Disproportionate presentations from radius of ED geographically (Mx)
Support rare/occasional/frequent users (Mx)
Smartphone/website access:
Lower use of smartphones in older age groups (but may be higher ED users) (patient, academic)
App overload (patient, IT, Mx)
Large number of website hits (IT)
What is motivation to download/USP of app (Mx)
Concerns by users about personal data collection (IT)
Connectivity concerns eg Dartmoor (IT); Need to enable (IT)

-Decision support based on presenting condition/clinical advice:
Concern re risk (Mx, IT, Mx, clinical) and re replication of other apps (Mx, Mx)
Support for benefits eg selfhelp video (Mx, clinician, patient, Mx, Mx, IT, Mx, Mx)
-Health and care self-care videos – clinical sign off (Hugh Kelly)
-Decision support with pharmacy, WICs, Dentists, GP, 111 advice:
-Necessary eg dos and above (IT, Mx, pt, Mx, Mx, academics)
-Who to attend eg mental health support – pt needs signposting (clinician)
-Link into other apps eg HandyApp, UofE app, to create a 'network' (Mx, Mx, academic)
-Simple/light for emergency use; balance simple and informative (acad, Mx, Mx, Mx)
-Simple visually on small screen (Mx, R15)
-Difference between 'providing information' and 'facilitating a decision':
Right choice, right reason (clinician, Mx, Mx)
Engage with senior clinical staff re risk, wording (Mx, IT/Mx, Mx)
Clear that decision remains with patient – just adding info (Mx, Mx, comms)
Ensure patient-centric, not provider centric eg acronyms, purpose, lay terminology (academic, Mx)
-Data/information:

Incorporating historical data to increase accuracy (IT, IT, Mx, Mx)
If behaviour changes, historical data may become inaccurate (IT)
Data feed accuracy concerns (IT, IT)

Capturing correct waiting time measure (IT)
Unable to capture other factors eg staff down, other 'human interaction' (IT)
Capturing minors only – not reflecting activity in rest of dept (Mx)
Incorporate 'Stress status' in ED as per website (IT)
Senior Mx support: Retain historical info if needed by senior Mx (IT)
Google translate (IT)
Why not Exmouth, Tiverton? (patient)
Autocall 111, 999 – including when out of range/data allowance? (Mx, Pt)

Table A2b-1 counts the number of references by participant type. These should be interpreted with care, as the workshop was a group activity, and so repetition was not expected however it provides an overview of important topics.

Table A2b 1 Summary of count of findings by theme, and by participant

	Patient rep	Manager	Clinician	Comms	IT
Use of MIU		5	1		1
Use of ED	1	6	3		
What matter to patients		10	2	1	4
Decision support	1	26	2	1	
Data/information	2	5			10

A2.1.3 Breakout session

A one-hour breakout session was held with the participants in groups of 4-6, using post-it notes and flipcharts. The following structure was followed:

A) Consider **on your own** re NHSquicker:
  i) What enthuses you?
  ii) What concerns do you have?
  iii) What improvements can you suggest?
B) In **groups**, discuss and record your ideas on a flipchart
  iv) Comments about the design and navigation
  v) Comments about the logo
  vi) Ideas for future content
C) Tech session (IT and developers only – not included in summary report)

A2.1.3.1 Breakout session categories

The data from the flipcharts and post-it notes are categorised below:

(i) What enthuses you?

Joint working between providers – integration
Live, location-based data
Empowering and informing patients of choices
Appropriate ED patients will receive better service

Patients – feel more empowered, self-treating, time-saving, faster treatments, reduced inconvenience
ED reduction in volume
Better resource utilisation
Making useful information available when/where/how people want it
Empowering for patients
Enabling patients to make own choices
Educates service users
Provides local up-to-date info
Open data
Simple to use
Continues the 'health economy' work across the patch
Joint working
Easy/simple to use and clear to understand, uses minimal storage
One app for all
Immense potential
Reduce pressure on EDs
Redirect to MIUs
Reduce anxiety
Urgent capacity vs number of patients
Patient focussed, really useful for patients
Choose right, wait less
As a patient – helps direct me to the closest place
Concept of one system across the NHS
Helps meet 4 hour performance target

(ii) **What concerns you?**

Risk of suboptimal outcome – clinical risk
Care with working – patient making decision about attendance
Responsibility and liability – make it right
Broad audience – health pyramid/health vs tech ability/access vs utility
Triage to get to right location? British Red Cross triage category?
Students as target audience – how to travel?
What value does app give over 'googling'
Light touch signposting
Omitting pharmacy and WIC misses important alternatives
'Emergency care' – not everything is an emergency – perceptions/knowledge/emotion
Not all public know what MIU/ED mean
Do we know enough about end-user behaviour to drive design?
Danger of not all options covered by app
Danger of all facilities not completely aligned with regards to metrics/frequency/log
Too much info, overcomplicating navigation with features
We don't know who/how/when/why – ASK
How do we overcome fear of change/exposure in NHS?
A very long list of hospitals to scroll through
Short films for self-treatment , clear instructions – good idea but concerned re risk around diagnosis
How to enforce people to keep the app on their phone

Governance/risk
Abbreviations used – people don't understand, eg ED
Lack of patient/user feedback or involvement
Unregulated websites/advice
What is the hook in for using the app?
Concern about NHS overcoming fear
Too much information
Making sure it runs smoothly and doesn't cause confusion
Different levels of service/ opening times

(iii) **What improvements do you suggest?**

Definitions re ED/MIU/UCC/WIC not all consistent
Include pharmacies near you
Use lay terminology – influencing public behaviour/choice
Should there be some science about behavioural/cultural change to help inform?
Student feedback
111/999 well known – so should this be a number if aiming for national rollout?
Short, snappy name
Evaluate – feedback forms, wait times, return visits, patient experience, patient stories, 4 hour target data, distribution of patients, trust awards for reaching targets, testimonies
Links to other apps
Capacity/ability to be seen rather than patient numbers (which need to be seen in context)
Show how stressed the unit is
Open data – reflect: time to absorb and be clear about purpose and communications
Potential link with other apps, websites
Diagnostic capacity

(iv) **Design and Navigation**

Auto-call
Concise and simple design
Patient/user choice re where to go (self-assess urgency)
As many sites/facilities as possible listed
Predictions – in one hour
Track location and radius pre-set – 2 routes?
Simple and light design
Keep purpose clear: reduce ED/MIU demand and redistribute load
Open app - ?Decision prompt – do I need to attend now? Continue button OR
Straight to list of providers aided by location – nearest and shortest wait time
Include self-help within the app – either replicate or make available links to 111/ChooseWell/NHSchoices/H&C videos
Under 'more information ' – definition about what an MIU can do/see
Some MIUs close early due to operational pressure – app may be misleading if not updated
If focus is on waiting times – is this patient centred? Patients will still have to wait.

Unintended consequence of empty EDs and MIUs underperforming against 4 hour target
Replicate 999 format/choices – What are my choices, what is near me?
Geography – some areas may still be 50 miles away, eg North Cornwall, so consider distance radius
Signpost to all services – pharmacy, GP, self-help
Include 111
Link to pharmacy, dentist, GP, MIC, UCC, ED
Tab between nearest and shortest waiting time on app
Waiting times – Do you need to attend an ED or MIU as initial question?
Initial question – do you think your condition is life-threatening? Call 999/111/ED.
Have you had an accident? Do you have a known medical condition?
Balance between simple/easy and more info eg clinical support
Be clear about app boundary – keep it simple
Something now better than perfect never – NHS fear of action
111 call button
Likert scale feedback within app
Include pharmacy/GP/MIU/ED/WIC – filter to expand to show additional services?
Feedback within app
Simple design with nudge
Proxy for experience – wait/stress
Need more – pharmacies, GPs, 111 etc
Pharmacies/GPs – info on services possible?
Link to 111
First click viewed as life-threatening

A2.2 Qualitative System Dynamics Workshops 18 July 2018, 27/28 June 2019.

NHSquicker QSD Workshops 18th July 2018 Notes Summary

Participants (n=5)

The aim was to capture the decision variables of users of the real-time data, from the perspective of staff (the collective preconceptions of the group members as to the components of decision). The workshop was led by an experienced QSD facilitator. The result was a causal loop diagram which conceptualised the decision mechanisms of a user.

Patient anxiety emerged as central in this workshop. The outcomes identified the following mechanisms:

- Anxiety increases the perceived need for treatment, the perceived severity of the condition, and in turn further increases anxiety levels.
- Confidence in the attendance choice reduces anxiety. Confidence in waiting time prediction increases knowledge about attendance choice.
- Access to information reduces risk aversion and reduces anxiety.

A subsequent QSD workshop investigated the effects of patient attendance decisions on the NHS. Here, the effect of reducing anxiety was seen to reduce the fraction of inappropriate attendances and improve the patient quality experience. This triangulates with the patient questionnaire findings.

A2.3 Questionnaire raw data

Questionnaire data was coded in Excel for analysis. A snapshot is included below in Figure A2b-1 of a subsection of the data.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	RespID	Locatio	Date	Shift	Respon	AgeGro	Gender	UsualH	LocalK	Postcod	Transp	DigNav	Accomp	Urgenc	Serious	AwareED	Av
2	1	1	29/07/2018	3	2	2	1	5	3	EX1	3	2	1	4	3	1	
3	2	1	29/07/2018	3	2	3	2	4	4	EX16	4	2	1	1	2	1	
4	3	1	29/07/2018	3	2	5	2	3	4	EX14	1	2	1	3	3	1	
5	4	1	29/07/2018	3	2	3	2	5	5	EX5	1	2	1	3	2	1	
6	5	1	29/07/2018	3	2	3	1	3	5	EX1	1	2	1	4	4	1	
7	6	1	29/07/2018	3	2	4	2	1	2	EX8	3	2	1	1	1	1	
8	7	1	29/07/2018	3	2	3	1	4	5	EX1	1	2	1	4	4	1	
9	8	1	29/07/2018	3	1	4	1	5	5	EX6	1	2	1	3	3	1	
10	9	1	29/07/2018	3	1	5	1	5	5	EX8	2	2	1	2	1	2	
11	10	1	29/07/2018	3	2	3	1	5	5	EX8	1	2	1	4	4	1	
12	11	2	29/07/2018	3	1	6	2	4	1	EX2	2	2	2	1	1	1	
13	12	2	29/07/2018	3	1	5	2	5	2	EX1	1	1	1	1	1	1	
14	13	1	29/07/2018	3	1	4	2	5	5	EX15	1	2	1	3	3	1	
15	14	1	22/07/2018	3	2	5	2	5	5	EX8	1	2	1	3	3	1	
16	15	1	22/07/2018	3	1	3	2	5	5	EX1	1	2	1	3	2	1	
17	16	1	22/07/2018	3	2	1	2	5	1	EX5	1	1	1	2	3	1	
18	17	1	22/07/2018	3	1	3	1	5	5	EX4	1	2	1	3	3	1	
19	18	1	22/07/2018	3	1	5	2	4	4	EX1	3	2	1	1	1	1	
20	19	1	22/07/2018	3	2	4	2	5	5	EX4	1	2	1	3	3	1	
21	20	1	22/07/2018	3	2	3	1	3	5	TA2	1	2	1	2	1	1	
22	21	1	22/07/2018	3	1	4	2	5	5	EX17	3	2	1	3	3	1	
23	22	1	22/07/2018	3	1	4	1	5	5	EX4	1	2	1	3	3	1	
24	23	1	22/07/2018	3	2	3	1	5	3	EX9	1	2	1	5	5	1	
25	24	1	22/07/2018	3	1	6	1	3	5	TQ14	1	2	1	3	3	2	
26	25	1	22/07/2018	3	1	6	2	3	5	TQ14	1	2	1	4	4	1	
27	26	1	22/07/2018	3	1	3	2	5	5	EX4	3	2	1	1	1	1	
28	27	1	22/07/2018	3	2	2	1	5	5	EX2	1	2	1	3	2	1	
29	28	1	22/07/2018	3	1	3	1	5	4	EX1	1	2	2	2	2	1	
30	29	1	22/07/2018	3	1	4	1	5	4	EX5	1	2	1	4	4	1	
31	30	1	22/07/2018	3	1	3	1	5	5	EX1	5	2	2	2	2	1	
32	31	1	06/07/2018	3	1	3	1	4	5	EX4	5	1	2	3	3	1	
33	32	1	06/07/2018	3	1	5	1	5	5	EX2	3	2	1	5	5	1	
34	33	1	06/07/2018	3	1	4	2	5	5	EX16	1	2	1	2	2	1	
35	34	1	06/07/2018	3	2	4	2	4	4	EX1	5	2	1	4	3	1	
36	35	1	06/07/2018	3	2	3	1	5	3	EX2	1	2	1	4	4	1	
37	36	1	06/07/2018	3	2	2	1	1	5	EX4	3	2	1	3	2	1	
38	37	1	06/07/2018	3	1	5	2	4	3	TQ13	3	2	1	4	4	1	
39	38	1	06/07/2018	3	2	1	2	4	3	EX1	1	1	1	3	3	1	
40	39	1	06/07/2018	3	2	6	2	3	4	EX4	2	2	2	3	2	1	
41	40	1	06/07/2018	3	1	4	1	4	4	EX1	1	2	2	1	2	1	
42	41	1	31/07/2018	4	2	2	1	5	5	EX1	1	2	1	3	3	1	
43	42	1	31/07/2018	4	2	3	2	5	4	PL19	1	1	1	3	3	1	
44	43	1	31/07/2018	4	2	3	2	4	1	EX1	1	2	1	5	5	1	
45	44	1	31/07/2018	4	1	5	1	4	5	EX6	2	2	2	5	4	1	
46	45	1	31/07/2018	4	1	4	2	4	3	EX2	1	2	1	3	2	1	
47	46	2	31/07/2018	4	1	5	1	2	1	TQ1	1	2	1	4	4	1	
48	47	1	31/07/2018	4	2	1	1	5	5	EX20	1	2	1	4	4	1	
49	48	1	07/08/2018	4	1	5	1	4	3	EX3	1	2	1	4	3	1	
50	49	1	07/08/2018	4	2	5	1	2	1	EX1	1	1	1	5	3	1	
51	50	1	07/08/2018	4	2	4	1	5	5	EX6	1	2	1	2	2	1	
52	51	1	07/08/2018	4	2	1	1	2	2	EX2	1	1	1	4	3	1	
53	52	2	07/08/2018	4	1	3	1	5	5	EX1	1	2	2	3	3	1	
54	53	2	07/08/2018	4	1	3	2	4	5	EX2	5	2	2	2	2	1	
55	54	2	07/08/2018	4	2	3	1	5	1	EX4	3	2	1	2	2	1	
56	55	2	07/08/2018	4	2	3	1	4	5	EX4	1	2	1	3	2	1	

Figure A2b-1 Snapshot of subsection of patient questionnaire data

Similarly, the open data was recorded in Excel by respondent ID and question number (Figure A2b-2). Responses were colour-coded for initial analysis. Subsequently, both sets of data were analysed in NVivo 12.

RespID	Qu.Number	Question wording	Response
		Waiting times	
		Planning	eg childcare, lifts, timing
		Demand/capacity	NHS side; patient side
		Parking	
		Would not have changed mind	
		Providing info	
		Would have gone elsewhere/changed mind	
		Predicted wait time usefulness	
		Other	Other suggestions eg wic or gp; reduce unnecessary use; manage expectations; reduce anxiety
	26	NHS Use	Help with parking
			Help with people picking you up
			Can help holiday-makers in our pub to find an open hospital/other facilities and where
	24	Current Wait Times	In emergency, I would likely choose the shortest wating time
	25	Predicted Wait Times	Not sure
	26	NHS Use	There always a choice between Exeter, Taunton and Dorchester for someone resident in W Dorset - would help.
108	22	Useful today?	Unsure - not busy here today (TSD ED)
109	24	Current Wait Times	To help us decide where to go
	25	Predicted Wait Times	Very useful to know waiting times
110	24	Current Wait Times	Don't use a smartphone, don't know what an app is. I work at a caravan park and have a list of hospital info but feel that our generation wouldn't use this
111	22	Useful today?	Yes, different location (here on holiday)
	25	Predicted Wait Times	Maybe would have gone to a different hospital with shorter waiting times
112	22	Useful today?	Might have been seen quicker elsewhere
	23	Useful in future	Yes, finding where to go
	24	Current Wait Times	So we don't waste anybodies time by being in the wrong place
	25	Predicted Wait Times	We could leave the house closer to the time of the appointment
	26	NHS Use	They could give patients some idea when they will be seen
			If somebody needs to be seen sooner than the predicted wait time they can can act before it is too late
			Could advise goining elsewhere (somewhere more suitable)
114	24	Current Wait Times	I don't have a smart phone but would be useful for my daughter
	25	Predicted Wait Times	Not very useful for me as I live within walking distance of the hospital
115	22	Useful today?	No - was told to come to specific department (labour ward)
	24	Current Wait Times	Save a long wait
			Direct to the most appropriate service available
	25	Predicted Wait Times	Very handy if multiple units provide appropriate service needed to situation
116	24	Current Wait Times	General info for relatives collect and drop off
			Timing of attendance

Figure A2b-2 Snapshot of subsection of open data from patient questionnaires

A2.4 NVivo 12 cross-tabs

NVivo has the capability of cross-tabulating open and closed data sets. These were exported to Excel. A large number of cross-tabs were generated. A sample is shown below in Table 2b-2.

Table A2b-2 Sample cross-tab open-coded data with two closed variables: Anxiety and Uncertainty re alternatives (%)

Closed data	anxiety=1	%	anxiety = 2	%
Open Data	NewUnsureAlternatives = 1 (62)	NewUnsureAlternatives = 2 (27)	NewUnsureAlternatives = 1 (16)	NewUnsureAlternatives = 2 (47)
1 : Balance demand and capa	14.52	33.33	25.00	12.77
2 : Staffing	6.45	11.11	12.50	8.51
3 : Information other than wa	14.52	29.63	31.25	19.15
4 : Other uses of RT data	19.35	7.41	6.25	8.51
5 : Save time and travel	9.68	0.00	6.25	6.38
6 : Wait time knowledge is us	72.58	70.37	81.25	59.57
7 : Consider if appropriate to	4.84	14.81	6.25	10.64
8 : Planning	25.81	33.33	12.50	21.28
9 : Predicted wait times are u	27.42	25.93	25.00	27.66
10 : Reduce anxiety	6.45	7.41	6.25	0.00
11 : To manage expectations	9.68	7.41	12.50	4.26
12 : When to go	9.68	3.70	18.75	12.77
13 : Where to go	33.87	51.85	43.75	31.91
14 : Would not have changed	16.13	7.41	37.50	23.40
Useful in future	32.26	29.63	31.25	23.40
15 : Total (unique)	51	22	14	34

A2.5 SPSS analysis

SPSS was used to analyse the closed data. An example of the output is below in Tables A2b-3/4 and Figure A2b-3. The output shows a simple 2*2 cross-tab for results of those who were referred to ED and those who needed reassurance regarding the severity of their condition. For this output, n=135, as the remainder of questionnaire participants were in a MIU or WIC. The results of chi-square analysis shows that there is a significance (2-sided) of p<0.25. This means that there is a significant difference between those who are referred, and those who are not, and the anxiety measure of needing reassurance, for those attending ED. Figure A2b-3 depicts the results in a bar chart. Those who are referred are less likely to be seeking reassurance as they have already been given that reassurance by being referred to ED.

Table A2b-3/4 SPSS sample output

Refererred * ReassuranceReSeverity Crosstabulation[a]

Count

		ReassuranceReSeverity		
		Yes	No	Total
Refererred	Yes	13	55	68
	No	25	42	67
Total		38	97	135

a. Location = RD&E ED or TSD ED

Chi-Square Tests[a]

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.525[b]	1	.019		
Continuity Correction[c]	4.662	1	.031		
Likelihood Ratio	5.595	1	.018		
Fisher's Exact Test				.022	.015
Linear-by-Linear Association	5.484	1	.019		
N of Valid Cases	135				

a. Location = RD&E ED or TSD ED

b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 18.86.

c. Computed only for a 2x2 table

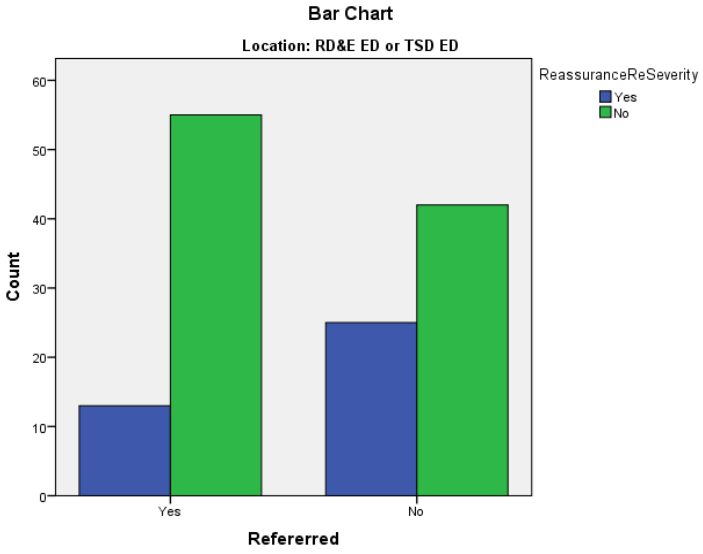


Figure A2b-3 Bar Chart for cross-tab Severity and need for reassurance

Appendix 3: Hybrid Model

A3.1 Hybrid model components of IHAF

A3.1.1 ED Data

A3.1.1.1 Available ED datasets

Table A3.1 ED datasets from the use-case and NHSquicker

Data set	Dates	Fields	Descriptions
NHSquicker (30 minutes)	3/01/2018 – 3/01/2019	Total Patients in ED Patients Waiting for first treatment Maximum Wait time in ED	Fields are available in near real-time for ED and MIUs.
AE Performance report (Daily)	1/04/16 - 19/09/18	day_of_week arrival_date total_attenders attends_2_hrs_or_less attends_4_hrs_or_less attends_over_4_hrs bed_delays gp_refs ae_admit ip_admissions discharges daily_4hr_compliance cummulative_4hr_compliance trolley_wait_between_4_to_12_hours trolley_wait_over_12_hours first_assess_within_15_mins	Daily reporting data of total daily attendance, wait duration category, number of patients with delayed admission, and 4 hour compliance against the target. Trolley wait category (waiting for admission) and time to triage (within 15 minutes) are also reported.
ED arrivals (per patient)	1/04/16 - 29/10/18	attendance_number patient_age sex postcode site_description arrival_date_time Weekday triage_category_description attendance_disposal_description diagnosis_01_description	Patient arrivals by date and time, age, gender, site, triage category, discharge type, diagnostic category
ED dataset (per patient)	1/08/15 – 31/07/16	Site department_type attendance_reason arrival_date at reception	Site (in urgent care network), arrival date and time, diagnosis,

		arrival_time at reception examination_date_time to see a doctor first_ews_date_time triage_date_time triage_category_description initial_assessment_date_time date_time_seen_by_clinician ae_time_seen_for_treatment visit_time_in_minutes (total time) visit_duration decision_date_time discharge_date_clockstops discharge_time left_department_date left_department_time attendance_category_description care_group_description attendance_disposal_description source_of_referral_description discharge_destination_description reason_for_delay_discharge_description diagnosis_01_description diagnosis_02_description diagnosis_03_description ae_investigation_1_description ae_investigation_2_description ae_investigation_3_description ae_treatment_1_description ae_treatment_2_description ae_treatment_3_description admission_date_time admitting_specialty admitting_specialty_description admission_source_description locations	examination date-time, triage date-time, triage category, treatment duration, total length of stay, number and type of treatments, number and type of investigations, discharge date-time, left department date-time, discharge type (e.g. admission, home, leave without treatment, died), reasons for delays in discharge or admission (e.g. waiting for bed, waiting for specialist review).

A3.1.1.2 NHSquicker data and ED attendance data (expanded from Section 6.4.2)

```
Call:
lm(formula = Patients.Waiting ~ Total.Patients)

Residuals:
    Min      1Q  Median      3Q     Max 
-7.3109 -1.5635 -0.1137  1.3415 19.4813 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)    -1.836196   0.092342  -19.89   <2e-16 ***
Total.Patients  0.207888   0.003065   67.82   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.398 on 5497 degrees of freedom
Multiple R-squared:  0.4555,    Adjusted R-squared:  0.4554 
F-statistic:  4599 on 1 and 5497 DF,  p-value: < 2.2e-16
```

Figure A3.1 Linear model output of Patients Waiting and Total Patients

```
> hist(lm.residuals, main = "Histogram of Total Patients residuals", col = "#00AFBB")
> qqnorm(lm.residuals, main = "Normal Probability Plot")
> qqline(lm.residuals, col="red")
> summary(lm.residuals)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
-7.3109 -1.5635 -0.1137  0.0000  1.3415 19.4813 
```

Figure A3.2 Residuals of Linear model Patients Waiting and Total Patients

Figures A3.1 and A3.2 are the results of a linear model of *Total Patients* and *Patients waiting*. The structure of the scatterplot of *Total Patients* and *Patients Waiting* indicates departures from the standard regression assumptions, with a 'floor' effect at zero (no less than zero patients can be waiting) lowering the mean of the residuals to zero, despite having a long positive tale. A residual plot that has a "fan shape" indicates a heterogeneous variance (non-constant variance). The residuals tend to fan out as error variance increases (Figure A3.3).

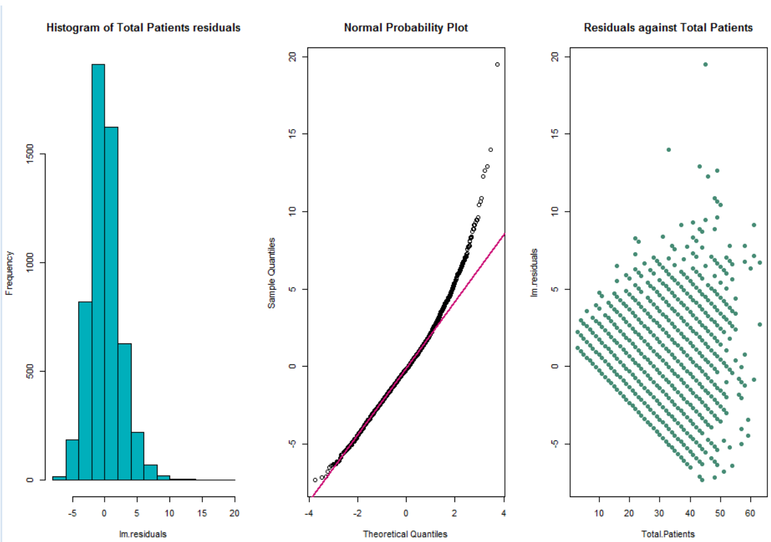


Figure A3.3 Plots of the residuals for normality and homoscedasticity

In the main scatterplot (Figure 6.11 in Chapter 6), a number of outliers can be seen in the top-right quartile, where the number of patients waiting to be seen has significantly departed from the line of best fit. This is confirmed in the scatterplot of the residuals against the predictor (*Total Patients*), where the variance in the residuals increases as the number of patients in the department increases (Figure 6.11). This effect can be explained by the busyness in the department. As the total number of patients in the department rises and the demand-capacity mismatch increases, a build-up of low acuity patients in the waiting area can occur if patients of higher urgency are present.

The following plots are scatterplots of *Patients Waiting* and *Total Patients* (and the residuals) with a view to examining where the variance begins to increase (Section 6.4.2), suggesting an appropriate trigger is between 40 and 45 *Total Patients*.

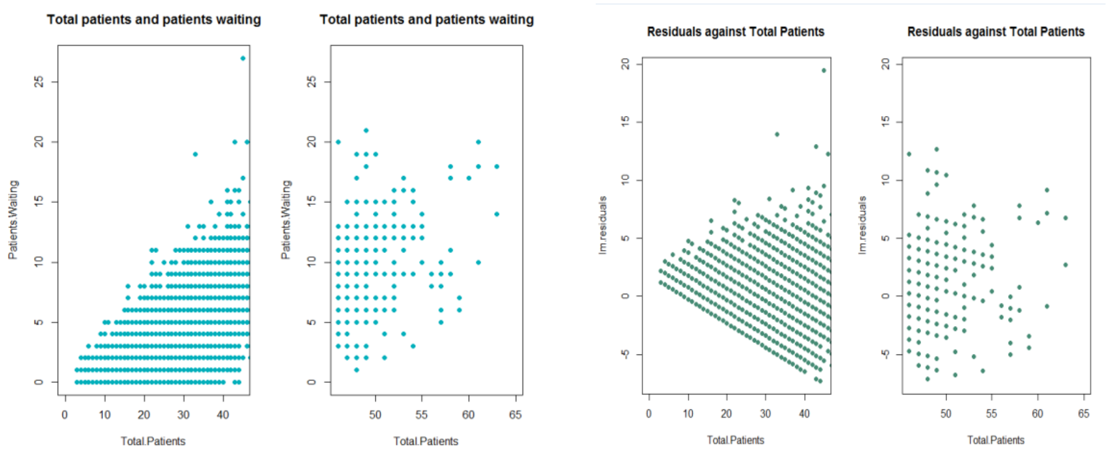


Figure A3.4 Scatterplots of 0-45 and 46-65 Total Patients, and the residuals

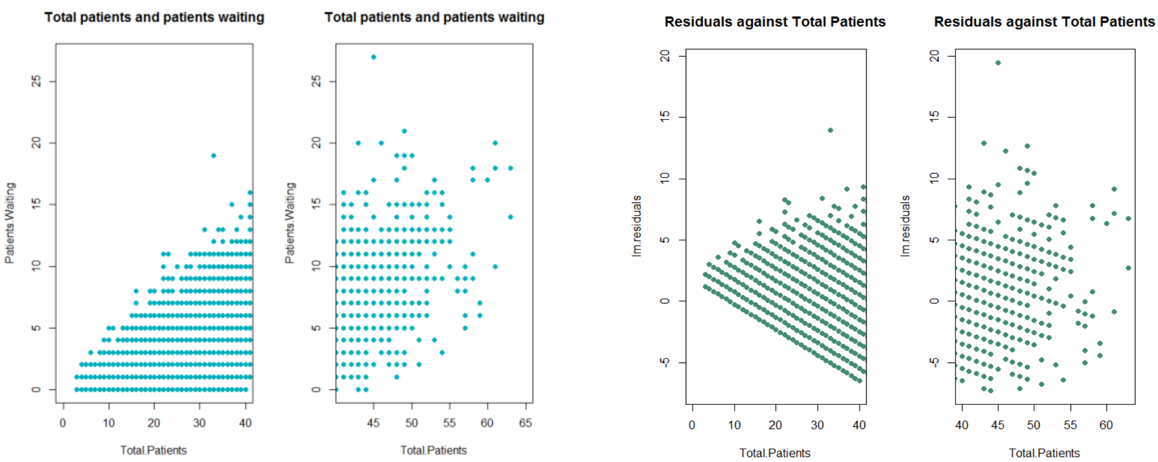


Figure A3.5 Scatterplots of 0-40 and 41-65 Total Patients, and the residuals

Figure A3.6 shows a scatterplot of *Total Number* of patients with *arrivals* 1, 2 and 3 hours ago, for visualisation.

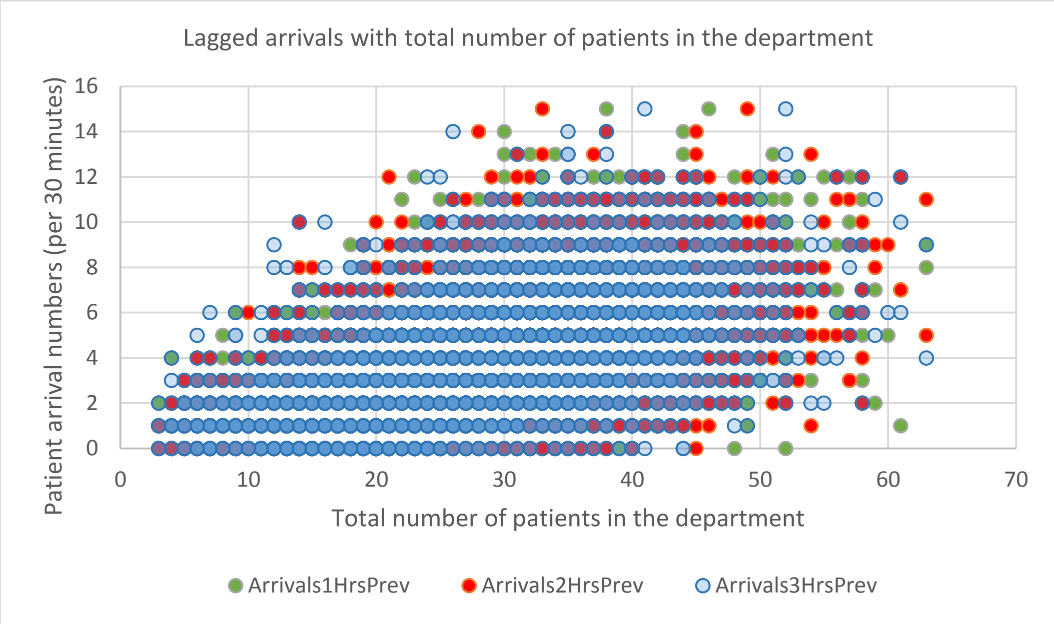


Figure A3.6 Scatterplot of lagged arrivals at 1 hour, 2 hours, and 3 previously with half hourly attendance

Figure A3.7 is a scatterplot of *Maximum Wait time* and *Total Number* of patients in the department 2, 3 and 4 hours ago for visualisation.

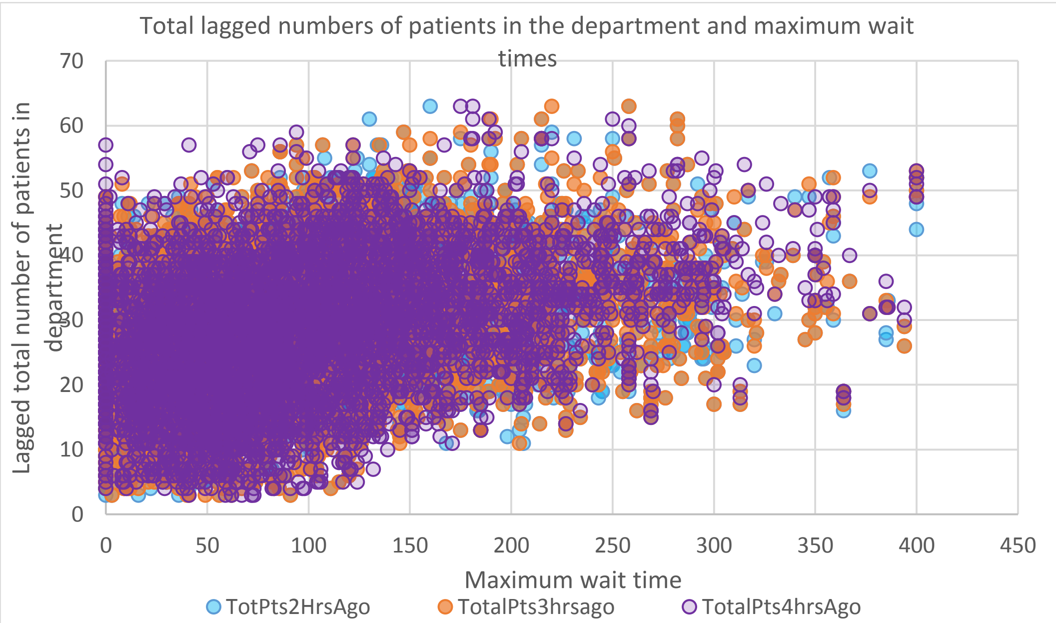


Figure A3.7 Scatterplot of lagged Total Patients with Maximum Wait time

A3.1.2 Diagnostic component: Time dependent Trigger

Section 6.4.3 creates a time-varying simulation trigger over a 24-hour period. To illustrate, *Total Patients* are plotted as a continuous line graph to assist visualisation (Figure A3.8) on a subset of data from 12:00 to 12:59 over a 115 day period, and as a scatterplot (Figure A3.9) against *Patients Waiting*.

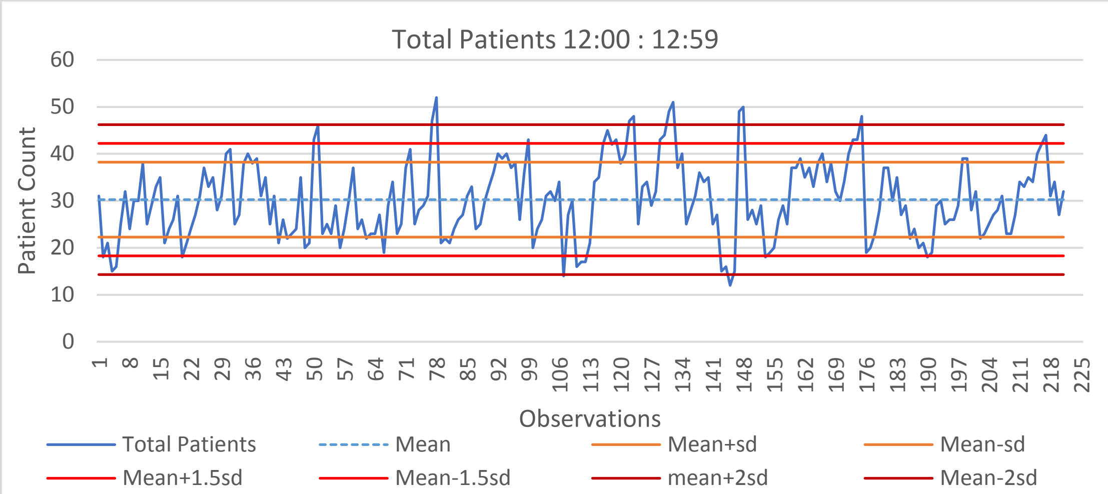


*Figure A3.8 115 days of 12:00 to 12:59 with mean and SD, 1.5*SD, 2*SD*

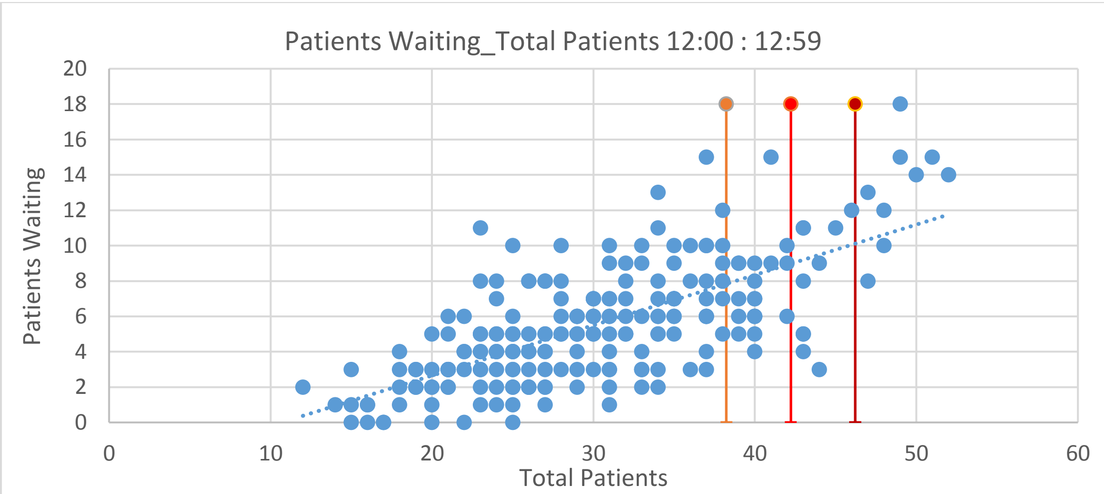


*Figure A3.9 Scatterplot of Patients Waiting and Total Patients, with SD, 1.5*SD, 2*SD for Total Patients*

Figure A3.10 plots the daily compliance (4-hour target) data with *Total Patients* data for 12:00:12:59 to examine at which point compliance tends to decrease. As with the plots in Chapter 6, Section 6.4.3, the compliance drops at one standard deviation.

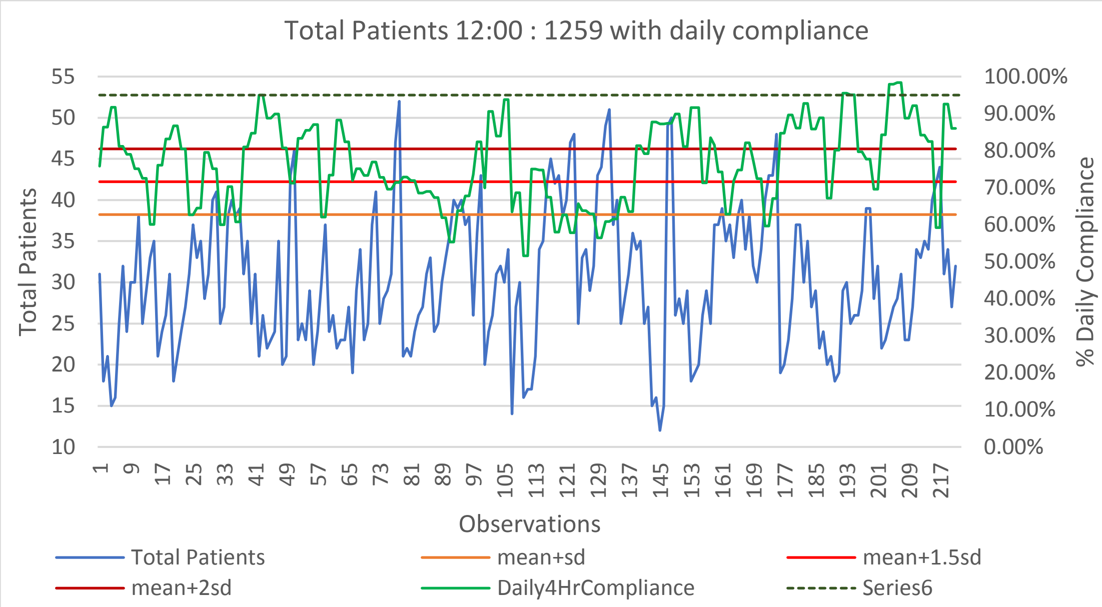


Figure A3.10 Total Patients (12:00 to 12:59) and Daily Compliance with the 4-hour target (24 hourly)

A3.1.3 Predictive analytics

There is an extensive body of work predicting demand for emergency services, using quantitative methods including linear regression (Jones et al., 2008; Ekstrom et al., 2015); machine learning (Khatri 2018; Yousefi et al., 2019); and time-series forecasting (Calegari et al., 2016; Choudhury, 2019). There has been significant interest in the use of climate factors for predicting ED demand, such as temperature and air quality, though these have shown mixed results, with Calegari et al. (2016) and Carvalho-Silva et al. (2018) demonstrating little to no additional predictive value. The field of machine learning is advancing rapidly, with an increase in publications applying these methods for healthcare forecasting. However, complex machine learning methods may not be a good approach where interpretability and clinician buy-in are priorities (Graham et al. 2018). For predicting emergency admissions, Wong et al. (2018) addressed the complexity of the approach as an implementation barrier in clinical practice. Additionally, these methods can require very large quantities of data, and data quality, collection and management requires substantial resources and commitment by healthcare stakeholders (Janke et al., 2016).

Time-series methods are part of a suite of predictive analytic methods which have shown considerable success in predicting emergency demand, in particular variations of auto-regressive moving averages (ARMA) as developed by Box and Jenkins (1976). These contain an autoregressive (AR) term (p), and a moving averages (MA) term (q). Where a single variable is observed at each time, the dataset is a *univariate time series*; where two or more variables are observed at each time, the dataset is known as a *multivariate time series*. Brownlee (2018) conceptualised ARMA problems as supervised learning problems, by using previous time steps (lags) as input variables and using the next time step as the output variable. The order of the variables is preserved, where the size of the lag is the number of previous time steps. ARMA models combine autoregressive and moving averages elements, but require fewer parameters than either used alone. The AR component uses the dependent (autocorrelated) relationship between an observation and a specified number of lagged observations, while the MA component uses the dependency between an observation and residual errors from a moving average model applied to lagged observations.

In ED forecasting, Aboagye-Sarfo et al. (2015) showed that ARMA and VARMA (vector autoregressive moving average) methods outperformed Winter's forecasting method, a widely-used univariate method for predicting seasonal data. Calegari et al. (2016) found that SARIMA (Seasonal Autoregressive Integrated Moving Average) provided better predictions of ED arrivals compared with more traditional seasonal approaches, and Choudhury (2019) found that SARIMA outperformed neural networks, and advanced seasonal models for predicting ED arrivals. SARIMA is an approach for modelling univariate time series data that contains a seasonal component. It contains additional seasonal terms which are similar to those in the ARIMA (p,d,q) model, where d is the degree of differencing. Differencing in statistics is a transformation applied to time-series data in order to make it stationary. A stationary time series' properties do not depend on the time at which the series is observed. The seasonal model is specified as SARIMA (p,d,q)(P,D,Q)s, where s is the seasonality and (P,D,Q) are the parameters influenced by the seasonal component. P uses the seasonally offset observation in the model, D is the order of seasonal difference, and Q is the order of moving averages or errors in the model. For this case study, SARIMA modelling has been chosen for creating short-term forecasts 2 and 4 hours ahead, as the ED data has a strong daily seasonality. Due to the availability of forecasting libraries, Python 3.7 is used for the forecast modelling.

A3.1.3.1 Time-series features

An observed time series can be seen as a realisation of a stochastic process (Chatfield & Xing, 2019) and is usually a sequence of real values $(x_1 \ldots x_e)$ taken at successive equally spaced points in time, from time *t=1* to time *t=e.* A time series can be plotted to obtain simple descriptive measures of the main properties of the series (Figure 6.17), and to visualise patterns, unusual observations and changes over time. These features include trends (a decrease or increase in the mean over time), seasonality (a regular repeating pattern related to the calendar), cyclic variations which are not regular, changes in the variance over time, and abrupt level changes. Additionally, the plots can identify outliers and missing values (Section 6.3.2). The data is not necessarily independent nor identically distributed, and the order of observations is important, because there is a dependency and changing the order could change the meaning of the data.

Given a time series, it is often useful to predict future values of the time series by utilising past longitudinal information to predict near future outcomes. It is therefore an appropriate method where there is numerical information available about the past, and it is reasonable to assume that some aspects of past patterns or sequences will continue into the future. Stationarity is discussed in Chapter 6, Section 6.5.2, however to introduce the concept, a stationary time series exhibits no trend, no systematic change in variance, and no seasonal variations. This means that the properties of one section of the data are similar to those of any other section. It does not mean that the series does not change over time, just that the *way* it changes does not itself change over time (Chatfield & Xing, 2019). Stationarity is a common assumption for many methods used in time-series analysis. Most time series data will violate this principle, however the term is often used to indicate that a stationary model can be fitted to a time-series by transforming a non-stationary time series into a stationary one, for example by removing the trend and/or seasonal variation, to model the variation in the residuals (Hyndman & Athanasopoulos, 2015). The first step in fitting an ARIMA model is to determine the order of the differencing needed to make the time series stationary. For some specific time point *r,* the observation x_{r-i} (i periods back) is called the *i*-th *lag* of x_r. A time series *Y* is generated by back-shifting another time series *X* by *i* time steps*.* A time series can be differenced until it becomes stationary, but the ACF and standard deviation (StD) should be inspected after each to determine whether further differencing in justified.

Where the variance is not constant over time, nonlinear transformation(s) such as logging and/or deflating and/or raising-to-some-power can convert the time series to a form where its local random variations are consistent (Nau, 2019).

Total Patients data is decomposed in Figure A3.11 to examine the seasonality and other components (Observed, Trend, Seasonal, Residual). This confirms that there is no long-term trend and the data has a clear daily seasonality.

```
res = sm.tsa.seasonal_decompose(quickerCordf.TotalPatientsS.dropna(),freq=48)
fig = res.plot()
fig.set_figheight(8)
fig.set_figwidth(15)
plt.show()
```

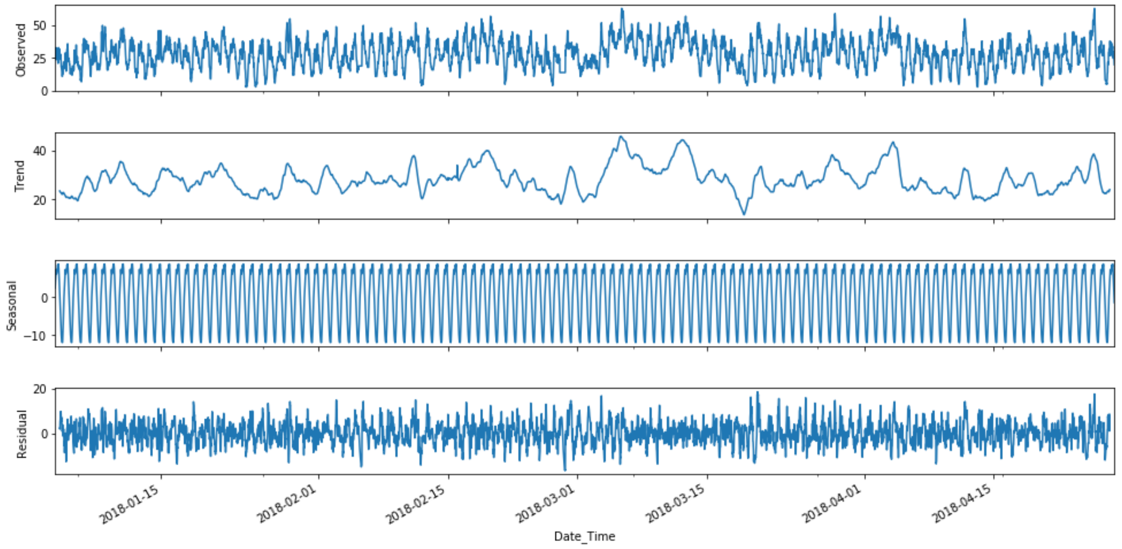


Figure A3-11 Total Patients decomposition

Autocorrelation is a feature of most time series, as the observations close together tend to be correlated, or serially dependent. Just as a correlation measures the strength of a relationship between two independent variables, autocorrelation measures the strength of the relationship between lagged values of a time series. It uses Pearson's correlation coefficient, returning a value between 1 and -1, where a value of 0 indicates no correlation. In most time series data, the data are correlated, which means that methods are required which deal with the inherently correlated structure, as apparently irregular variation may be explained in terms of probability models such as AR or MA.

The AR component of ARIMA models use the dependent (autocorrelated) relationship between an observation and a specified number of lagged observations, while the MA component uses the dependency between an observation and residual errors from a moving average model applied to lagged observations. Each is specified explicitly as a parameter as an integer value in the specification ARIMA (p,d,q). The stronger the correlation between the output variable and a lagged variable, the more weight the AR model will apply to that variable in the model. If there is little or no correlation between an output variable and its lag variables, the time-series problem may not be predictable (Brownlee, 2018).

One of the simplest ARIMA models is AR(1), or naïve forecast, which uses a linear model to predict the value at the present time using the value at the previous time. This is an autoregressive model of order 1, where the order indicates how many previous lags are used to predict the current time. This can provide a baseline performance as a point of comparison, to give an indication of how well other models will perform on the forecasting problem. The naïve forecast (AR(1)) reflects the autocorrelation, with a RMSE of 3.204 (Chapter 6, Figure 6.20). The data is split into training (0.83) and test sets (0.17), the model is run by predicting the output value as the same as the input value, and the RMSE is calculated. Measures of accuracy are discussed in the next section.

An *autocorrelation plot* will show the correlation coefficients for each lag variable, giving a good indication of which lag variables will be good candidates for use in a predictive model, and how the relationship between the lag values changes over time. Autocorrelation Functions (ACF) and Partial Autocorrelation Functions (PACF) can help to choose the parameters of the ARMA or ARIMA model. The ACF is a plot of the autocorrelation of a time series. A PACF summarises the correlation between observations in a time series with the relationships of intervening observations removed.

An ACF plot provides the lag number along the x-axis and the correlation coefficient value between -1 and +1 on the y-axis. The plot also includes 95% and 99% confidence interval for the correlation values. Correlation values above these lines are more significant than those below the line, providing a threshold for selecting more relevant lag values. A PACF summarises the correlations for an observation with lag values that are not accounted for by prior lagged observations. A time series with no autocorrelation is known as 'white noise', where the ACF should be close to 0, with some random variation (Hyndman & Athanasopoulos, 2015).

The model has an AR component if the ACF trails off after a lag and has a hard cut-off in the PACF after a lag. This lag is taken as the value for p. The model has a moving average component if the PACF trails off after a lag and has a hard cut-off in the ACF after the lag. This lag value is taken as the value for q. The model is a mix of AR and MA if both the ACF and PACF trail off (Brockwell & Davis, 2016).

The 'residuals' in a time-series are what is left after a model is fitted. Using time series analysis, the features in the data can be used to make forecasts, while 'noise' is the variability in the observations which cannot be explained by the model. It is usually equal to the difference between the observations and corresponding fitted values (Hyndman and Athanasopoulos, 2015). Ideally, the forecasting model will result in residuals which are uncorrelated, as correlations between residuals indicates that there is information left in the residuals which can be used for computing forecasts. Residuals should have a mean of zero, otherwise the forecasts are biased, and should have a constant variance and be normally distributed to make prediction intervals easier to calculate.

A3.1.3.2 Evaluating forecast accuracy

The accuracy of forecasts can only be determined by considering how well a model performs on new data that was not used for fitting the model. For this reason, as previously with the naïve forecasts, the set of data should be partitioned into a training set and a test set. The training set is used to estimate parameters of the forecasting method, and the test set is used to evaluate its accuracy. Hyndman and Athanasopoulos (2015) recommend an 80/20 split, although this value depends upon the size of the dataset and how far ahead forecasts need to be made, such that the test set is at least as large as the maximum required forecast horizon. As this study is interested in short term forecasts, small test sets of less than 20% are often used.

A model which fits the data well will not necessarily provide accurate forecasts. Over-fitting a model to data means having too many parameters such that it fits the training set well, but is then unable to forecast accurately. For example, Nau

(2019) advises using models where at least one of p and q is no larger than 1 to avoid overfitting.

A forecast "error" is the difference between an observed value and its forecast, hence, unlike a residual, they are calculated on the test set. Errors can be measured using scale-dependent or percentage measures. The two most common scale-dependent measures are:

- Mean Absolute Error (MAE): mean($|e_t|$)

  This is calculated as the average of the absolute forecast error values, and is useful when comparing forecast methods applied to a single time series, or several time series with the same units, as the units of the error are the same as the units of prediction. Minimising the MAE will result in forecasts of the median.

- Root mean squared error (RMSE): $\sqrt{\text{mean}(e^2_t)}$

  The *Mean Squared Error (MSE)* is calculated as the average of the squared forecast error values. This both forces errors to be positive, and puts more weigh on the large errors. MSE scores are the squared units of predictions, and taking the square root of the MSE transforms them back into the original unit of the predictions. This is known as the *Root Mean Squared Error (RMSE)*. A RMSE of zero indicates no error. Minimising the RMSE will result in forecasts of the mean.

Percentage errors are unit-free, and are useful for comparing performance between datasets.

- $p_t = 100e_t/y_t$
  Mean Absolute Percentage Error (MAPE): mean($|p_t|$)

  If $y_t = 0$, then for any t in the time period of interest, the MAPE will be infinite and will have extreme values where y_t is close to zero

As the intention in this study is to compare models using the same dataset, scale-dependent errors are used, and RMSE is chosen, RMSE values are always slightly higher than MAE values, which becomes more pronounced as the prediction errors increase. This is a benefit of using RMSE over MAE as RMSE penalises larger errors (Brownlee, 2015). For this study, cross-validation is used, where there are a series of test sets, each consisting of a single observation. One-step forecasts, or rolling forecasts, involve taking the regression coefficients learned by the model from the training data, and using it to make predictions in a rolling manner across the test dataset. As each step in the test set is executed, the prediction is made using the coefficients, and stored. Hyndman and Athanasopoulos (2015) recommend that a good way to choose the best forecasting model is to find the model with the smallest RMSE computed using time-series cross-validation. However it is important to note that RMSE does not generalise across multiple samples when doing cross-validation (Dinov, 2018).

A3.1.3.3 ARIMA

ARIMA is a generalisation of the simpler ARMA method, which adds integration (I), the use of differencing of raw observations (i.e. subtracting an observation from an observation at a previous time step) to make the time-series stationary, in particular to remove a trend or seasonality. For an ACF to make sense, the series must be 'weakly stationary', that is, the ACF for any particular lag is the same, regardless of where it is along the time series. The means that the mean is the same for all of t, the variance is the same for all of t, and the covariance (and correlation) between x_t and x_{t-1} is the same for all t. Most series are not stationary. A continual upward trend, for example, is a violation of the requirement that the mean is the same for all t. Distinct seasonal patterns also violate that requirement, and are exhibited in the Total Patients dataset.

Each ARIMA parameter is specified explicitly as an integer value in the specification ARIMA (p,d,q). Its analysis assumes that the time series data is stationary, so it needs to be made stationary by differencing the series (the d parameter) and then testing statistically that the result is stationary. The parameters are defined as:

AR (p) – The autoregressive component, lag order, or number of lag observations

I (d) – Trend difference order (integration). Over-differencing can result in additional complexity and the addition of extra serial correlation.

MA (q) – The size of the moving average window, or order of the moving average. This moving average process is an autoregression of the time series of residual errors from prior predictions. In other words it corrects future forecasts based on errors made on recent forecasts.

Stationarity means that the statistical properties of a process generating a time series do not change over time. It does not mean that the series does not change over time, just that the *way* it changes does not itself change over time (Chatfield & Xing, 2019). Stationarity is a common assumption for many methods used in time-series analysis.

The first step in fitting an ARIMA model is to determine the order of the differencing needed to make the time series stationary, however over-differencing can introduce negative autocorrelation and increase the standard deviation (Nau, 2019). For some specific time point r, the observation x_{r-i} (i periods back) is called the i-th *lag* of x_r. A time series Y is generated by back-shifting another time series X by i time steps. A time series can be differenced until it becomes stationary, but the ACF and SD should be inspected after each to determine whether further differencing in justified. For non-seasonal data, first-order differencing may be sufficient. For seasonal data, a seasonal difference is recommended (Hyndman & Athanasopoulos, 2015), while a first order difference may also be required. A seasonal difference is the difference between an observation and the previous observation from the same season. For Total Patients data, this requires subtracting each observation from the same time in the previous cycle (48) to create a new time series. This is necessary, otherwise the model assumes that the seasonal pattern will fade over time (Nau, 2019).

The Augmented Dickey Fuller Test (ADF) is a unit root test for stationarity in a time-series. This is done in statsmodels using adfuller, for analysis of a univariate process in the presence of serial correlation (Statsmodels, 2019b), on the seasonally differenced data, to determine the need for first order differencing (Figure A3-12).

```
In [189]: from statsmodels.tsa.stattools import adfuller
          from matplotlib import pyplot

          # create a differenced series
          def difference(dataset, interval=1):
              diff = list()
              for i in range(interval, len(dataset)):
                  value = dataset[i] - dataset[i - interval]
                  diff.append(value)
              return Series(diff)

          X = quickerCordf.TotalPatientsS.values
          X = X.astype('float32')
          # difference data
          obsv_in_day = 48
          stationary = difference(X, obsv_in_day)
          stationary.index = quickerCordf.TotalPatientsS.index[obsv_in_day:]
          # check if stationary
          result = adfuller(stationary)
          print('ADF Statistic: %f' % result[0])
          print('p-value: %f' % result[1])
          print('Critical Values:')
          for key, value in result[4].items():
              print('\t%s: %.3f' % (key, value))
          # save
          stationary.to_csv('stationary.csv', index=False)
          # plot
          stationary.plot()
          pyplot.show()

          ADF Statistic: -12.023747
          p-value: 0.000000
          Critical Values:
                  1%: -3.432
                  5%: -2.862
                  10%: -2.567
```

Figure A3-12 ADF statistic on seasonally differenced Total Patients Time Series

Figure A3-12 shows the ACF and PACF plots of the seasonal differenced data. There is still daily seasonality present in the data (Appendix 3, Figure A3-13). This indicates that it may be worth considering a better model of seasonality, such as modelling it directly, rather than attempting to remove it from the model using seasonal differencing. However the parameters of an ARIMA model are explored.

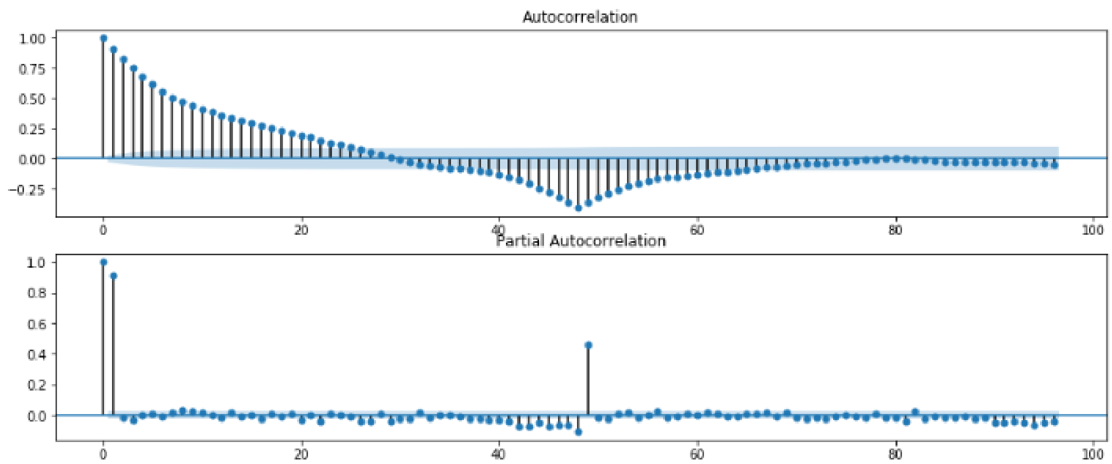


Figure A3-13 ACF and PACF of seasonal differenced Total Patients data (lag = 96)

The first significant ACF lag can indicate the order of the MA (q) parameter, and this is 1. The first significant PACF lag can indicate the order of the AR (p) parameter, which is also 1. An ARIMA (1,0,1) is therefore a starting point on the seasonally differenced data. Using an 80:20 train-test split, ARIMA (1,0,1) is implemented on a sample of the seasonally differenced data as a 7-day rolling forecast (Figure A3-14).

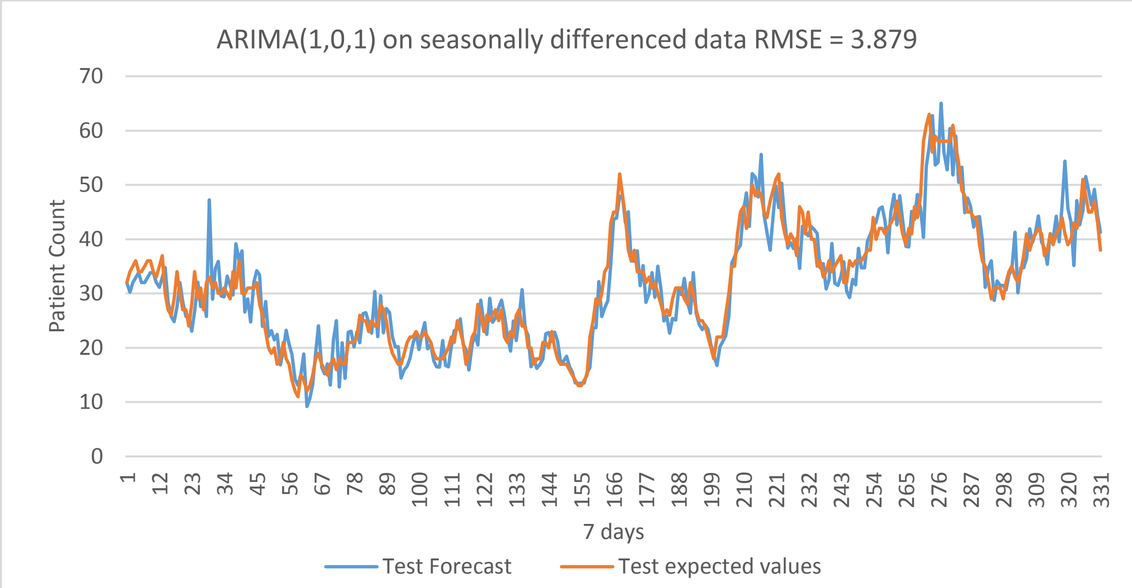


Figure A3-14 ARIMA (1,0,1) on seasonally-differenced Total Patients

This provides a good starting point however the one-step rolling RMSE are still higher than the naïve model. To confirm this analysis, a grid search of a suite of ARIMA parameters was conducted to check that there is not an ARIMA model that can outperform ARIMA (1,0,1) in test performance (using the seasonal stationary data). This is done by searching p, d, q values for the combination with the best performance by searching all combinations of p(0-3), d(0-2) and q(0-3), making 18 possible combinations (Figure A3-15). Higher parameters could be investigated, for example within seasonal patterns, however parsimonious models are usually considered best (Nau, 2019) and the seasonal differencing confirmed that the data was stationary.

```
ARIMA(0, 0, 1) RMSE=6.633
ARIMA(0, 0, 2) RMSE=5.404
ARIMA(0, 1, 1) RMSE=3.933
ARIMA(0, 1, 2) RMSE=3.936
ARIMA(1, 0, 0) RMSE=3.879
ARIMA(1, 0, 1) RMSE=3.879
ARIMA(1, 0, 2) RMSE=3.876
ARIMA(1, 1, 0) RMSE=3.933
ARIMA(1, 1, 1) RMSE=3.879
ARIMA(2, 0, 0) RMSE=3.879
ARIMA(2, 0, 1) RMSE=3.879
ARIMA(2, 0, 2) RMSE=3.878
ARIMA(2, 1, 0) RMSE=3.935
ARIMA(2, 1, 1) RMSE=3.879
Best ARIMA(1, 0, 2) RMSE=3.876
```

Figure A3-15 Results of grid search of p(0-3), d(0-2) and q(0-3) for ARIMA (p,d,q)

Using the grid search, the best ARIMA model is ARIMA (1,0,2) with a RMSE of 3.876 and MSE of 15.02 (Figure A3-14). This marginally outperforms the ARIMA (1,0,1) tested above. Models with higher numbers of parameters failed to converge and were not returned. The ARIMA(1,0,2) one-step ahead forecasts for 999 observations is plotted below (Figure A3-16). The model performs fairly well, predicting an expectation of error of 3.812 total patients.

The next step is to check residual errors. These should ideally be normally distributed with a mean of zero. This can be done using summary statistics and plots to investigate the residual errors from the ARIMA (1,0,2) model (Figure A3-17). This returns descriptive statistics of the residual errors. The mean and median are very close to 0 and the range is very slightly shifted to the right but they approximate normal, meaning there should be no bias in the forecasts. The first graph is a frequency histogram of the residual errors between the test set and the model forecasts, and the second is a probability density distribution.

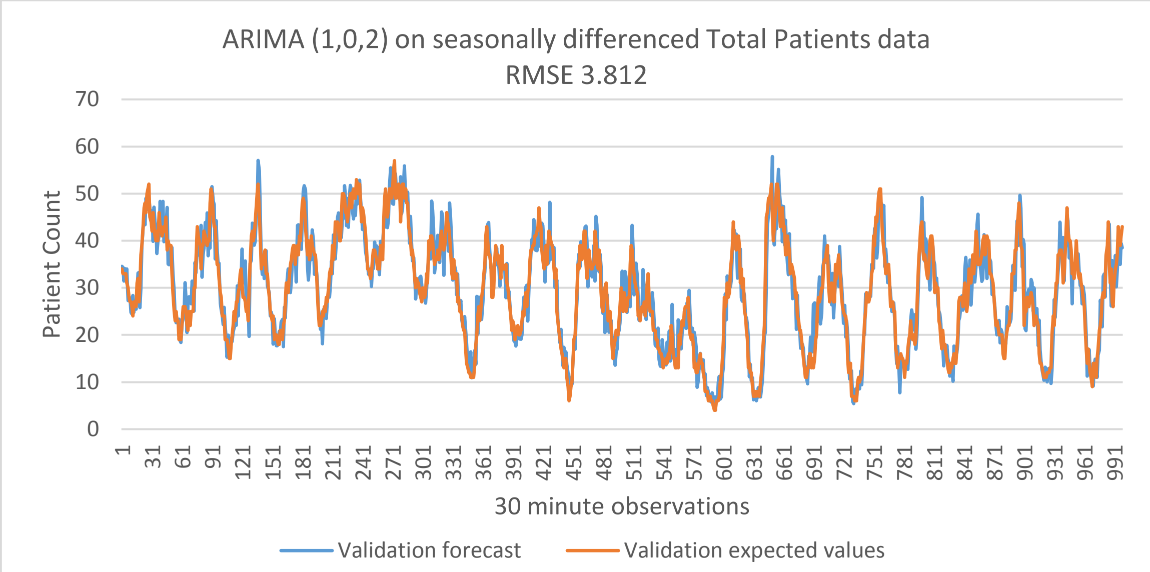
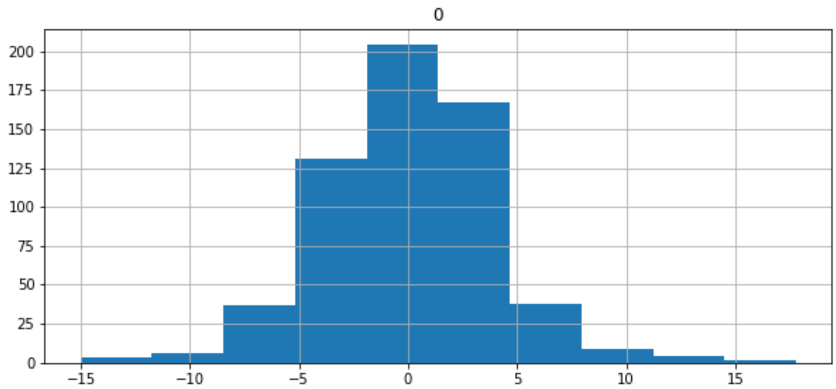


Figure A3-16 ARIMA (1,0,2) on seasonally differenced data

```
count   600.000000
mean      0.030132
std       3.879510
min     -15.027067
25%      -2.317411
50%       0.036860
75%       2.363799
max      17.791899
```



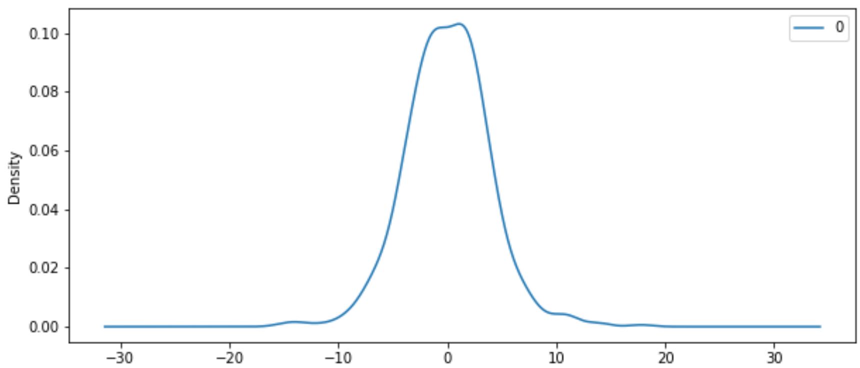


Figure A3-17 Summary statistics and plots of the residuals on ARIMA(1,0,2) on seasonally differenced data

The time series of the residual errors is checked for autocorrelation. If present, it would suggest that a model has more opportunity to capture the temporal structure in the data. The results (Figure A3-18) suggest that there is still autocorrelation present in the data, particularly at the seasonal (48 observations) points, with significant spikes at 48, and again in the PACF at 96 and 144. The seasonal ARIMA model incorporates both non-seasonal and seasonal factors and was therefore chosen for the *Predict* component of IHAF (Chapter 6, Section 6.5.3).

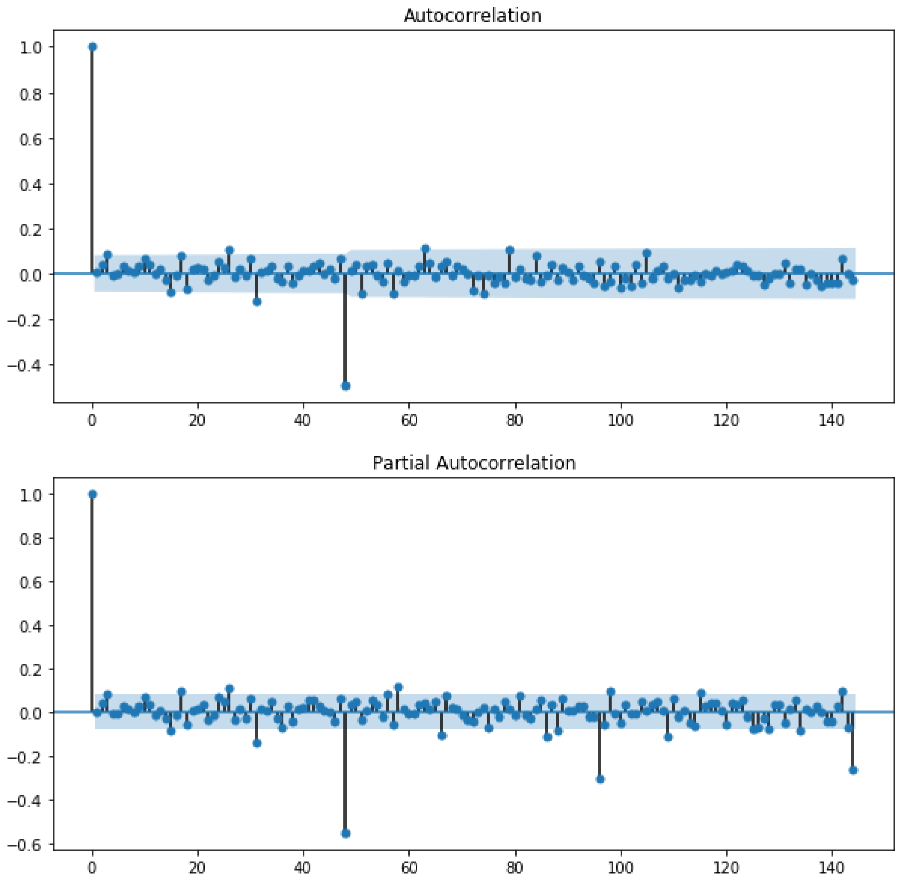


Figure A3-18 ACF and PACF of residuals of ARIMA (1,0.2)

A3.1.3.4 SARIMA

A summary of SARIMA (1,1,2)(1,0,1)[48] model fit is in Figure A3-19.

The Ljung-Box statistic (Q) is 35.28. The Ljung-Box test is a diagnostic tool used to test the lack of fit of a time series model. The test is applied to the residuals of a time series after fitting an ARMA model to the data. The test examines *h* autocorrelations of the residuals. If the autocorrelations are very small, it can be concluded that the model does not exhibit significant lack of fit, in other words, the Q-statistic needs to be non-significant, which it is. Hyndman

(2014) recommends using h=min(2s,T/5) where s is the period of seasonality and T is the length of the time series. In this case, the minimum *is* 2*s, so 96 lags were used to analyse Q, which is non-significant in most lags, using the 'ljungox' method in Statsmodels. The Jarque-bera test is a test for normality, requiring the 'jarquebera' method. This is significant to $p<0.05$. The Goldfeld-Quandt test is a statistical test for heteroscedasticity, requiring the method 'breakvar'. The null hypotheses is homoscedacity, and in this case this is rejected, suggesting that variance is actually increasing over time in this sample.

```
                                 Statespace Model Results
==========================================================================================
Dep. Variable:                                  y   No. Observations:                 1248
Model:             SARIMAX(1, 1, 2)x(1, 0, 1, 48)   Log Likelihood               -2996.929
Date:                            Wed, 12 Feb 2020   AIC                           6005.857
Time:                                    15:46:43   BIC                           6036.378
Sample:                                         0   HQIC                          6017.356
                                           - 1248
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.8449      0.024     35.671      0.000       0.798       0.891
ma.L1         -0.8887      0.035    -25.336      0.000      -0.957      -0.820
ma.L2         -0.0818      0.031     -2.643      0.008      -0.142      -0.021
ar.S.L48      -0.0260      0.028     -0.927      0.354      -0.081       0.029
ma.S.L48      -0.9867      0.106     -9.334      0.000      -1.194      -0.780
sigma2         7.7981      0.813      9.596      0.000       6.205       9.391
===================================================================================
Ljung-Box (Q):                       35.28   Jarque-Bera (JB):                32.52
Prob(Q):                              0.68   Prob(JB):                         0.00
Heteroskedasticity (H):               0.82   Skew:                            -0.02
Prob(H) (two-sided):                  0.05   Kurtosis:                         3.81
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Figure A3-19 SARIMA(1,1,2)(1,0,1)[48] model fit

The residuals are examined below to confirm the fit of the model (Figure A3-20). They approximate a normal distribution, which is a useful confirmation of the PIs, and there is no significant autocorrelation in the ACF, suggesting the chosen model is a good fit for the seasonally differenced Total Patients data.

```
              0
count  48.000000
mean    0.008252
std     2.710339
min    -6.685738
25%    -1.070181
50%     0.040956
75%     1.427273
max     6.185776
```

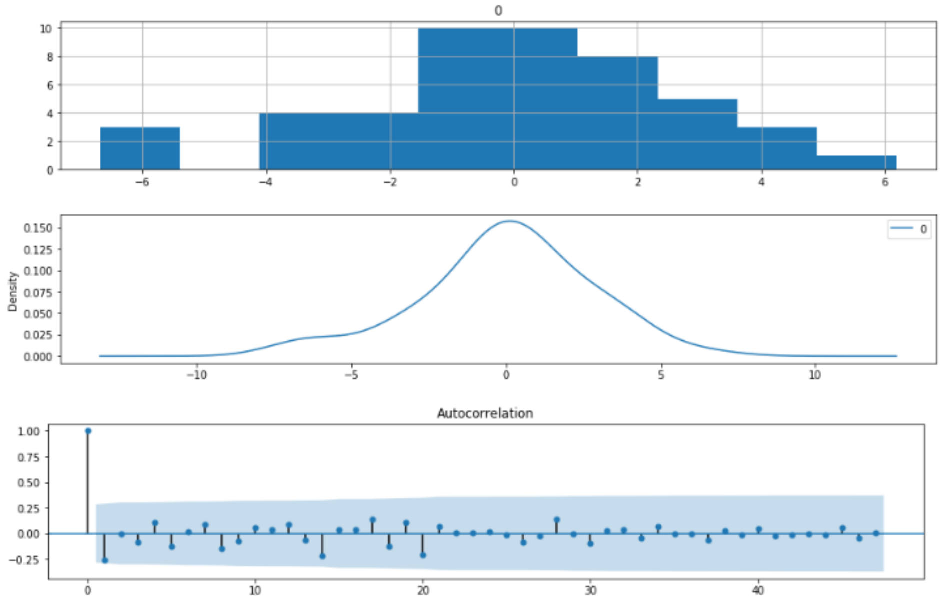


Figure A3-20 Summary statistics and summary plots of the residuals of SARIMA(1,1,2)(1,0,1)[48]

A3.1.4 Resampling the data

The data is resampled to reduce the granularity, and provide one-step ahead forecasts. Feature engineering is used to resample the data, as multistep forecasting places a significant burden on existing data by assuming the accuracy of each step ahead. This is done by calculating the average Total Patients in the resampled time period.

Naïve forecasts are repeated as a baseline on each resampled dataset. As can be expected, as the granularity of the data reduces, the naïve forecasts lose accuracy. These are plotted in Figures A3-21 to A3-24.

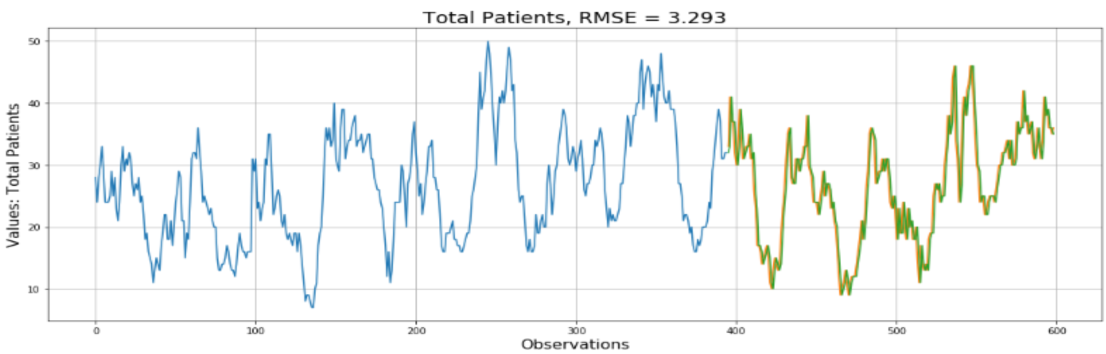


Figure A3-21 Baseline Naïve forecasts with RMSE on Total Patients data
---- predicted values ---- expected values

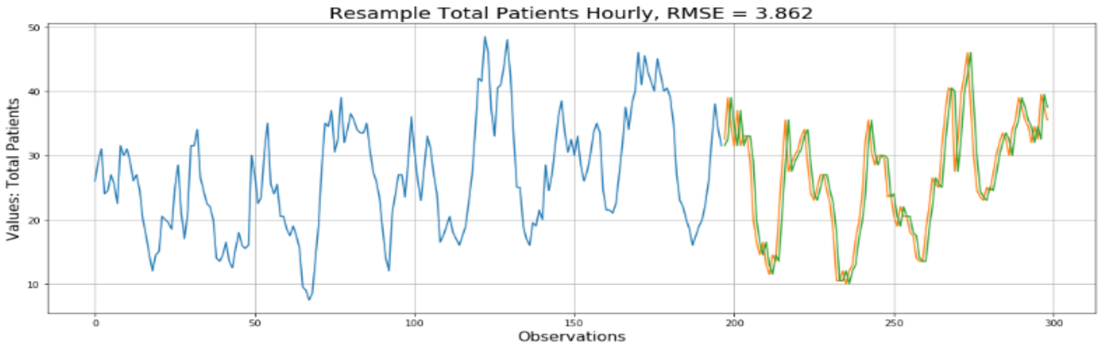


Figure A3-22 Baseline Naïve forecasts with RMSE on 1-hour resampled Total Patients data ---- predicted values ---- expected values

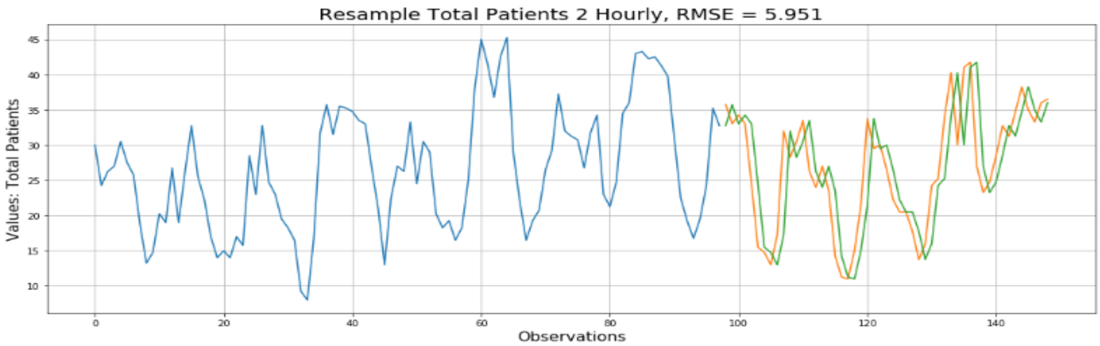


Figure A3-23 Baseline Naïve forecasts with RMSE on 2-hour resampled Total Patients data ---- predicted values ---- expected values

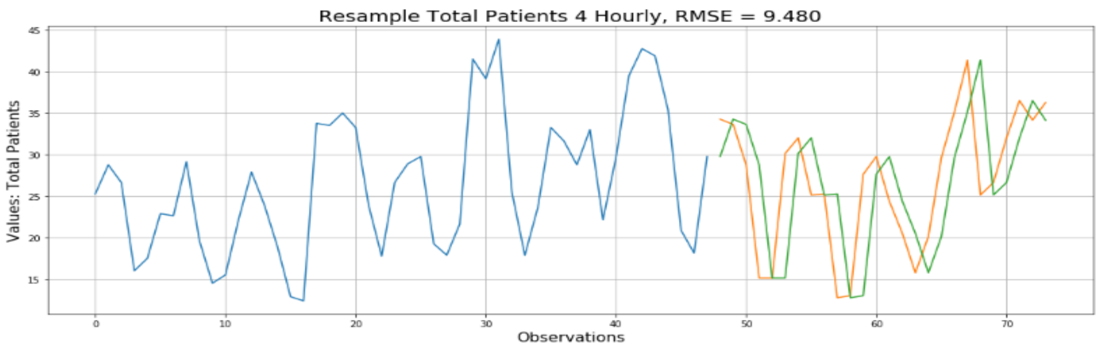


Figure A3-24 Baseline Naïve forecasts with RMSE on 4-hour resampled Total Patients data ---- predicted values ---- expected values

One-step SARIMA forecasts are performed on each resampled dataset. These are plotted in Figures A3-25 to A3-27. Each outperforms the naïve forecasts.

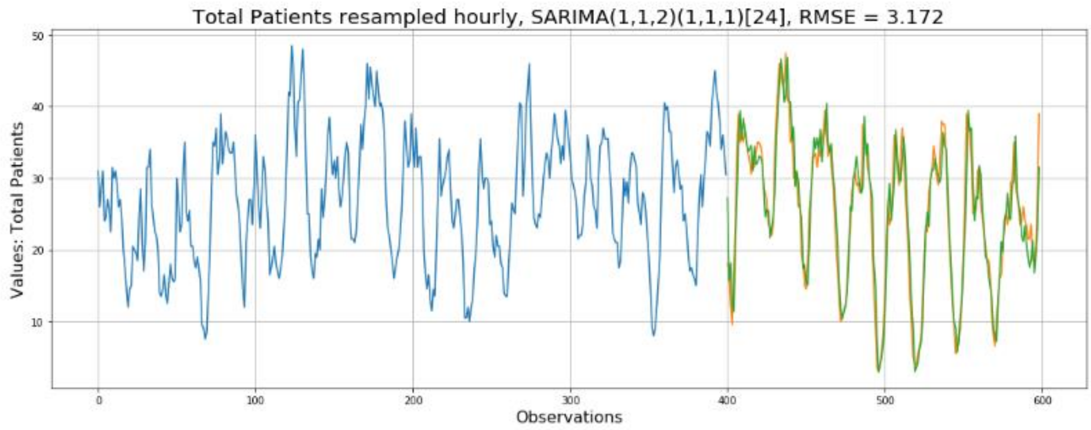


Figure A3-25 SARIMA forecasts with RMSE on 1-hour resampled Total Patients data
---- predicted values ---- expected values

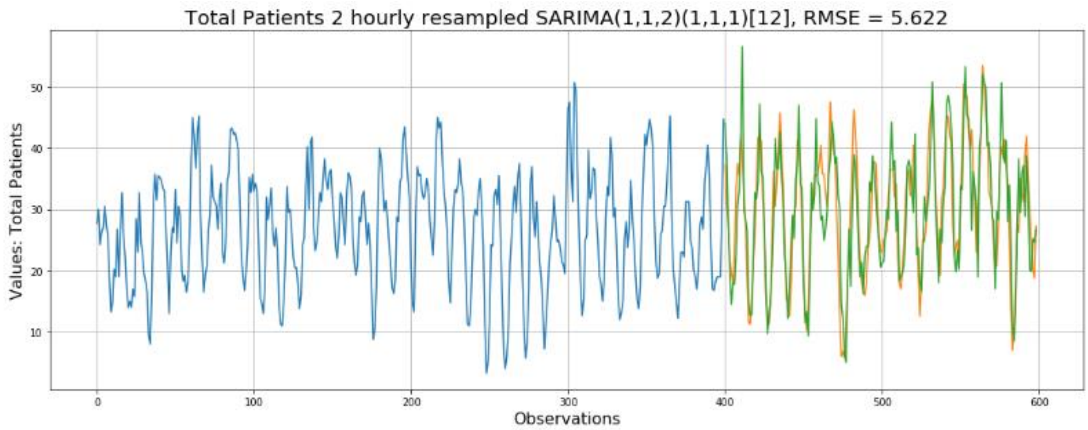


Figure A3-26 SARIMA forecasts with RMSE on 2-hour resampled Total Patients data
---- predicted values ---- expected values

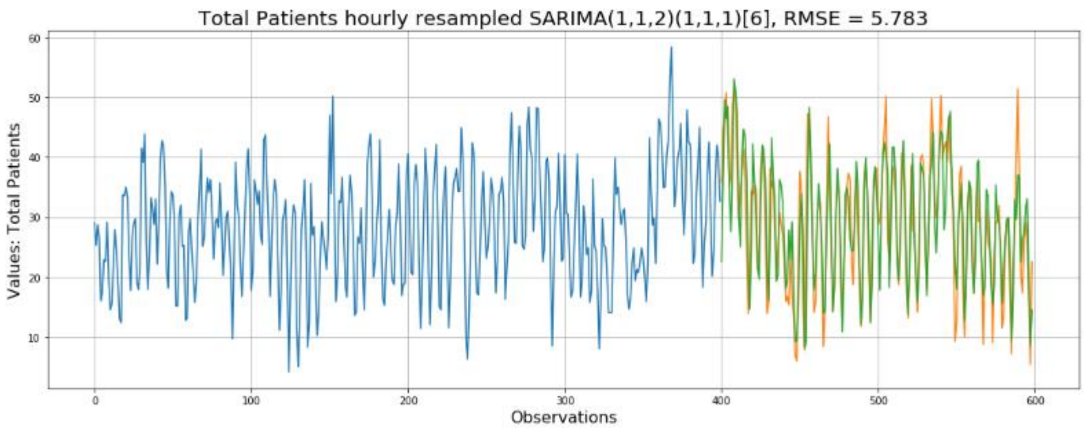


Figure A3-27 SARIMA forecasts with RMSE on 4-hour resampled Total Patients data
---- predicted values ---- expected values

Figure A3-28 illustrates the diagnostic results of the residuals on the 2-hourly resampled data. This allows forecasts to be updated every 30 minutes, for 2 hours and 4 hours ahead. Examination of the residuals and ACFs shows that the residuals are approximately normally distributed with a small peak of remaining autocorrelation at 1, which may be relevant. Model summary results show that the Jarque-Bera test for normality of the residuals is significant at p<0.001; the Goldfeld-Quandt test for heteroscedasticity of residuals is not significant. The null hypotheses is homoscedacity, and across the full dataset, this is supported with p=0.9684, indicating the variance is stationary. The Ljung-Box test for serial correlation however has some non-significant p-values in early lags, indicating that some serial correlation remains. This can be visualised in the ACF of the residuals at 12 and 24 hours, reported as non-significant.

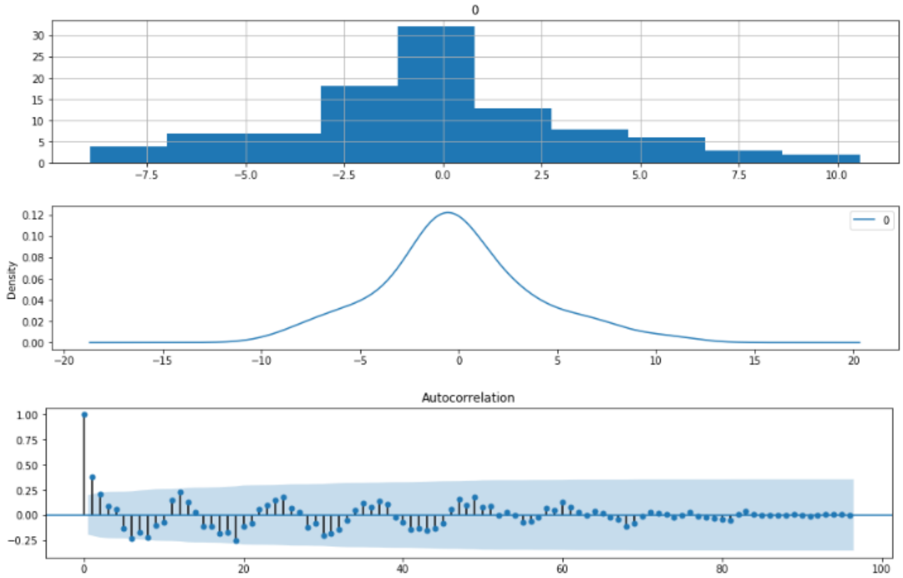


Figure A3-28 Summary statistics and plots of the residuals on SARIMA(1,1,2)(1,0,1)[12] on 2 hourly resampled seasonally differenced Total Patients data

A3.1.5 Forecasting (expanded from Chapter 6, Section 6.5.5).

The model is fitted to be used later for making predictions. Using get_forecast allows prediction intervals and multi-step forecasting (Figure A3-29).

```
forecastci = model_fit.get_forecast(steps = 2)
df95 = DataFrame(forecastci.summary_frame(alpha=0.05))
df80 = DataFrame(forecastci.summary_frame(alpha=0.20))

print(df80)
print(df95)

x_labels = [0,2,4]

ax = pyplot.gca()
plt.rcParams["figure.figsize"] = [15, 6]
plt.title('4-hour forecasts using 2-hrly Total Patients with PIs', fontsize = 20)
plt.xlabel('Forecasts', fontsize = 16)
plt.ylabel('Total Patients', fontsize = 16)
plt.grid(True)
plt.tight_layout()
plt.xticks(x_labels)
df80.plot(kind='line', y='mean', ax=ax)
df80.plot(kind='line', y='mean_ci_lower', color = 'orange', linestyle ='--', ax=ax)
df80.plot(kind='line', y='mean_ci_upper', color = 'orange', linestyle ='--', ax=ax)
df95.plot(kind='line', y='mean_ci_lower', color = 'pink', linestyle ='--', ax=ax)
df95.plot(kind='line', y='mean_ci_upper', color = 'pink', linestyle ='--', ax=ax)
pyplot.show()
```

```
y       mean   mean_se  mean_ci_lower  mean_ci_upper
0  29.515698  4.512923      23.732154      35.299241
1  31.991627  6.095951      24.179352      39.803903
y       mean   mean_se  mean_ci_lower  mean_ci_upper
0  29.515698  4.512923      20.670532      38.360864
1  31.991627  6.095951      20.043783      43.939472
```

Figure A3-29 Multi-step forecasts with prediction intervals

However it also required the model to be retrained each time it is needed. Having trained the model on all available data, instead, the model and its parameters is saved so that it does not need to relearn the regression coefficients each time a prediction is needed (Figure A3-30). The Statsmodels module in Python has built-in functions to save and load models by calling save() and load() on the fitted SARIMAX Results object (Statsmodels, 2019). The predictions can now be made using ARResults.predict.

```
# save model to file
model_fit.save('2Forecastmodel2Hr.pkl')

#save half-hourly dataset
np.save('30minDataset', dfTotPat)

#save TotPat2hr Dataset
np.save('2Dataset2Hr.npy', TotPat2Hr)
```

```
from statsmodels.tsa.ar_model import ARResults

# load model and data
model2 = ARResults.load('2Forecastmodel2Hr.pkl')
data = np.load('2Dataset2Hr.npy')


# make predictions
prediction1 = model2.predict(start=len(data), end=len(data))
multistepdata = np.append(data, prediction1)
prediction2 = model2.predict(start=len(multistepdata), end = len(multistepdata))
multistepdata = np.append(data, prediction2)

#Estimate prediction intervals
print('Prediction in 2 hours: %.f' % prediction1[0])
PIU1_80 = prediction1[0] + (1.28* df80.mean_se[0])
PIL1_80 = prediction1[0] - (1.28* df80.mean_se[0])
print('          PIU2_80: %1.f' % PIU1_80)
print('          PIL2_80: %1.f' % PIL1_80)


print('Prediction in 4 hours: %.f' % prediction2[0] )
PIU2_80 = prediction2[0]  + (1.28* df80.mean_se[1])
PIL2_80 = prediction2[0]  - (1.28* df80.mean_se[1])
print('          PIU4_80: %1.f' % PIU2_80)
print('          PIL4_80: %1.f' % PIL2_80)
```

```
Prediction in 2 hours: 34
          PIU2_80: 40
          PIL2_80: 28
Prediction in 4 hours: 31
          PIU4_80: 39
          PIL4_80: 24
```

Figure A3-30 Using save() and load() to make 2 and 4 hour forecasts with prediction intervals

Now the forecast model needs to be kept updated, once the next real observation is made available by *NHSquicker*. This requires updating the data set used as inputs to make the subsequent prediction (Figure A3-31).

The following steps are required:

- The new observation is recorded. In Figure A3-31, the unrealistic figure 120 is manually inputted.
- The 30 minute dataset and 2 hour dataset are loaded and converted from numpy arrays to dataframe objects. This is a two-dimensional data structure that allows Pandas to manipulate the data. The 30 minute dataset is indexed with the original date-time index so that an additional row can be added.

- In Figure A3-31, the new observation is inputted using pd.to_datetime. However for the integrated model, the new observation can be added in real-time using pd.Timestamp.now(). The index is retained using ignore_index=False.
- To control the size of the dataset, it is saved with the first value removed as a new value has now been added to the end.
- Save() is used to resave the 30 minute dataset.
- It is now resampled 2 hourly and saved as the 2 hour dataset. The last five observations of each is printed to ensure each has saved correctly.

```
# update the data for the SARIMA model with a new obs
# get real observation from NHSquicker.  Assume it is 120 for tracking

observation = 120
# Load the saved data
data = np.load('2Dataset2Hr.npy')
data30 = np.load('30minDataset.npy')

#convert to dataframe for appending new observation and removing first observation
df = DataFrame(data)
df30 = DataFrame(data30)

# update and save observation as new row at current date/time; remove first obs
df30.index = dfTotPat.TotPat.index
ts = pd.to_datetime("2018-12-27 17:00:02.000000", format="%Y-%m-%d %H:%M:%S.%f")
#Use this when using current data
#ts = pd.Timestamp.now()
new_row = pd.DataFrame([[observation]], columns = [0], index=[ts])
df30 = pd.concat([df30, pd.DataFrame(new_row)], ignore_index=False)
df30 = df30[1:]
print(df30.tail())

np.save('30minDataset', df30)

#resample 2 hourly
df = df30.resample("2H").mean()
print(df.tail())

np.save('2Dataset2Hr.npy', df)
```

Figure A3-31 The datasets are updated with the new observation

It was mentioned that pd.Timestamp.now() can be used to update the date-time index in the 30 minute data in the Python forecast model code, while the new observation needs to be inserted directly into the forecast code as a variable. This creates a new row in the 30 minute dataset. This new dataset is then used to update the 2 hour dataset for making forecasts.

Using `java.io.InputStreamReader;` the code in Figure A3-32 is able to call on the python forecast model “SARIMA2_4.py” and insert arguments.

```
Process p = Runtime.getRuntime().exec("python SARIMA2_4.py " + args[0] + "" + args[1] + " " + args[2]);
BufferedReader in = new BufferedReader(new InputStreamReader(p.getInputStream()));
String read=null;

while ((read = in.readLine()) != null)
{
    System.out.println("value is : "+read);
}
```

Figure A3-32 Java code to insert new observation into python forecast model

This means the new observations in the Python forecast model (Figure A3-30) can be replaced with the code snippet in Figure A3-33.

```
print(sys.argv[1])
observation = int(sys.argv[1])
```

Figure A3-33 Python code to receive new observation from Java

A similar procedure is required for returning the forecasts to be sent to the simulation engine.

A3.2 Prescribe component of IHAF

A3.2.1 Historical data

This subsection provides additional material for Chapter 6, Section 6.7, specifically Step 3: Data collection and analysis.

Arrivals at the ED vary by hour of day and day of week, so an arrival schedule was constructed for each (Table A3.2). This enables entities in the simulation model to arrive per hour of day and day of week.

Table A3.2 Table 0-1 Hourly arrival rate schedule for each day of week

	min Sun	min Mon	min Tue	min Wed	min Thur	min Fri	min Sat
00	0.109	0.088	0.088	0.087	0.092	0.094	0.101
01	0.096	0.070	0.069	0.064	0.073	0.073	0.077
02	0.085	0.056	0.056	0.054	0.060	0.052	0.067
03	0.066	0.039	0.046	0.046	0.048	0.046	0.061
04	0.073	0.052	0.050	0.040	0.046	0.047	0.059
05	0.058	0.051	0.049	0.049	0.047	0.046	0.052
06	0.063	0.055	0.052	0.045	0.047	0.045	0.053
07	0.066	0.066	0.067	0.060	0.051	0.056	0.059
08	0.119	0.119	0.110	0.117	0.106	0.104	0.113
09	0.168	0.188	0.188	0.171	0.170	0.160	0.168
10	0.213	0.234	0.202	0.194	0.194	0.177	0.197
11	0.209	0.222	0.210	0.179	0.201	0.199	0.194
12	0.216	0.239	0.204	0.196	0.196	0.195	0.204
13	0.203	0.224	0.203	0.209	0.201	0.184	0.193
14	0.202	0.218	0.184	0.197	0.210	0.191	0.197
15	0.196	0.196	0.184	0.180	0.193	0.184	0.189
16	0.208	0.217	0.194	0.194	0.198	0.187	0.181
17	0.190	0.217	0.200	0.191	0.194	0.212	0.192
18	0.205	0.253	0.228	0.237	0.226	0.208	0.197
19	0.205	0.202	0.203	0.213	0.199	0.182	0.176
20	0.197	0.199	0.196	0.212	0.196	0.185	0.180
21	0.182	0.192	0.158	0.172	0.167	0.158	0.163
22	0.140	0.136	0.147	0.136	0.134	0.139	0.138
23	0.109	0.105	0.095	0.105	0.105	0.116	0.118

The distribution of triage categories were found to be relatively stable per year of available data (Figure A3-34), 2016 – 2018. This enables entities to be allocated a triage category upon arrival into the system according to a probability distribution. In Chapter 6 (Section 6.7) it can be seen that the daily arrival patterns per triage category follows the overall arrival pattern.

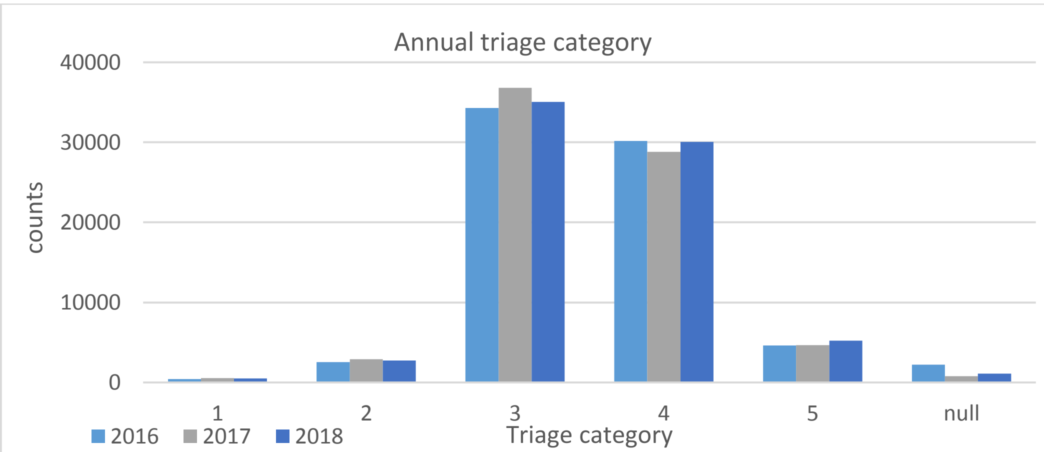


Figure A3-34 Triage category probability distribution calculated using three years of data (2016-2019)

Length of stay in ED are wide and flat, vary little between triage categories, and do not differentiate between time spent in treatment and time waiting for treatment. For behavioural reasons (i.e. working to targets), ED lengths of stay all peak sharply at the four hour mark, distorting the distribution. For illustration, Category 5 lengths of stay are plotted in Figure A3-35 from the ED use-case.

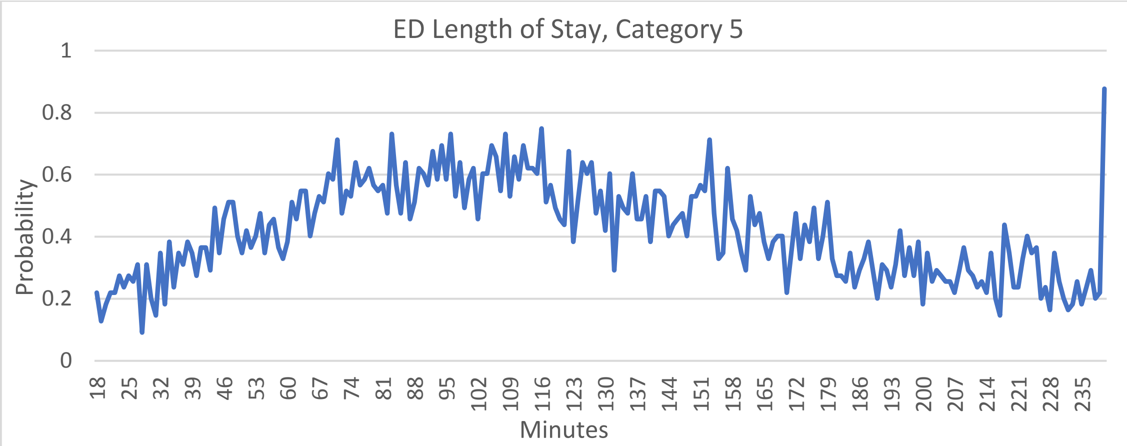


Figure A3-35 Distribution of length of stay of Category 5 patients

A better approach is to determine the proportion of patients who had no treatment, one treatment, two treatments, and three treatments, for each triage category, as captured in the ED dataset. This includes all treatment options, including resuscitation, drug administration by all methods, splints, plaster, dressings, and minor operations. From the data, the percentage of patients by triage category requiring each number of treatments is tabulated in Table A3.3. In table A3.4, the percentage of patients who required 1,2 or 3 treatments, 2 or 3 treatments, or only 3 treatments is calculated. From here, the probability of having zero treatments, one treatment only, two treatments only, and three treatments only is calculated (Table A3.5). Note all calculations are rounded to integers.

Table A3.3 Percentage of patients per triage category who had 0,1,2,or 3 treatments

Triage category	Zero treatment (%)	One treatment (%)	Two treatments (%)	Three treatments (%)
1	50	17	10	23
2	42	20	12	25
3	48	25	12	15
4	48	21	14	17
5	49	17	15	19

Table A3.4 Proportion of patients per triage category who had 1,2,or 3, 2 or 3, or 3 treatments

Triage category	1,2,3 treatments (%)	2,3 treatments (%)	3 treatments (%)
1	50	33	23
2	58	38	25
3	52	27	15
4	52	31	17
5	51	34	19

Table A3.5 is calculated by determining the proportion, for example, of those in Category 1 who had 2 or 3 treatments = 33/50 = 0.65. Of those, 23/33 had 3 treatments = 0.69. These are used as conditional probabilities in the simulation model.

Table A3-5 Conditional probabilities for numbers of treatments per triage categories

Category	Had 1,2,3 treatments	Had no treatment
1	0.50	0.50
2	0.58	0.42
3	0.52	0.48
4	0.52	0.48
5	0.51	0.49
Category	Had 2,3 treatments	Had no treatment
1	0.65	0.35
2	0.65	0.35
3	0.52	0.48
4	0.59	0.41
5	0.66	0.34
Category	Had,3 treatments	Had no treatment
1	0.69	0.31
2	0.68	0.32
3	0.56	0.44
4	0.56	0.44
5	0.55	0.45

A staff nurse provided estimates of treatment durations for first and subsequent treatments per triage categories, in triangular distributions, and resources required. These are displayed in Table A3-6.

Table A3.6 Estimated treatment durations for first, second and third treatments per triage categories

Category	First treatment	Resources
1	Triangular (20, 50, 100)	1 consultant, 1 junior doctor, 1 nurse
2	Triangular (20, 40, 70)	1 nurse, 1 junior doc OR 1 consultant, 1 nurse
3	Triangular (20, 40, 60)	1 nurse, 1 junior doc OR 1 consultant, 1 nurse
4	Triangular (20, 40, 60)	1 junior doc OR 1 nurse OR 1 consultant
5	Triangular (20, 40, 60)	1 junior doc OR 1 nurse practitioner
Category	Subsequent treatments	Resources
1	Triangular (15, 20, 60)	1 junior doc OR 1 consultant OR 1 nurse
2	Triangular (10, 15, 20)	1 junior doc OR 1 consultant OR 1 nurse
3	Triangular (10, 15, 20)	1 junior doc OR 1 nurse
4	Triangular (10, 15, 20)	1 junior doc OR 1 nurse practitioner OR 1 nurse
5	Triangular (10, 15, 20)	1 junior doc OR 1 nurse practitioner OR 1 nurse

From the ED data, the proportion of patient who required zero investigations, and one or more investigations were calculated (Table A3.7) and are used as conditional probabilities in the simulation. Table A3.8 shows the estimated distribution of service times for internal and external investigations, and the resource requirements.

Table A3.7 Percentages of patients requiring internal and external investigations per triage category

Category	Zero internal investigation (%)	Internal investigation (%)	Zero external investigation (%)	External investigation (%)
1	22	78	30	70
2	19	81	33	67
3	33	67	53	47
4	50	50	57	43
5	86	14	61	39

Table A3.8 Estimated investigation durations and resources per triage category

Category	Internal investigation	Resources
1	Triangular (10,15,45)	Consultant and junior doc OR consultant and nurse OR junior doc and nurse
2	Triangular (10,15,45)	Junior doc OR nurse
3	Triangular (10,15,45)	Junior doc OR nurse
4	Triangular (10,15,45)	Junior doc OR nurse OR nurse practitioner
5	Triangular (10,15,45)	Junior doc OR nurse OR nurse practitioner
Category	External investigation	Resources
1	Triangular (45,60,75)	
2	Triangular (35,60,75)	
3	Triangular (25,45,60)	
4	Triangular (45,120,190)	
5	Triangular (45,90,240)	

Table A3.9 displays patient discharges as a proportion of all discharges, by triage category. These are divided into those who are admitted, those who died in the department, those who were discharged to any destination, and those who left without treatment or refused treatment. Additionally, those who 'could have gone to MIU' are estimated by adding those who were coded as any of the following: 'Discharge – follow up treatment by GP', 'Discharge – no follow-up', 'Left department before being treated/Did not wait', and left department having refused treatment/self-discharged'. This was for later use in developing scenarios.

Table A3.9 'Disposal' destination percentage for patients by triage category

'Disposal'	Category 1 (%)	Category 2 (%)	Category 3 (%)	Category 4 (%)	Category 5 (%)
Admit	74	78	53	31	7
Discharge	11	21	45	65	84
Died	14	1	0.1	0	0
LWBS	0.5	0.5	2	5	8
Could have gone to MIU	3	8	23	35	51

Patients who ‘walk-in’ are triaged, usually be a triage nurse (nurse practitioner) but occasionally by a consultant, when ‘minors’ are busy and ‘majors’ are quiet. Estimated triage durations are in Table A3.10.

Table A3.10 Estimated triage service time and resources

Triage service time	Resources
Triangular (8,9,15)	Nurse practitioner OR consultant

As *NHSquicker* currently does not provide real-time information about admission or discharge, these delays are calculated by proportion of triage category as a mean percentage increase (Table A3.11) directly from the ED data, by comparing the mean length of stay (LoS) in ED of those given a ‘delay reason’ with the mean LoS for those without a delay. This is important because downstream delays (e.g. lack of appropriate bed, theatre delay) will increase the ED length of stay, and numbers in the department. The percentage increase, and the percentage of patients affected (per triage category) are used in the model in a ‘delay’ to replicate downstream delays for the appropriate proportion of patients.

Table A3.11 Recorded delays for discharge or admission (LoS = Length of Stay)

Category	% Delayed	Mean LoS delayed (mins)	Mean LoS no delay (mins)	**% increase**
1	29	340	137	**148**
2	32	348	161	**116**
3	28	340	159	**113**
4	21	325	151	**116**
5	7	318	121	**162**

A3.3 Discrete-Event Simulation

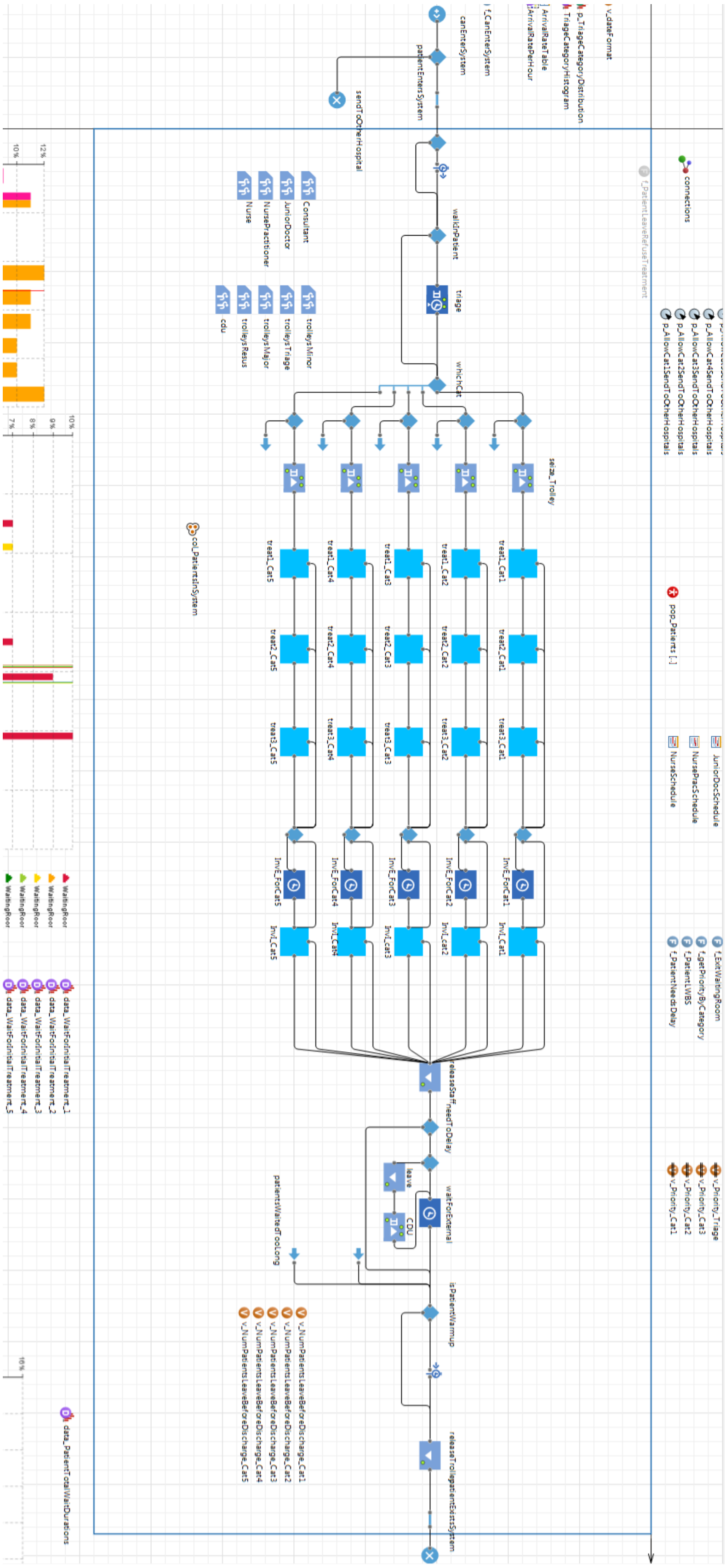


Figure A3-36 DES model (landscape)

Figure A3-36 provides a screenshot of the DES model. The following sections detail the model in detail, in three parts: (i) Model initialisation, warm-up and triage with accompanying screenshot; (ii) Patient agents and treatment blocks with accompanying screenshot; (iii) Discharge/admission delays, exit system, with accompanying screenshots.

A3.3.1 Model initialisation, warm-up and triage

The first section of the simulation is illustrated in Figure A3-37. The components are now described in detail, labelled as (a) – (i) in this subsection.

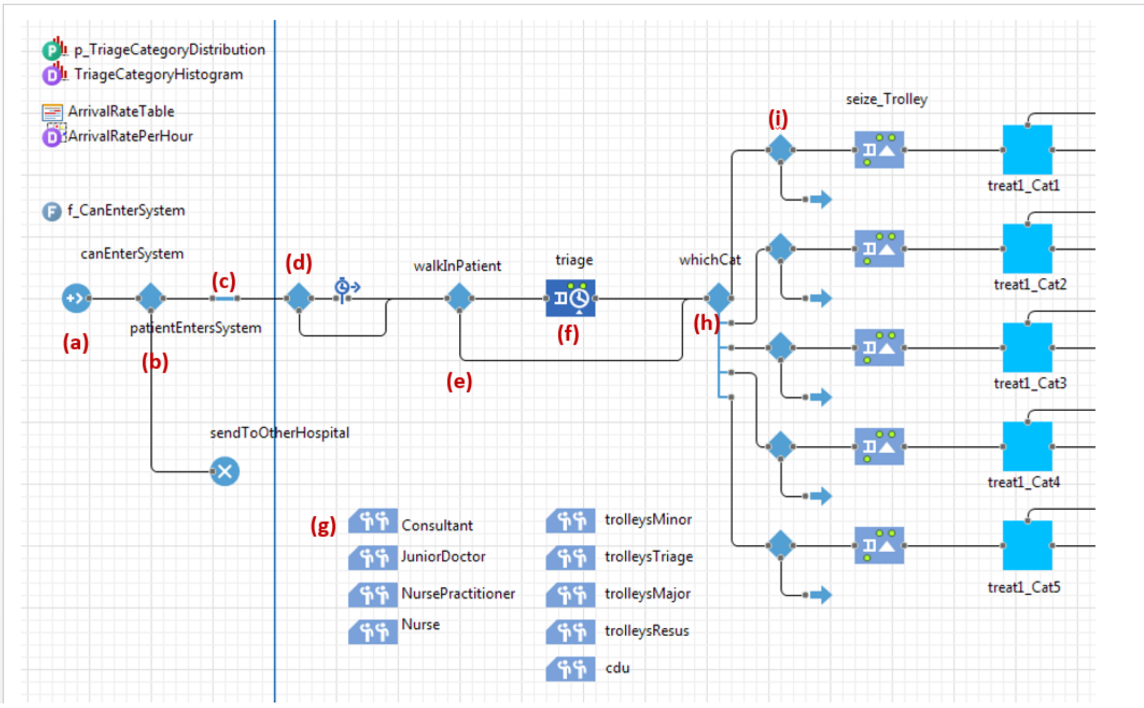


Figure A3-37 Model section 1: Patient arrivals and triage

(a) Patients enter the model according to the rate schedule defined in Table A3.2 and are allocated a triage category according to the distribution in Figure A3-34. Prior to simulation run, a user-control 'control button' and 'slider' enable the user to input 'start model at current time' and 'number of days used for warm-up'. The selected warm-up period (default = 3 days; Chapter 6, Section 6.7.1) is added to the current date-time so that data collection starts at the current date-time.

This is executed using the following code:

```
// start model with actual current time?
if (cb_StartAtSystemTime.isSelected() == true) {
	Date currentSystemDate = new Date(); // automatically gets system date
	// adjust for warmup period (start that much earlier)
	LocalDateTime currentDate = LocalDateTime.now();
	LocalDateTime startDate =
currentDate.minusDays((long)slider_warmup_Duration_days.getValue());
	Date startDateToUse =
Date.from(startDate.atZone(ZoneId.systemDefault()).toInstant());
	getEngine().setStartDate(startDateToUse);
}
else {
	// do nothing, using setup from "Model time" section above
}
```

As explained in the previous section, this is important because the simulation is intended to run for a very short time (2-4 hours), so the above code allows the model to start at the appropriate time of day and day of week in the arrival schedule, and to have the appropriate staff resources available for the particular time of day.

(b) Upon entry, a conditional 'select output' block allows patients to enter the system or be sent to another hospital. For the baseline model, this is switched off; for scenarios it is accessed at runtime using a control button and sliders. These are described in the experimentation section.

(c) The small 'plain transfer' block is used to define an action to be executed when an agent passes through this point in the model. Here, the patient is simply added to a 'collection' to keep track of the number of patients in the system at any one time. A corresponding 'plain transfer' block removes patients immediately prior to the exit block.

A parameter 'warm-up duration days' is linked to an interactive user-slider 'number of days used for warm-up' for defining the duration of warm-up. This parameter is used in a state chart to move between 'warm-up' state and 'running' state using timeout (Figure A3-38). The run time can be set in the 'running' state, and for calibration and validation this is set to 7 days, with a 3 day warm-up. After 7 days, timeout stops the model, and .finish is used to retain the results. All plots use only the 'running' data.

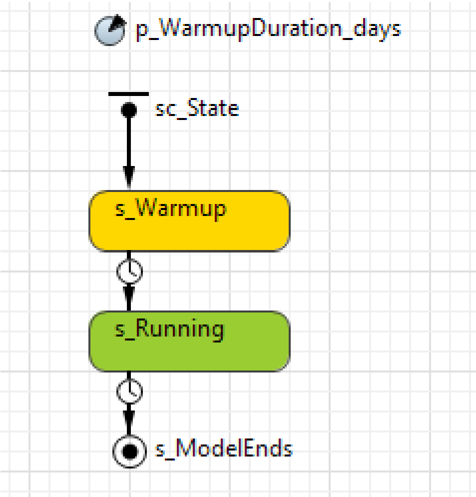


Figure A3-0-1 Setting the warm-up and running time.

(d) A ‘select output’ conditional block is placed after the ‘plain transfer’ block. If the simulation is in ‘running’ state, the patient enters a ‘Time measure start’ block to start measuring entry time of patients. If ‘false’, the patient is labelled ‘in warm-up’, and deviates around the ‘time measure start’ block. At the end of the model prior to removal from the simulation, a corresponding conditional block directs patients to ‘time measure end’ if patients are not labelled ‘in warm-up’, or bypasses the time measure if patients are ‘in warm-up’. This provides a relatively straightforward way of creating a warm-up period, and excluding it from model results, as AnyLogic does not have an easy method for setting a warm-up period.

(e) At this point, patients are defined as ‘walk-in’ or ‘arrive by ambulance’ (as per Table 6.19). This simplification is set as a probability in an output block, where those who arrive by ambulance or air ambulance bypass triage and go straight to treatment. Those who walk-in are triaged. This is acceptable as ambulance delays aren’t captured in this model due to the focus on low-acuity patients, however these are an important part of system performance and capturing ambulance handovers/delays would support a more flexible model for future work.

(f) The triage block is a ‘service block’ which seizes resources, delays the agent, and at the end of the delay, releases the resources. It contains a queue component, which initiates the start of the ‘waiting room’, which will be described shortly when the patient agent is described. The seized resources and the delay time (service time) are as defined in Table 6.18 in the previous section (Step 3). As triage can be performed by a nurse practitioner as a high priority, or by a consultant as a low priority, a set of priority variables allocate prioritisation in a simplified hierarchy, where triage priority falls between Categories 3 and 4. This means that consultants will always prioritise treatment for patients in categories 1, 2 or 3 over triage, while nurse practitioners will always prioritise triage over treatment of patients in categories 4 or 5. Following triage, patients enter a 5-point conditional output block toward treatment.

(g) Resources are defined in resource pools. Trolleys are static resources, with fixed numbers as described in Section 6.7. Staff resources are defined within an Option List, which allows each staff type to share the parameter ‘Staff_Type’ to define them as consultants, junior doctors, nurse practitioners or nurses. This is a simplification, as in practice there are other staff types (e.g. Matron, Senior Matron, healthcare assistants are all nurses; first year in practice, second year in practice, registrar are all junior doctors). The capacity for each staff type is defined using estimated schedules of three shifts/day.

(h) Patients have a single parameter (triage category) which is allocated on entry as defined in Table 6.18 in the previous section. For treatments, patients are directed down one of five pathways, according to allocated triage category. From here, the number of required treatments, treatment distribution times, and resources required are as defined in Tables 6.13 and 6.14 in the previous section. The five treatment pathways are retained for visualisation and demonstration.

(i) Five ‘output blocks’ immediately after allocation to a triage category pathway are the probability of dying in the department per triage category, and are set using sliders at model initialisation, with defaults as per Table 6.17. The assumption is that if a patient dies, it will occur before treatment starts.

A3.3.2 Patient agents and treatment blocks

Following allocation to a treatment pathway, investigations and treatments take place, according to triage category. This is illustrated in Figure A3-39 and discussed in more detail below, labelled (a) – (e) in this subsection. .

(a) Patients who enter the system seize a trolley which they hold for the duration of their treatment. The ‘seize block’ embeds a queue object where the agent waits for the resource, as treatment cannot begin until a trolley is available. Time spent in this queue is added to the ‘waiting room’, and time waiting for initial treatment starts here. Category 1 seizes a resus trolley, Categories 2 and 3 seize a majors trolley, and Categories 4 and 5 seize a minors trolley. Once the resource is seized, agents leave the block immediately, and the resource is retained until it is released, before the patient leaves the department.

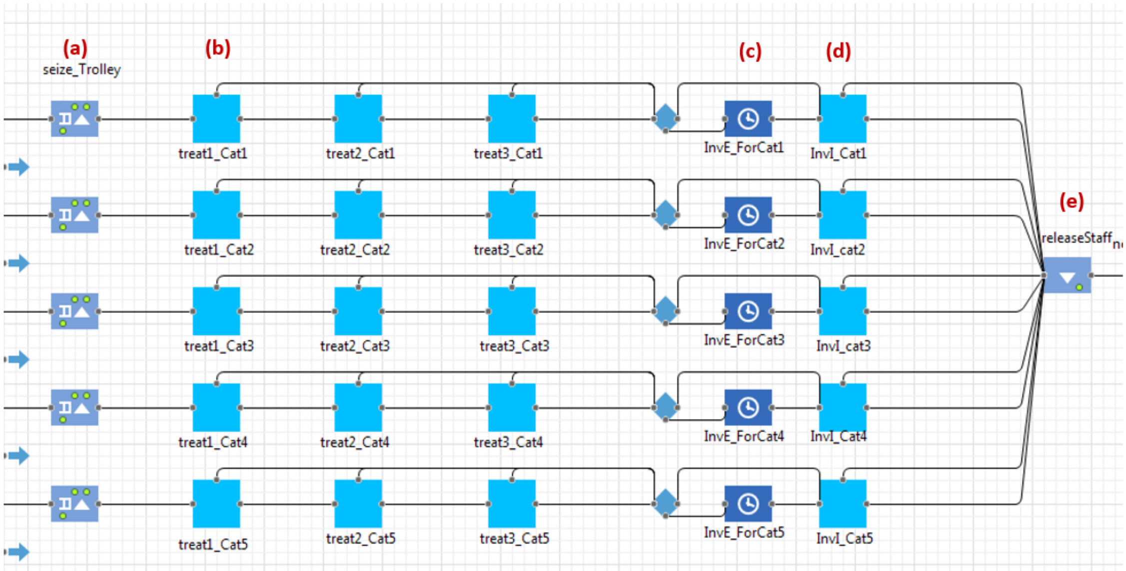


Figure A3-39 Model section 2: Treatment and investigation blocks

The ‘waiting room’ is defined as a state chart within the Patient agent (Figure A3-39) for the behavioural component ‘Leave without being seen (LWBS)’. Patients can be in either a ‘waiting’ state or a ‘not waiting’ state. The transitions between these states occur on a message trigger which occurs in the queue object of the seize trolley block. On entry to this object for each triage category, the waiting state is initiated. Once the trolley is seized (and the patient leaves the block), the waiting state is ended, and the patient re-enters the ‘not waiting’ state via a message trigger. While ‘waiting’, the variable ‘v_TimerWaitForInitialTreatment’ starts, to measure the wait time for initial treatment by category. Additionally, a function in Main counts the number of patients in the waiting room, and patients are added to ‘collections’, by triage category. Once the trolley is seized, patients exit the waiting room.

As patients undergoing treatment may be waiting for staff resources, they may also enter a ‘waiting state’ at this point. This is logged within the ‘waiting’ state using the variables ‘TotalTimeWaited_mins’ and ‘TimeStartedWait’. An internal transition within the ‘waiting’ state triggers LWBS with a timeout. This is done using the function ‘getPatientWaitLimit’. If this function (per triage category) is less than the time already spent waiting, the patient leaves. Currently the function ‘getPatientWaitLimit’ is set using parameters called within Main sliders, with defaults as per Table A3-40. However future work will investigate setting the LWBS function as a linear relationship with the real-time *Maximum Wait* Time, or *Total Patients*, as described in Chapter 6, Section 6.7.1.

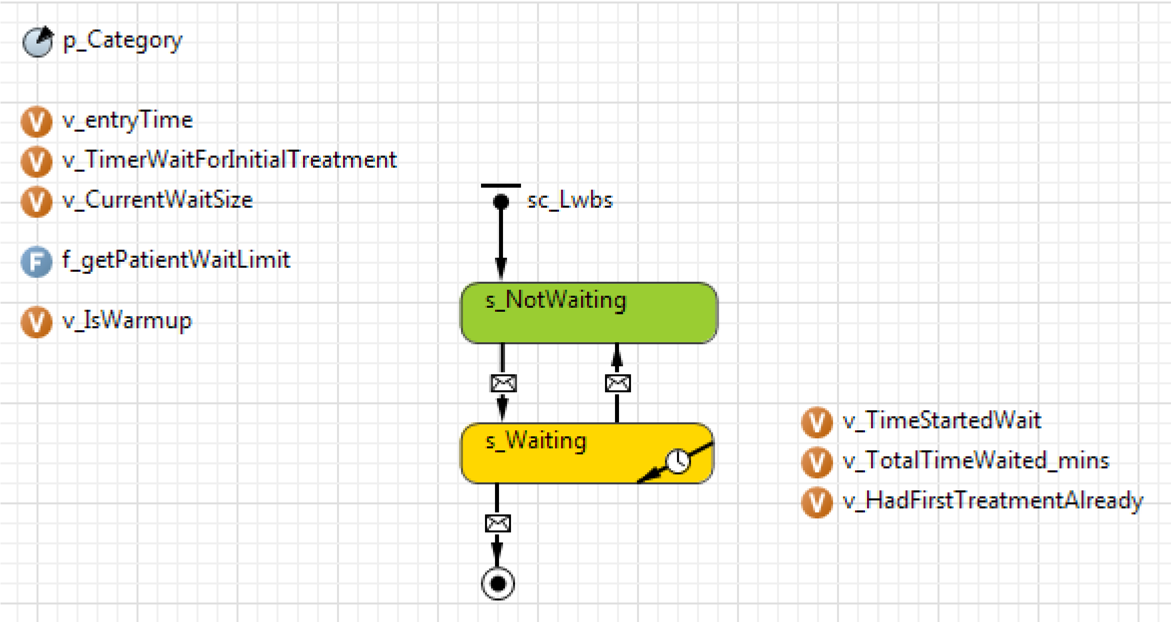


Figure A3-40 Patient agent state chart

(b) The blue 'treatment' boxes are individual sub-models which contain a range of parameters (Figure A3-41). The 'icon' is the blue square, which is the entry and exit to the sub-model. Patients leave the icon toward the output block, which calls the parameter p_Probability, to determine whether treatment occurs in this block. These are set according to the probabilities calculated in Table A3-5. If false, patients return directly to the icon and will progress to the next sub-model. If true, patients enter the 'waiting' state in the queue block within the 'seize staff' block. While waiting, if this is the patient's initial treatment (set in 'seize staff' using the patient variable v_HadFirstTreatmentAlready, set to == True), the initial treatment duration is logged using the function seen in Figure A3-40 if v_HadFirstTreatmentAlready == False. Staff are seized (p_ResourceSets), and treatment undertaken (p_Duration_min) according to Table A3-6. At the end of each treatment, staff resources are released and patients re-enter the 'icon' which returns them back to the Main model to progress to the next treatment sub-model.

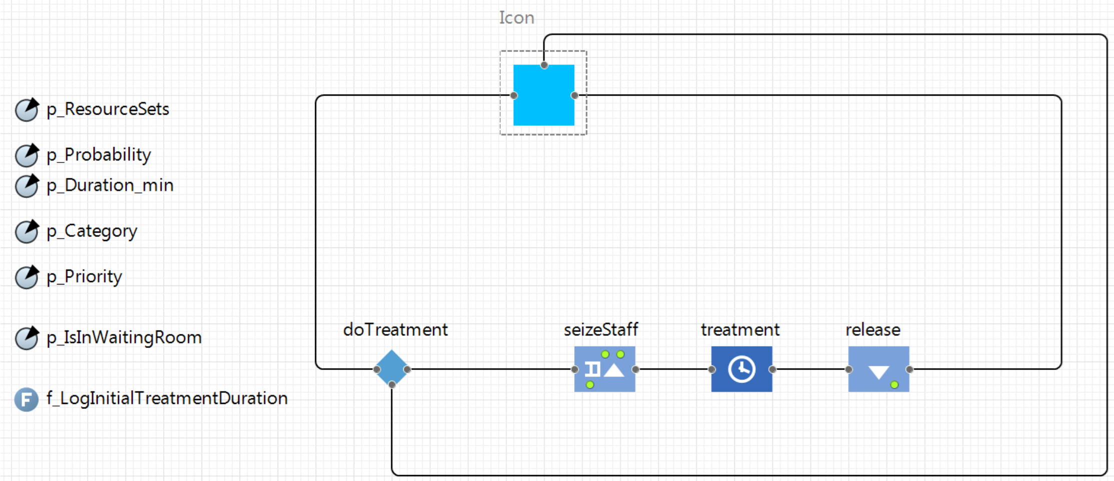


Figure A3-0-2 Treatment sub-models, with parameters

An example (first treatment for category 1 patient) is shown in Figure A3-42 to illustrate the parameters, and treatment prioritisation, set within each treatment block. Note that the parameter p_IsInWaitingRoom allows patients to return to the waiting room when waiting for treatment. This is set for Categories 4 and 5 only, and is called in the 'seize staff' block within the treatment sub-model.

(c) External investigations (e.g. Xray, ultrasound scan or other scans) use a simple 'delay' block given a probability as per Table A3-7 defined in the previous section. It is assumed that no resources are required and that the external waiting time is built into the delay time distributions, hence categories with lower prioritisation have longer delays.

(d) Internal investigations use the same treatment block sub-model as treatments, with the probability of needing one or more investigations as per Table A3-7, investigation times and resources required as per Table A3-8, and priority set by treatment category.

(e) Following all treatments and investigations, staff are released.

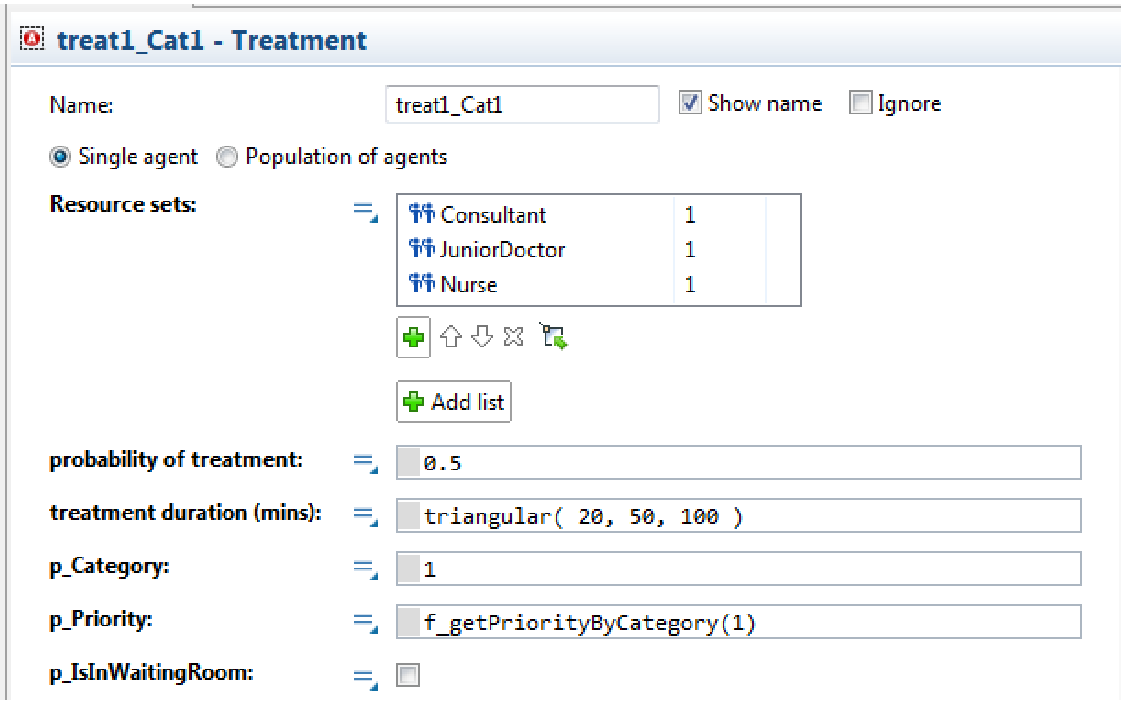


Figure A3-2 Setting the parameters for each treatment block in Main

A3.3.3 Discharge/admission delays, exit system

Figure A3-43 illustrates the final part of the model, discussed below as (a) – (e).

(a) Staff have been released however there may be delays before discharge or admission. These are coded in the ED dataset, for example: *waiting for transport, waiting for a specialist, waiting for a bed*. The proportion of patients who are delayed are set as per Table A3-11 using a function f_PatientNeedsDelay, as probabilities in an output block.

(b) Categories 4 and 5 release their trolleys and enter the CDU, seizing a CDU chair. Categories 1, 2 and 3 retain their trolleys. All enter a delay block using a probability distribution derived from Table A3-11. Where CDU is full, Categories 4 and 5 patients wait in their trolleys.

(c) Two 'exit arrows' after the delay take patients who are deceased (assumed to occur prior to any treatment taking place) and patients who LWBS. Patients who LWBS are added to the variables v_NumPatientsLeaveBeforeDischarge_Cat for plotting.

(d) Following this, patients who are not in the 'warmup' state enter the TimeMeasureEnd for plotting the lengths of stay in the department.

(e) Finally, all resources (trolleys) are released, patients are removed from the patient collection, which is an array of the number of patients in the system, and patients exit the system.

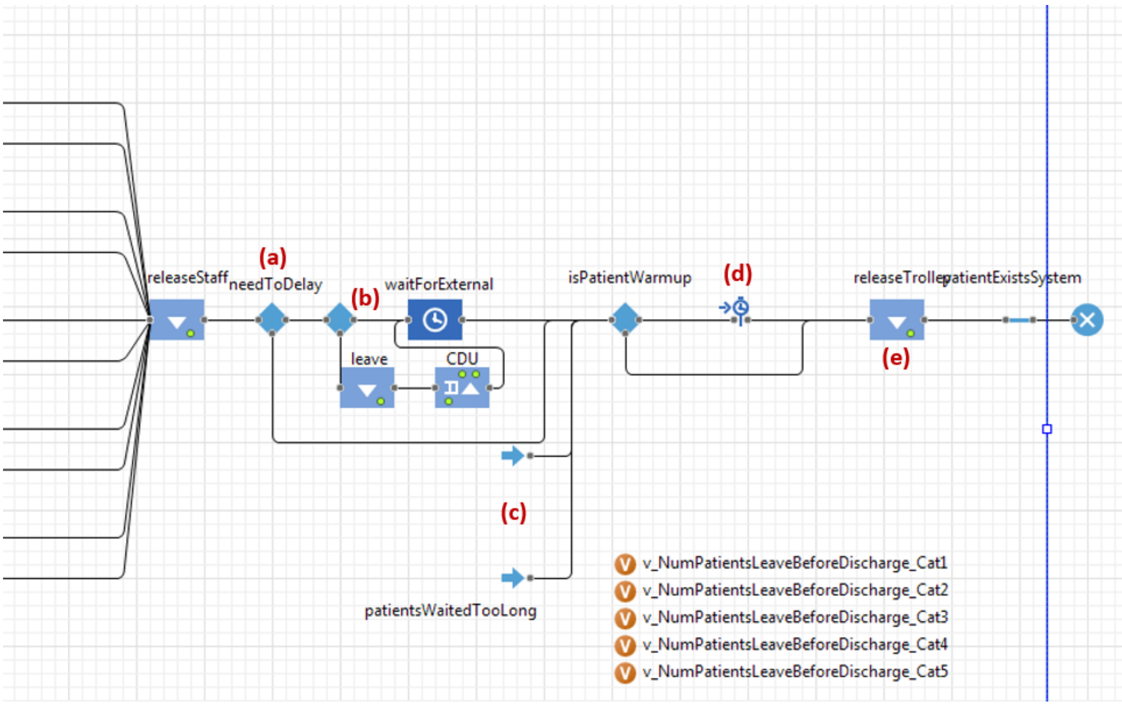


Figure A3-43 Patients release resources, may be delayed, and exit the system

A3.3.4 Validation: simulated Patients Waiting by Triage category

NHSquicker data provides Patients Waiting data overall, but not segmented by triage category. The simulation divides this output by triage category and provides 2-D histograms over-laying 150 replications. The total is a good fit with NHSquicker data, and the following graphs illustrate the findings per triage category, which align with expected values (A3-44 – A38).

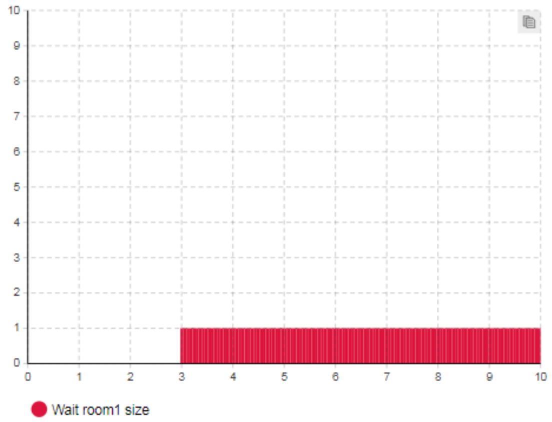


Figure A3-44 Category 1 number of patients waiting

y-axis = patient numbers, x-axis = days

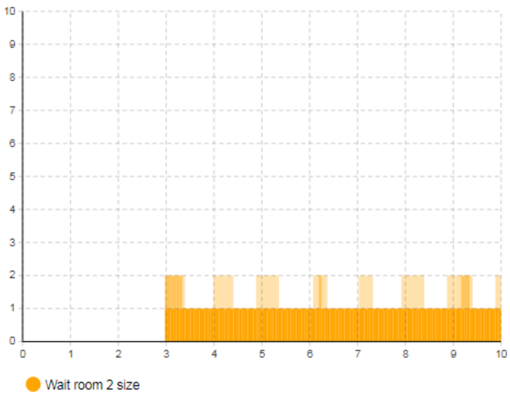


Figure A3-45 Category 2 number of patients waiting

y-axis = patient numbers, x-axis = days

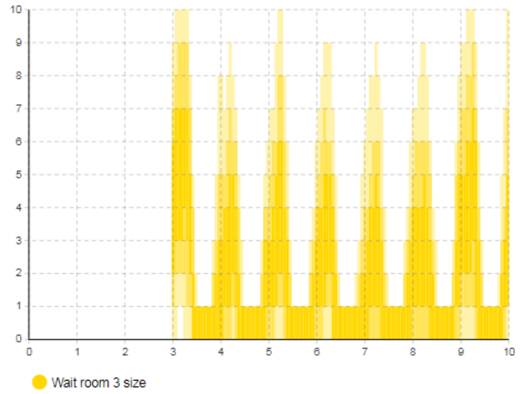


Figure A3-46 Category 3 number of patients waiting

y-axis = patient numbers, x-axis = days

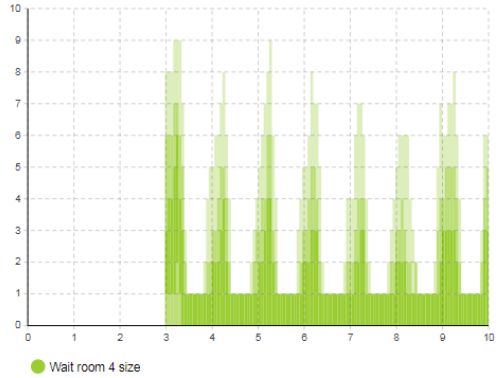


Figure A3-47 Category 4 number of patients waiting

y-axis = patient numbers, x-axis = days

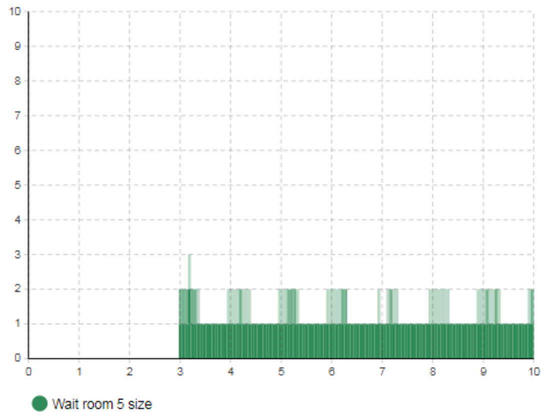


Figure A3-48 Category 5 number of patients waiting

y-axis = patient numbers, x-axis = days

Appendix 4

Participant Information Sheet

Title of Project: A Real-time Hybrid Systems Modelling Approach for Emergency Healthcare Short-Term Decision-support

Researcher name: Alison Harper

Purpose of the research:

In the NHS, data is being produced at an increasing rate and volume. Subject to the legal framework governing the use of healthcare data, "secondary uses" of data are essential for supporting a safe, efficient, and equitable health service. It is increasingly advantageous for healthcare organizations to improve performance by creating a data-driven decision-making culture.

While the NHS typically relies on historical reporting, or using data feeds in 'real time' to support rapid decision-making by enabling staff to know and react to 'what is happening'. This can improve efficiency and agility for decision-making and has repeatedly been shown to deliver value in healthcare, through process efficiency, reduced costs and improved patient care. With increasing availability of healthcare operational data, opportunities exist to develop advanced applications for real-time operational decision-support. However little is known about how useful these tools are likely to be in practice.

Why have I been approached?

For patient decision-making, applications such as NHSquicker (www.nhsquicker.co.uk), developed by the University of Exeter Business School and [redacted], provide real-time wait-time data to patients to support attendance decisions for shaping demand across the urgent care network.

This project has leveraged the value in this available real-time data to investigate the further support of short-term decisions toward controlling ED crowding using advanced analytics. It has two components:

- Forecasting of patient numbers in ED (2-4 hours ahead): Early warning of crowding may support the mobilisation of resources to prevent crowding before it occurs.
- Simulation of ED: A computer model of ED mimics the behaviour of the department. It is possible to experiment with the model, rather than the real system to determine the best escalation action to take to prevent overcrowding, identifying bottlenecks at the input, throughput or output stage.

In order to evaluate this work, and inform further research in this area, three different staff types are being invited for interview: NHS clinicians, NHS managers, and NHS IT staff. We are interested in obtaining the following information toward future work in short-term decision support applications in ED:

- How you or your department manage an escalating situation
- How you use existing real-time tools for decision support
- How you believe real-time data can support patient attendance decisions
- What kind of decision support might help, and what it might look like
- Your understanding of how forecasting and simulation might support decisions.

What would taking part involve?

Semi-structured interviews will last approximately 45 minutes to one hour. You will be asked to sign a consent form prior to commencement. You will only be identified as one of the 'staff types' outlined above, and all data will be anonymised.

What are the possible benefits of taking part?

Data-driven decision tools are showing increasing advantage for organisational decision support, but little is known about how these can be used in healthcare. This study will inform further research understanding the risks and benefits of such tools, and testing implementations in practice toward the development of more generic tools that can be more widely applied in the NHS.

What are the possible disadvantages and risks of taking part?

We don't anticipate that there are any risks associated with your participation.

What will happen if I don't want to carry on with the study?

At any point during the interview you may request to stop the interview without giving a reason and, if requested, you may withdraw from the research at any time and ask that your data be destroyed.

How will my information be kept confidential?

The University of Exeter processes personal data for the purposes of carrying out research in the public interest. The University will endeavour to be transparent about its processing of your personal data and this information sheet should provide a clear explanation of this. If you do have any queries about the University's processing of your personal data that cannot be resolved by the research team, further information may be obtained from the University's Data Protection Officer by emailing dataprotection@exeter.ac.uk or at www.exeter.ac.uk/dataprotection

Interviews will be recorded and transcribed for analysis. The transcriptions are stored on an encrypted laptop, and will be held for twelve months. All data will be anonymised, identified only as one of the three staff types described above. Access to the interview transcript will be limited to the research investigator and academic colleagues as part of the research process. The actual recording will be destroyed.

What will happen to the results of this study?

The results may be disseminated in academic publications and/or conference papers. Any summary interview content, or direct quotations from the interview, that are made available through academic publication or other academic outlets will be anonymised so that you cannot be identified, and care will be taken to ensure that other information in the interview that could identify you is not revealed.

Who is organising and funding this study?

This research is funded by the Economic and Social Science Research Council (ESRC) via the SWDTP.

Who has reviewed this study?

This project has been reviewed by the University of Exeter Business School Research Ethics Committee (Reference Number eUEBS000905), and the researcher holds an honorary contract at [redacted] until September 2020.

Further information and contact details:

Alison Harper

Centre for Simulation, Analytics and Modelling, University of Exeter Business School, Stocker Road, Exeter, EX4 4PY

Email: ah596@exeter.ac.uk

Tel: 07922 109 779

If you are not happy with any aspect of the project and wish to complain, please contact:

Professor Navonil Mustafee,

Email: n.mustafee@exeter.ac.uk

Gail Seymour, Research Ethics and Governance Manager

g.m.seymour@exeter.ac.uk, 01392 726621

Thank you for your interest in this project

UNIVERSITY OF EXETER

CONSENT FORM

Title of Project:

A Real-time Hybrid Systems Modelling Approach for Emergency Healthcare Short-Term Decision-support

Name of Researcher: Alison Harper

Please initial box

1. I confirm that I have read the information sheet dated 08/03/2020 (V2) For the above project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without my legal rights being affected. ☐

3. I understand that relevant sections of the data collected during the study, may be looked at by members of the research team, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

4. I understand that taking part involves anonymised interview transcripts to be used for the purposes of analysis, which will be held securely for a period of up to on year. These may be published in an academic publication ☐

5. I agree to take part in the above project. ☐

______________________	______________________	______________
Name of Participant	Date	Signature
______________________	______________________	______________
Name of researcher	Date	Signature

Interview Guide

<u>THINK ABOUT THE LAST TIME ED WAS UNDER SIGNIFICANT PRESSURE and you felt that there were potential risks to patient safety</u>:

(Clinician – first-hand; Manager/IT – org. strategies)

1. **Your behaviour to manage an escalating situation:**

- Describe the nature of the situation
- What did the organisation do, how did it respond?
- How does it know when to make a decision? (triggers)
- What are the goals at the various decision points?
- Are there any situations in which the decision would have turned out differently?
- What information is available at the time of the decision?
- At any stage, were you uncertain about either the reliability or the relevance of the information that you had available to help with your decision?
- What was the most important piece of information that you used to help with the decision?
- Was there any stage during the decision-making process in which you found it difficult to process and integrate the information available?
- At any stage, were you uncertain about the appropriateness of the decision?
- Were there any other alternatives available to you other than the decision you made?
- What other escalation decisions were taken by others?
- How does your behaviour change if you think it might get busier later on?

2. **Existing tools and policies:**

a. **Technology**

i. **Forecasting**

ii. **Real-time data –** eg Overcrowding tool, Symphony, Ambulance

iii. **Simulation/OR model –** eg Excel spreadsheet

b. **Policies**

i. What escalation policies currently implemented based on this?

ii. Larger policy-based decisions, smaller hour-to-hour changes in behaviour?

- Overcrowding tool – how is it currently used?
- How do you currently use it?
- What other decision-support is available to you?

3. NHSquicker

- Do you, or other staff you are aware of, use it?
- How do you use it?
- Does it help staff decision making?
- Reliable and accurate?
- Supports the wider system?
- Unintended consequences for staff?
- Any indirect consequences, positive or negative, ST or LT?
- How could it be made more useful for staff?
- Which staff use it? Which staff should be using it? What stops people using it?

4. Evaluate IHAF

a. Diagnostic measure

- Does Total Number in the department represent overcrowding to staff? Is it useful or meaningful to you?
- What would be better?

b. Predictive measure –

- How might you use short-term forecasts of Total Patients?
- What escalations could be used in 2-4 hours? What isn’t possible in that time?
- What forecasts would be most useful?
- What would you do with them?
- Does something need to change to enable this?
- Which staff would use it? Which staff should be using it?
- What would make it not useful?
- Any risks or unexpected or indirect consequences? ST or LT?

- Any barriers to use?

c. Prescriptive measure:

- Who might use this tool?
- Would you use it? Do you want it?
- What escalation actions are possible – input, throughput, output?
- What actions would you take?
- What are the barriers? How could these be tackled?
- What might its impact be? Negative or positive? ST/LT?
- Could it be useful more widely, beyond ED, or beyond the hospital? Are there possible contextual differences?
- How might it help patients?
- How might it help the wider system?

5. Ideal decision support tool

- What would it consist of? What data? Where would it sit? How would it be used? Who would use it? What would it tell you?
- How best to access it, eg in dashboard, app, website, other?
- How could this knowledge be used? What escalation activities? How far ahead would be useful?
- What are the barriers? Technical, organisational, other. How can they be tackled?
- Who wouldn't want it? Why?
- What would influence your trust in the tool?

6. Patient decision-making: NHSquicker, patient:

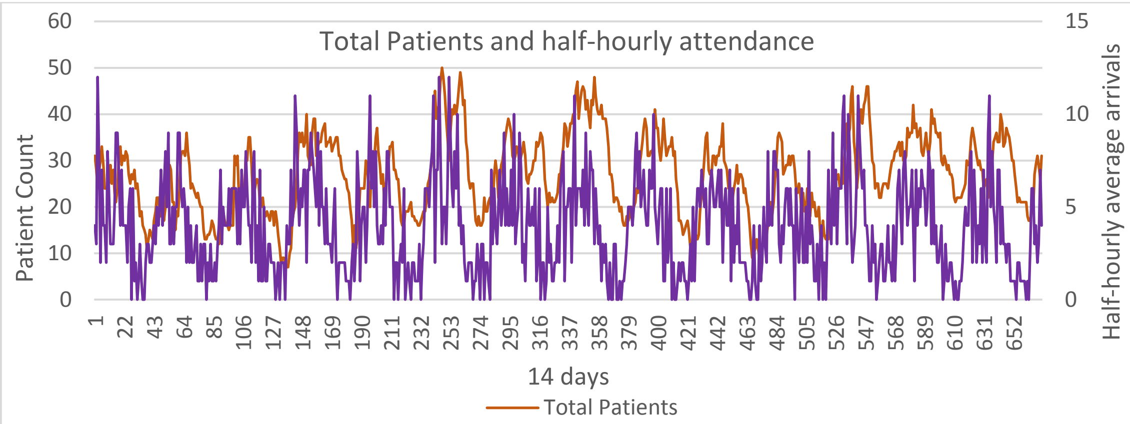
- Are you aware of any patients who have used this information?
- How do you feel about patients having access to additional information to support attendance decisions? How do others feel?
- Forecasts? Other more useful info for attendance decisions?
- Do you believe it changes patient attendance behaviour?
- Do you believe it has benefits at the system level?
- Any unintended or indirect consequences? ST or LT?
- How could it be made more useful for patients? Immediate and LT

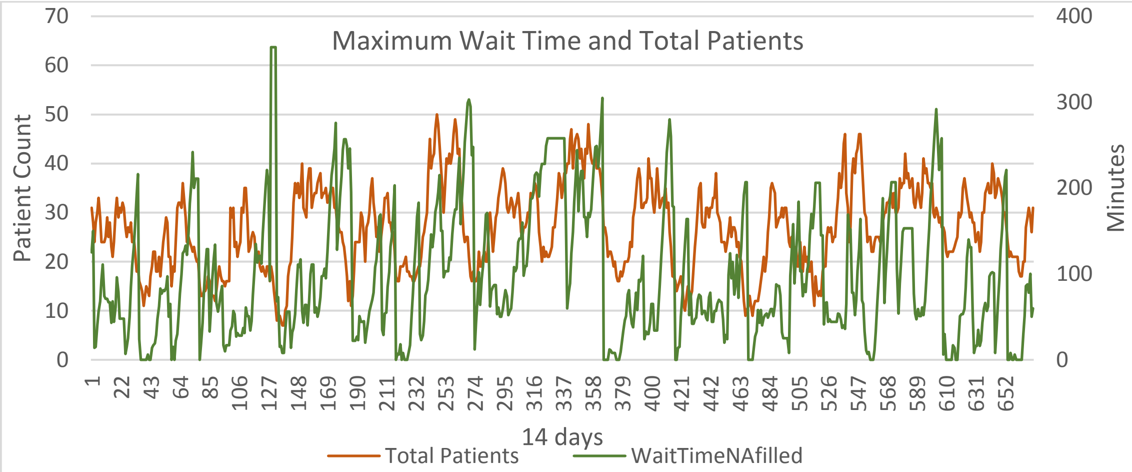
- Given questionnaire information, how do you feel about patients having access to additional information to support attendance decisions? Unexpected consequences? Barriers to use?

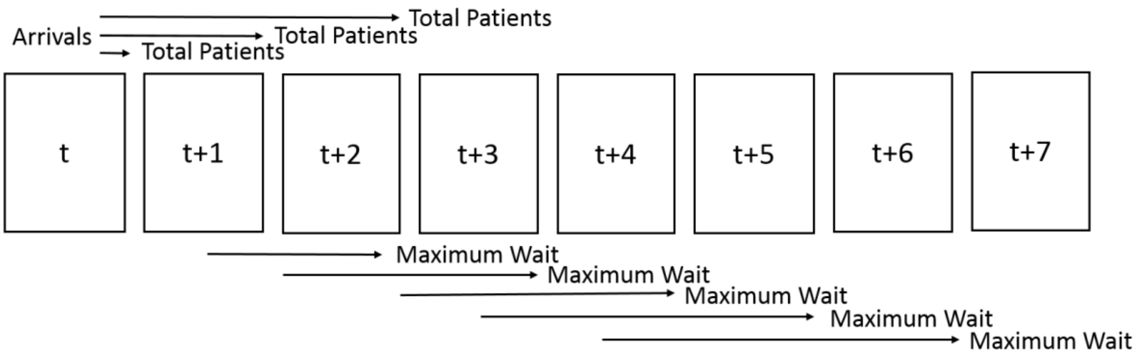
	General	SA	Real-time Data	Existing Forecasts	Simulation /Excel	IHAF	Triggers
Managing ED at peak periods (NHS)	X	X				=>	=>
Existing Tools		X	X	X	X	=>	=>
Existing Policies	X				X	=>	X
NHSquicker and Forecasts		X	X	X			X
Ideal real-time tool		X				X	=>
(NHSquicker) - Patients	X		X			X	=>

Interview props for demonstration and evaluation

<u>Trigger: Total Patients</u>

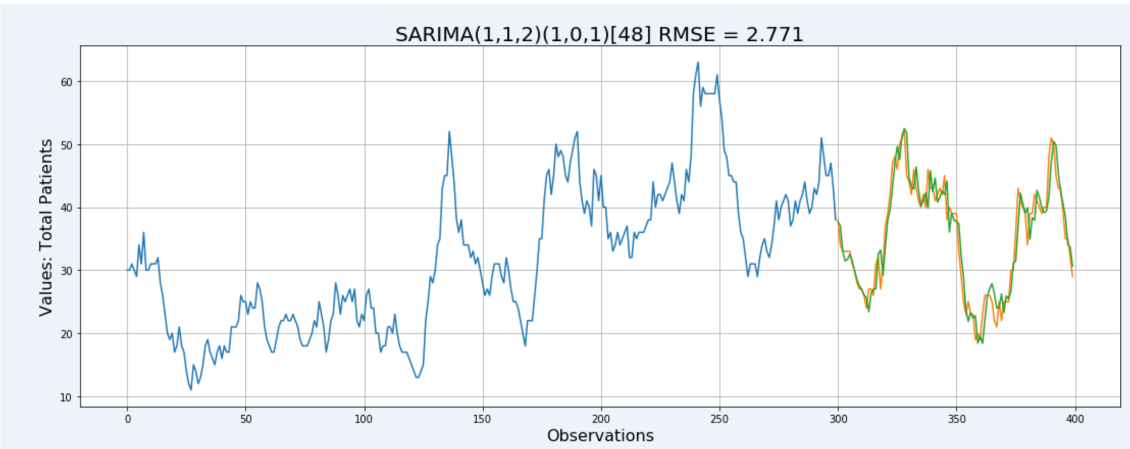


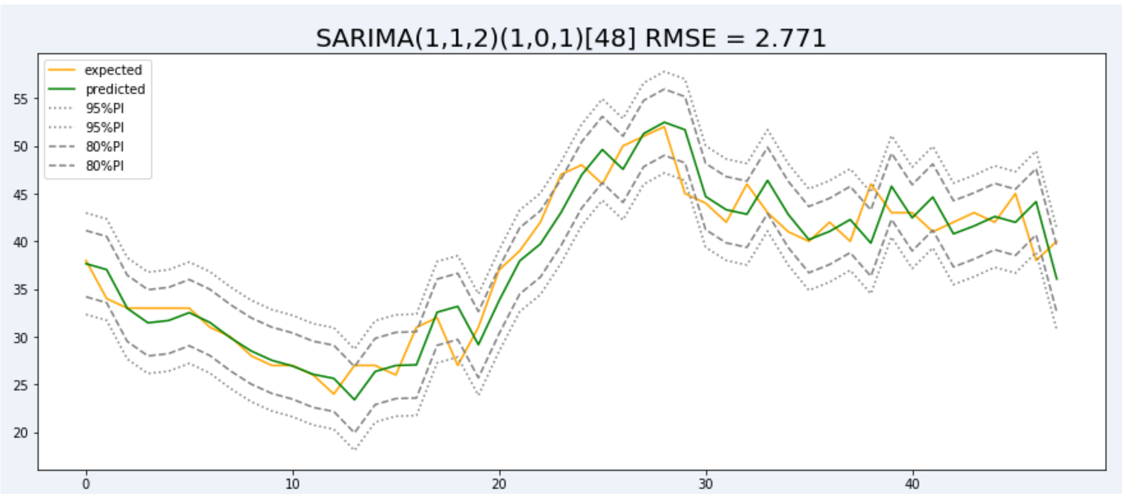




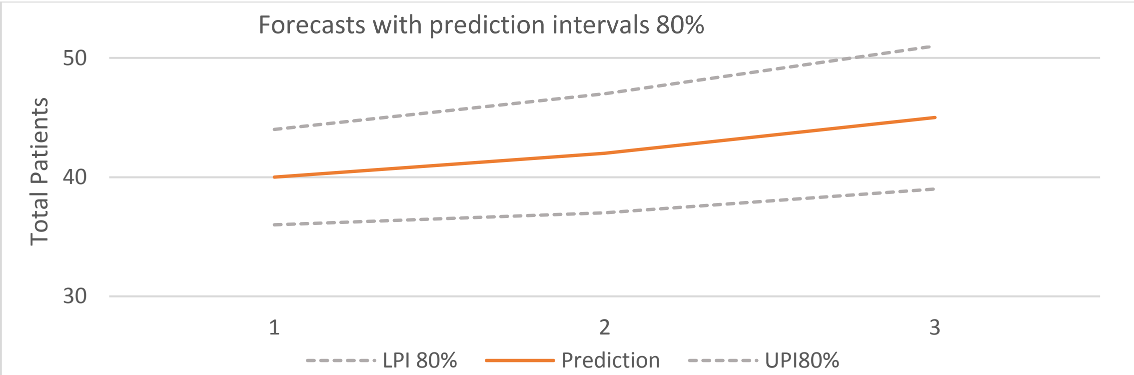
Hour	00.00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00
1.5 sd	41	39	36	33	30	29	28	27	28	31	35	39
Hour	12.00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
1.5 sd	42	44	46	47	46	45	47	48	48	48	47	45

Forecasts: Total Patients

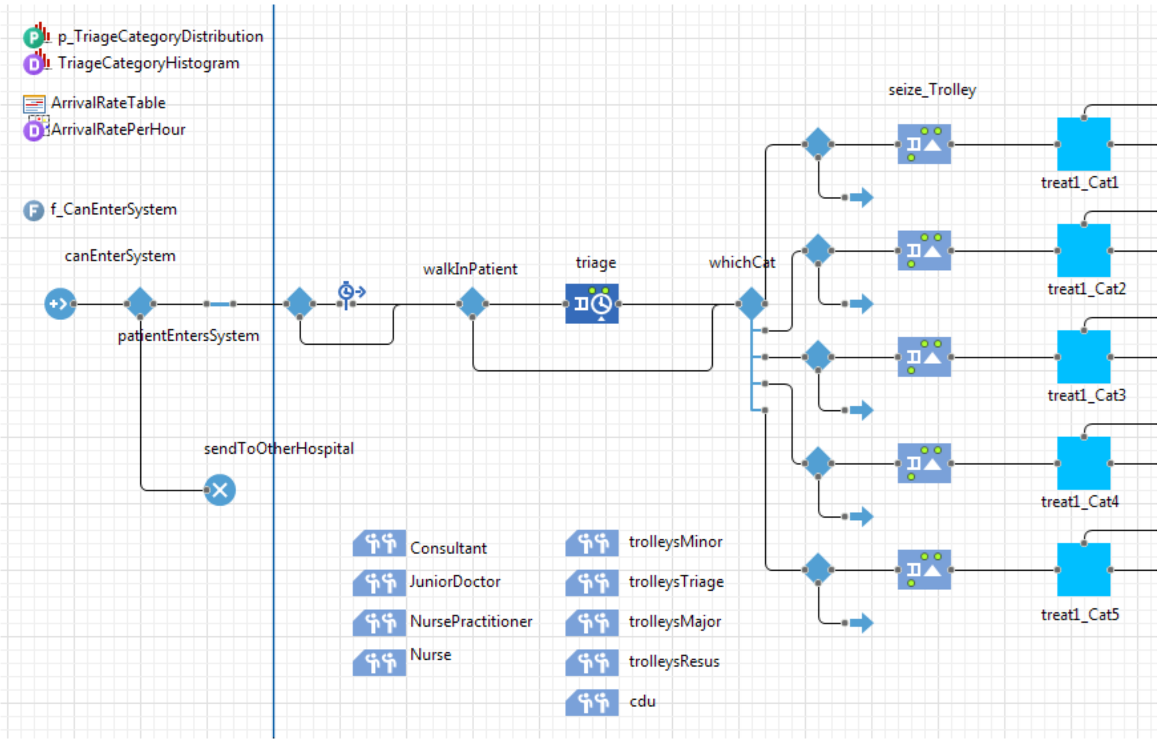


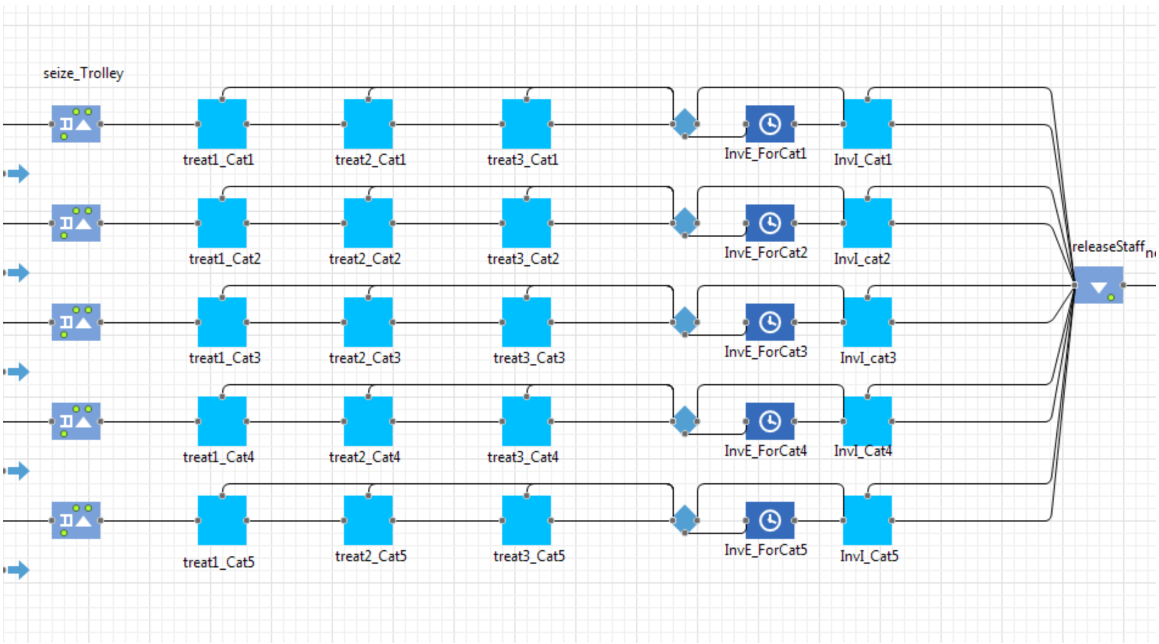


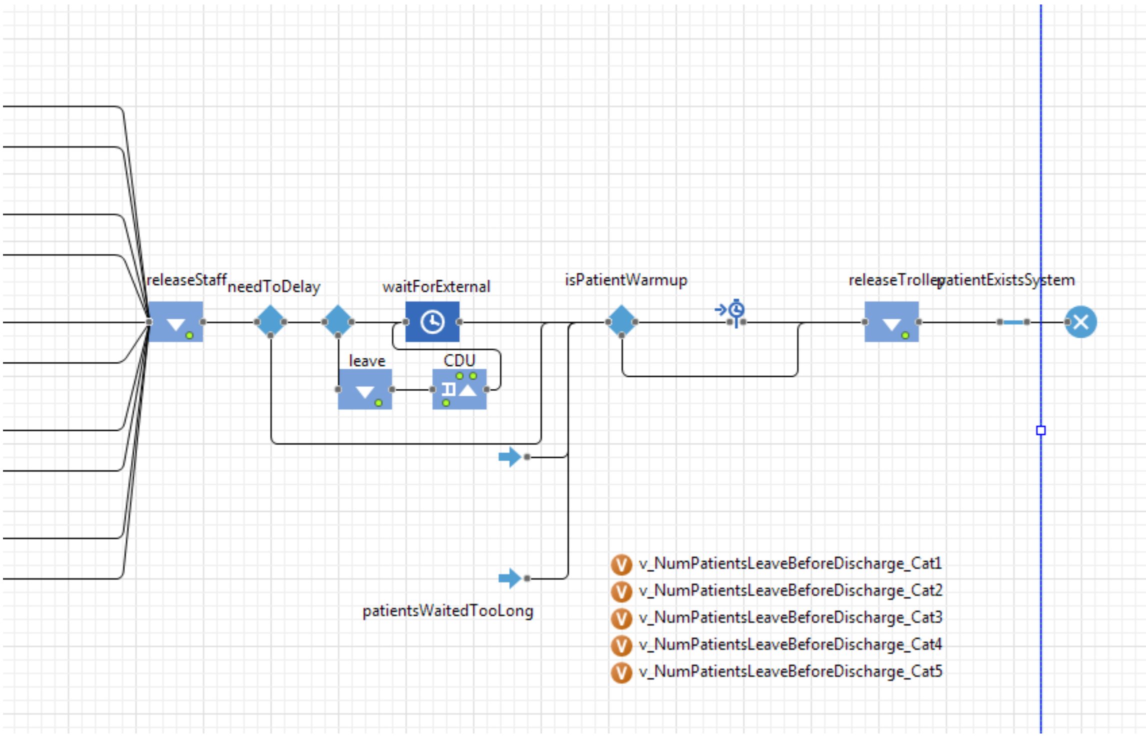
	Forecast	PI (80%)
1 hour forecast	40	36 - 44
2 hour forecast	42	37 - 47
4 hour forecast	44	38 - 50



Simulation: Real time

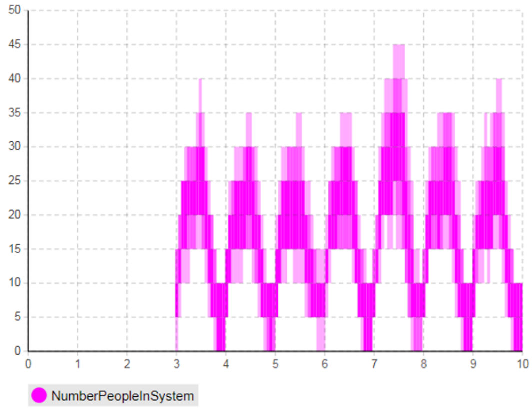




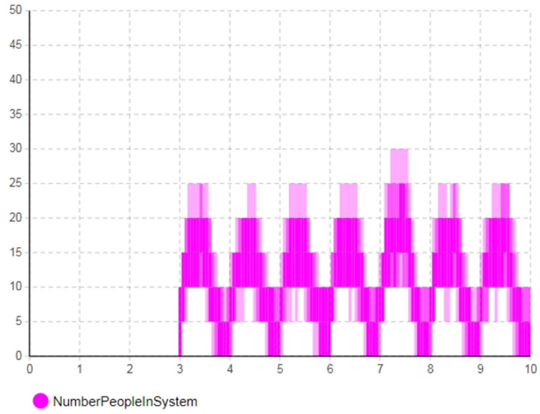


- Baseline – proportion of patients LWBS (leave without being seen) per triage category, calibrated to 2018 data
- Scenario 1 – Redirect all Cat 4 and 5 patients when number of patients in department reaches hourly trigger
- Scenario 2 – Redirect a proportion of Cat 4 and 5 patients to MIU when number of patients in department forecasted to reach hourly trigger in 2-4 hours' time

- Scenario 3 – Redirect a proportion of Cat 4 and 5 patients to MIU when number of patients in department forecasted to reach hourly trigger in 2-4 hours' time, and given sufficient capacity in MIU

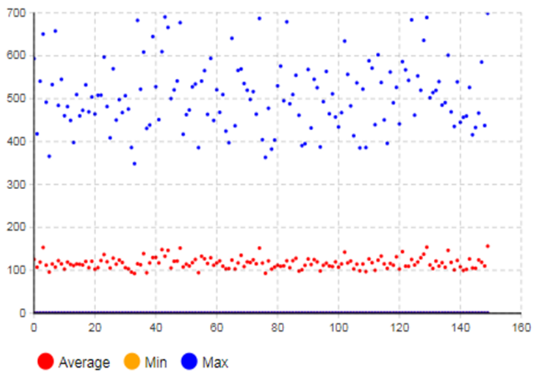


Baseline 2D histogram Total number Department

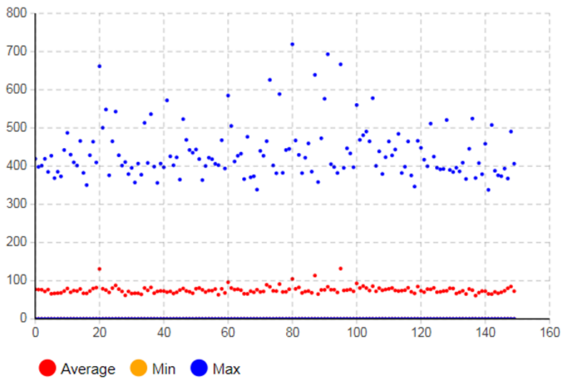


Redirect 50% Cat 5; 30% Cat 4; 15% Cat 3

Preliminary results: 150 repetitions
3 day warmup
7 day runtime



Baseline Length of Stay in Department (min, mean, max)



Redirect 50% Cat 5; 30% Cat 4; 15% Cat 3

References

Abend, G. (2008). The meaning of ‘theory’. *Sociological theory*, *26*(2),173-199.

Abdelbari, H. & Shafi, K. (2017). A computational Intelligence-based Method to ‘Learn’ Causal Loop Diagram-like Structures from Observed Data. *System Dynamics Review*, *33*(1), 3–33.

Abir, M., Goldstick, J.E., Malsberger, R. *et al.* (2019). Evaluating the impact of emergency department crowding on disposition patterns and outcomes of discharged patients. *Int J Emerg Med* 12**,** 4

Aboagye-Sargo, P., Mai, Q. Sanfilippo, F. M., Preen, D. B., Stewart, L. M. & Fatovich. D. M. (2015). A Comparison of Multivariate and Univariate Time-Series Approaches to Modelling and Forecasting Emergency Department Demand in Western Australia. *Journal of Biomedical Informatics* 57,62-73.

Abohamad, W., & Arisha, A. (2013). Simulation-based framework to improve patient experience in an emergency department. *European Journal of Operational Research*, 224(1), 154–166.

Abohamad, W., Ramy, A., & Arisha, A. (2017). A hybrid process-mining approach for simulation modeling. In *2017 Winter Simulation Conference (WSC)* (pp. 1527-1538). IEEE.

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

Aceto, G., Persico, V. & Pescapé, A. (2020). Journal of Industrial Information Integration Industry 4.0 and Health : Internet of Things , Big Data , and Cloud Computing. *Journal of Industrial Information Integration*. https://doi.org/10.1016/j.jii.2020.100129

Acito, F. & Khatri, V. (2014). Business analytics : Why now and what next ? *Business Horizons*, *57*(5), 565–570.

Ackermann, F., & Eden, C. (2011). Strategic management of stakeholders: Theory and practice. *Long range planning*, *44*(3), 179-196.

Adra, H. (2016). Real-time predictive and prescriptive analytics with real-time data and simulation. In *2016 Winter Simulation Conference (WSC)* (pp. 3357–3364). IEEE.

Afilalo, J., Marinovich, A., Afilalo, M., Colacone, A., Leger, R., Unger, B., & Giguere, C. (2004). Nonurgent emergency department patient characteristics and barriers to primary care. *Academic Emergency Medicine*, *11*(12), 1302-1310.

Agarwal, S., Banerjee, J., Baker, R., Conroy, S., Hsu, R., Rashid, A., … Habiba, M. (2012). Potentially avoidable emergency department attendance: Interview study of patients’ reasons for attendance. *Emergency Medicine Journal*, *29*(12), 2009–2012.

Ågerfalk, P. J. (2013). Embracing diversity through mixed methods research. *European Journal of Information Systems*, *22*(3), 251-256.

Ahalt, V., Argon, N. T., Ziya, S., Strickler, J., & Mehrotra, A. (2018). Comparison of emergency department crowding scores: a discrete-event simulation approach. *Health care management science*, *21*(1), 144-155.

Alavi, M., & Leidner, D. (1999). Knowledge management systems: issues, challenges, and benefits. *Communications of the Association for Information systems*, *1*(1), 7

Albert, S. & Bradley, K. (1997). *Managing Knowledge: Experts, Agencies and Organisations*. Cambridge University Press.

Allotey, P., Reidpath, D. D., Ghalib, H., Pagnoni, F., & Skelly, W. C. (2008). Efficacious, effective, and embedded interventions: implementation research in infectious disease control. *BMC Public Health*, *8*(1), 1-6.

Almagooshi, S. (2015). Simulation Modelling in Healthcare: Challenges and Trends. *Procedia Manufacturing*, *3*, 301-307

Altaf, M. S., Liu, H., Al-Hussein, M., & Yu, H. (2016). Online simulation modeling of prefabricated wall panel production using RFID system. In *Proceedings of 2016 Winter Simulation Conference* (WSC) *(pp.* 3379–3390). IEEE.

Alvanchi, A., Lee, S., & AbouRizk, S. (2011). Modeling framework and architecture of hybrid system dynamics and discrete event simulation for construction. *Computer‐Aided Civil and Infrastructure Engineering*, *26*(2), 77-91.

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. Academy of Management Review, 36(2), 247-271.

Alvesson, M., & Sandberg, J. (2020). The problematizing review: A counterpoint to Elsbach and Van Knippenberg's argument for integrative reviews. Journal of Management Studies. https://doi.org/10.1111/joms.12582

Alvesson, M., & Willmott, H. (Eds.). (2003). *Studying management critically*. Sage.

Alvesson, M., & Willmott, H. (2012). *Making sense of management: A critical introduction*. Sage.

Alyasin, A., & Douglas, C. (2014). Reasons for non-urgent presentations to the emergency department in Saudi Arabia. *International Emergency Nursing*, *22*(4), 220–225.

Amalberti, R., & Vincent, C. (2020). Managing risk in hazardous conditions: improvisation is not enough. *BMJ quality & safety*, *29*(1), 60-63.

Anandaciva, S. (2019). Five reasons why this NHS winter may be different. The Kings Fund. Available at: https://www.kingsfund.org.uk/blog/2019/12/five-reasons-why-nhs-winter-may-be-different?utm_source=twitter&utm_medium=social&utm_term=thekingsfund. Accessed: 22.06.19

Andersen, R. M. (1968). Behavioral Model of Families' Use of Health Services. Research Series No. 25. Chicago, IL: Center for Health Administration Studies, University of Chicago.

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: does it matter? *Journal of health and social behavior*, 1-10.

Andersen, R. M. (2008). National health surveys and the behavioral model of health services use. *Medical care*, *46*(7), 647-653.

Andersen, R. M., Rice, T. H., & Kominski, G. F. (2011). *Changing the US health care system: Key issues in health services policy and management*. John Wiley & Sons.

Andersen, R. M., Davidson, P. L., & Baumeister, S. E. (2014). Improving access to care. *Changing the US health care system: key issues in health services policy and management. San Francisco: Jossey-Bass*, *36*(3), 33-69.

Andrews, H., & Kass, L. (2018). Non-urgent use of emergency departments: populations most likely to overestimate illness severity. *Internal and Emergency Medicine*, *13*(6), 1–8.

Annan, S. M., & Banks, J. (1992). Design of a Knowledge-Based On-Line Simulation System to Control a Manufacturing Shop Floor. *IIE Transactions (Institute of Industrial Engineers)*, *24*(3), 72–83.

Archer, M. S., & Archer, M. S. (2003). *Structure, agency and the internal conversation*. Cambridge University Press.

Archer, M. S. (Ed.). (2013). *Social morphogenesis*. Springer Science & Business Media.

Archer, M. S. (2016). Reconstructing sociology: The critical realist approach. *Journal of Critical Realism*, *15*(4), 425-431.

Archer, M., Bhaskar, R., Collier, A., Lawson, T., & Norrie, A. (Eds.). (2013). *Critical realism: Essential readings*. Routledge.

Artman, H., & Garbis, C. (1998). Situation awareness as distributed cognition. In *Proceedings of ECCE* (Vol. 98). Citeseer.

Ash, J. S., Gorman, P. N., Lavelle, M., Stavri, P. Z., Lyman, J., Fournier, L., & Carpenter, J. (2003). Perceptions of physician order entry: results of a cross-site qualitative study. *Methods of information in medicine*, *42*(04), 313-323.

Ash, J. S., Sittig, D. F., Seshadri, V., Dykstra, R. H., Carpenter, J. D., & Stavri, P. Z. (2005). Adding insight: a qualitative cross-site study of physician order entry. *International journal of medical informatics*, *74*(7-8), 623-628.

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

Atenstaedt, R., & Evans, K. (2015). Emergency departments: Better safe than sorry? *Emergency Nurse*, *23*(4), 20–22.

Augusto, V., Xie, X., Prodel, M., Lamarsalle, L., & Baptiste, J. (2016). Evaluation of discovered clinical pathways using process mining and joint agent-based discrete-event simulation. In T. M. K. Roeder; P. I. Frazier; R. Szechtman; E. Zhou; T. Huschka; and S. E. Chick; (Ed.), *Proceedings of the 2016 Winter Simulation Conference* (pp. 2135–2146). IEEE.

Augusto, V., Viallon, A., & Murgier, M. (2018). A modelling and simulation framework for intelligent control of emergency units in the case of major crisis. In *Proceedings of the 2018 Winter Simulation Conference M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds*. (pp. 2495 – 2506). IEEE.

Aydt, H., Turner, S. J., Cai, W., & Low, M. Y. H. (2008a). Symbiotic simulation systems: An extended definition motivated by symbiosis in biology. *Proceedings - Workshop on Principles of Advanced and Distributed Simulation, PADS*, *b*, 109–116.

Aydt, H., Turner, S. J., Cai, W., Low, M. Y. H., Lendermann, P., Gan, B. P., & Ayani, R. (2008b). Preventive What-If Analysis in Symbotic Simulation. In J. W. F. eds. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson (Ed.), In *Proceedings of the 2008 Winter Simulation Conference* (pp. 750–758). IEEE.

Aydt, H., Turner, S. J., Cai, W., & Low, M. Y. H. (2009a). Research issues in symbiotic simulation. In *Proceedings 2009 Winter Simulation Conference*, (pp.1213–1222). IEEE.

Aydt, H., Turner, S. J., Cai, W., & Ayani, R. (2009b). Symbiotic Simulation Model Validation for Radiation Detection Applications. *Proceedings of the 2009 ACM/IEEE/SCS 23rd Workshop on Principles of Advanced and Distributed Simulation*, 11–18.

Babitsch, B., Gohl, D., & Von Lengerke, T. (2012). Re-revisiting Andersen's Behavioral Model of Health Services Use: a systematic review of studies from 1998–2011. *GMS Psycho-Social-Medicine*, *9*.

Back, J., Ross, A. J., Duncan, M. D., Jaye, P., Henderson, K., & Anderson, J. E. (2017). Emergency Department Escalation in Theory and Practice: A Mixed-Methods Study Using a Model of Organizational Resilience. *Annals of Emergency Medicine*, *70*(5), 659–671.

Backman, A. S., Blomqvist, P., Lagerlund, M., Carlsson-Holm, E., & Adami, J. (2008). Characteristics of non-urgent patients. *Scandinavian Journal of Primary Health Care*, *26*(3), 181–187.

Backman, A. S., Blomqvist, P., Svensson, T., & Adami, J. (2010). Health care utilization following a non-urgent visit in emergency department and primary care. *Internal and Emergency Medicine*, *5*(6), 539–546.

Badri, M. A., & Hollingsworth, J. (1993). A Simulation Model for Scheduling in the Emergency Room. *International Journal of Operations & Production Management*, *13*(3), 0-0.

Bae, J., Bae, H. Kang, S. & Kim, Y. (2004). Automatic Control of Workflow Processes Using ECA Rules. *IEE Transactions on Knowledge and Data Engineering*. 16(8): 1010-1023.

Bahrani, S., Tchemeube, R. B., Mouttham, A., & Amyot, D. (2013). Real-time simulations to support operational decision making in healthcare. In *Proceedings of the 2013 Summer Computer Simulation Conference* (p. 53). Society for Modeling & Simulation International.

Baker, C. (2017) Accident and Emergency Statistics: Demand, Performance and Pressure. House of Commons Library Briefing Paper, Number 6964. Available at: https://commonslibrary.parliament.uk/research-briefings/sn06964/ Accessed: 03.09.19

Baker, S.E. & Edwards, R. (2012) *How many qualitative interviews is enough.* Discussion Paper. National Centre for Research Methods (NCRM).

Balci, O. (1989). How to assess the acceptability and credibility of simulation results. In *Proceedings of the 21st Conference on Winter Simulation (WSC) (pp.* 62–71). IEEE.

Balci, O., & Nance, R. E. (1985). Formulated problem verification as an explicit requirement of model credibility. *Simulation*, *45*(2), 76-86.

Balci, O. (1990) Guidelines for Successful Simulation Studies. In *Proceedings of the 1990 Winter Simulation Conference,* edited by O. Balci et al. (pp. 25-32). IEEE.

Balci, O. (2012). A life cycle for modeling and simulation. *Simulation*, *88*(7), 870-883.

Banks, J., Carson, J., Nelson, B.L. & Nicol, D.M. (2001) *Discrete-event Simulation*. Prentice Hall, NJ.

Barnes, S., Hamrock, E. Toerper, M. Siddiqui S. & Levin. S. (2016). Real-time Prediction of Inpatient Length of Stay". *J.Am.Med.Inform.Assoc.* 23(1): 2-10.

Batalden, P. B., & Davidoff, F. (2007). What is "quality improvement" and how can it transform healthcare? *Quality and Safety in Health Care*, *16*(1), 2–3.

Batalden, M., Batalden, P., Margolis, P., Seid, M., Armstrong, G., Opipari-Arrigan, L., & Hartung, H. (2016). Coproduction of healthcare service. *BMJ Quality and Safety*, *25*(7), 509–517.

Batarseh, F. A., & Latif, E. A. (2015). Assessing the Quality of Service Using Big Data Analytics. With Application to Healthcare. *Big Data Research*, *1*, 1–12.

Beach, L. R., M. Chi, G. Klein, P. Smith, and K. Vicente (1997). 'Naturalistic decision making and related lines' in Naturalistic decision making. C. E. Zsambok and G. Klein (eds). Mahwah, NJ: Lawrence Erlbaum.

Beache, S. K., & Guell, C. (2016). Non-urgent accident and emergency department use as a socially shared custom: A qualitative study. *Emergency Medicine Journal*, *33*(1), 47–51.

Beaulieu, H., Ferland, J. A., Gendron, B., & Michelon, P. (2000). A mathematical programming approach for scheduling physicians in the emergency room. *Health care management science*, *3*(3), 193-200.

Bélanger, J., Venne, P., & Paquin, J.-N. (2010). The What, Where and Why of Real-Time Simulation. *Planet RT*, *1*(0), 37–49.

Berggren, P., Prytz, E., Johansson, B., & Nählinder, S. (2011). The relationship between workload, teamwork, situation awareness, and performance in teams: A microworld study. *Proceedings of the Human Factors and Ergonomics Society*, 851–855.

Berwick, D. (2013). A promise to learn, a commitment to act: the Berwick Review into Patient Safety. *National Advisory Group on the Safety of Patients in England, Department of Health, London*.

Bevan, G. and Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Admin*istration, 84(3): 517–538, 2006.

Bhaskar, R. (2014). *The possibility of naturalism: A philosophical critique of the contemporary human sciences*. Routledge.

Bhaskar, R. (2013). *A realist theory of science*. Routledge.

Bhattacharjee, P., & Ray, P. K. (2014). Patient flow modelling and performance analysis of healthcare delivery processes in hospitals: A review and reflections. *Computers and Industrial Engineering*, *78*, 299–312.

Bittner, K. (2002). *Use case modeling*. Addison-Wesley Longman Publishing Co., Inc..

Blackwell, D. L., Martinez, M. E., Gentleman, J. F., Sanmartin, C., & Berthelot, J. M. (2009). Socioeconomic status and utilization of health care services in Canada and the United States: findings from a binational health survey. *Medical care*, 1136-1146.

Blessing, L. T. M. (1994). *A process-based approach to computer-supported engineering design* (pp. 1-369). University of Twente, Enschede.

Blessing, L. T., & Chakrabarti, A. (2009). *DRM: A design research methodology* (pp. 13-42). Springer London.

Boisot, M., & Canals, A. (2004). Data, information and knowledge: Have we got it right? *Journal of Evolutionary Economics*, *14*(1), 43–67.

Boivin, A., Lehoux, P., Lacombe, R., Burgers, J., & Grol, R. (2014). Involving patients in setting priorities for healthcare improvement: a cluster randomized trial. *Implementation Science*, *9*(1), 24.

Bonci, A., & Ingegneria, D. (2016). ScienceDirect A database-centric approach approach for the and control of approach for the and control of physical systems in modeling , of the future . A database-centric approach for the simulation and control of physical systems in the factory of the p. *IFAC-PapersOnLine*, *49*(12), 249–254.

Bond, K., Ospina, M. B., Blitz, S., Afilalo, M., Campbell, S. G., Bullard, M., ... & Rowe, B. H. (2007). Frequency, determinants and impact of overcrowding in emergency departments in Canada: a national survey. *Healthcare quarterly (Toronto, Ont.)*, *10*(4), 32-40.

Booker, M. J., Simmonds, R. L., & Purdy, S. (2014). Patients who call emergency ambulances for primary care problems: a qualitative study of the decision-making process. *Emergency Medicine Journal*, *31*(6), 448-452.

Boriboonsomsin, K., M. J. Barth, W. Zhu, and A. Vu. 2012. "Eco-routing Navigation System Based on Multisource Historical and Real-time Traffic Information". *IEEE Transactions on Intelligent Transportation Systems.* 13(4): 1694-1704.

Bostrom, R. P., & Heinen, J. S. (1977). MIS problems and failures: A socio-technical perspective. Part I: The causes. *MIS quarterly*, 17-32.

Bourdieu, P. (1990). *In other words: Essays towards a reflexive sociology*. Stanford University Press.

Bowen, A. B. (2019). How do emergency nurse practitioners experience managing acutely unwell patients in minor injury units? An Interpretative Phenomenological Analysis. *International emergency nursing*, *43*, 99-105.

Box, G. E. P. and G. M. Jenkins. (1976) Time Series Analysis: Forecasting and Control. Revised Edition. Holden Day: California.

Boyle, J., Jessup, M., Crilly, J., Green, D., Lind, J., Wallis, M., … Fitzgerald, G. (2012). Predicting emergency department admissions. *Emergency Medicine Journal*, *29*(5), 358–365.

Boyle, J., Ireland, D., Webster, F., & O'Sullivan, K. (2016). Predicting demand for hospital capacity planning. *3rd IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2016*, 328–331.

Boyle, A., & Higginson, I. (2018). This emergency department crisis was predictable - and partly preventable. *BMJ (Online)*, *360*(January), 2018. https://doi.org/10.1136/bmj.k64.

Bradley, E. H., McGraw, S. A., Curry, L., Buckser, A., King, K. L., Kasl, S. V., & Andersen, R. (2002). Expanding the Andersen model: The role of psychosocial factors in long-term care use. *Health services research*, *37*(5), 1221-1242.

Brailsford, S. (2005). Overcoming the barriers to implementation of operations research simulation models in healthcare. *Clinical and investigative medicine*, *28*(6), 312.

Brailsford, S. C. (2008, December). System dynamics: What's in it for healthcare simulation modelers. In *2008 Winter simulation conference* (pp. 1478-1483). IEEE.

Brailsford, S. C., Harper, P. R., Patel, B., & Pitt, M. (2009). An analysis of the academic literature on simulation and modelling in health care. *Journal of Simulation*, *3*, 130–140.

Brailsford, S. C., Bolt, T. B., Bucci, G., Chaussalet, T. J., Connell, N. a. D., Harper, P. R., … Taylor, M. (2013). Overcoming the barriers: a qualitative study of simulation adoption in the NHS. *Journal of the Operational Research Society*, *64*(2), 157–168.

Brailsford, S. C., Carter, M. W., & Jacobson, S. H. (2017). Five decades of healthcare simulation. In *2017 Winter Simulation Conference (WSC)* (pp. 365-384). IEEE.

Brailsford, S. C., Eldabi, T., Kunc, M., Mustafee, N., & Osorio, A. F. (2019). Hybrid simulation modelling in operational research: A state-of-the-art review. *European Journal of Operational Research*, *278*(3), 721-737.

Breen, B. M., & McCann, M. (2013). Healthcare providers' attitudes and perceptions of "inappropriate attendance" in the Emergency Department. *International Emergency Nursing*, *21*(3), 180–185.

Brem, A., & Voigt, K. I. (2009). Integration of market pull and technology push in the corporate front end and innovation management—Insights from the German software industry. *Technovation*, *29*(5), 351-367.

Brim, C. (2008). A descriptive analysis of the non-urgent use of emergency departments. *Nurse Researcher*, *15*(3), 72–88.

British Medical Association (2018) Working in a system that is under pressure. BMA 20180174. Available at: https://www.bma.org.uk/-/media/files/pdfs/collective%20voice/influence/key%20negotiations/nhs%20pressures/working-system-under-pressure-bma-council-report-mar-2018.pdf?la=en Accessed: 20.09.19.

British Medical Association (2019) NHS Pressures – Winter 2018/2019: A hidden crisis. BMA 20190262. Available at: https://www.bma.org.uk/-

/media/files/pdfs/collective%20voice/policy%20research/nhs%20structure%20and%20delivery/winter-pressures-report-2018-19.pdf?la=en Accessed: 20.09.19.

Brock, M. E., Vert, A., Kligyte, V., Waples, E. P., Sevier, S. T., & Mumford, M. D. (2008). Mental models: An alternative evaluation of a sensemaking approach to ethics instruction. *Science and Engineering Ethics*, *14*(3), 449–472.

Brocklesby, J., & Cummings, S. (1996). Foucault plays Habermas: An alternative philosophical underpinning for critical systems thinking. *Journal of the Operational Research Society*, *47*(6), 741-754.

Brocklesby, J. (2016). The what, the why and the how of behavioural operational research - An invitation to potential sceptics. *European Journal of Operational Research*, *249*(3), 796–805.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*. Springer.

Brownlee, J. (2018). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.

Brunkhorst, H., Kreide, R., & Lafont, C. (Eds.). (2017). *The Habermas Handbook*. Columbia University Press.

Bryman, A., & Bell, E. (2007). Business research strategies. *Business research methods*, 226-238.

Bumblauskas, D., Nold, H., Bumblauskas, P., & Igou, A. (2017). Big data analytics: transforming data to action. *Business Process Management Journal*. 23(3): 703-720.

Burger, K., White, L. & Yearworth, M. (2019). Developing a smart operational research with hybrid practice theories. *European Journal of Operational Research.* 277: 1137-1150.

Calegari, R., F. Fogliatto, F. Lucini, J. Neyeloff, R. Kuchenbecker, and B. Shaan. (2016). Forecasting Daily Volume and Acuity of Patients in the Emergency Department. *Computational and Mathematical Methods in Medicine*, ID 3863268. https://doi.org/10.1155/2016/3863268.

Callegaro, M., Murakami, M. H., Tepman, Z., & Henderson, V. (2015). Yes–no answers versus check-all in self-administered modes A systematic review and analyses. *Interntional Journal of Marketing Research*, *57*(2), 203–223.

Callen, J.L., Braithwaite, J., Westbrook, J.I. (2008). Contextual Implementation Model: A Framework for Assisting Clinical Information System Implementations, *Journal of the American Medical Informatics Association*, 15(2): 255–262

Carayon, P., Bass E. J., Bellandi, T., Gurses, A. P., Hallbeck, M.S. & Mollo, V. (2011). Sociotechnical systems analysis in health care: a research agenda, IIE *Transactions on Healthcare Systems Engineering*, 1:3, 145-160.

Carayon, P., Hancock, P., Leveson, N., Noy, I., Sznelwar, L., & Van Hootegem, G. (2015). Advancing a sociotechnical systems approach to workplace safety–developing the conceptual framework. *Ergonomics*, *58*(4), 548-564.

Carayon, P. (Ed.). (2016). *Handbook of human factors and ergonomics in health care and patient safety*. CRC Press.

Care Quality Commission (2018). Under pressure: Safely managing increased demand in emergency departments. May 2018. Available at: https://www.cqc.org.uk/publications/themed-work/under-pressure-safely-managing-increased-demand-emergency-departments

Carlsson, S. A. (2006). Towards an information systems design research framework: A critical realist perspective. In *Proceedings of the First International Conference on Design Science Research in Information Systems and Technology, Claremont, CA* (pp. 192-212).

Carvalho-Silva, M., Monteiro, M. T. T., de Sá-Soares, F., & Dória-Nóbrega, S. (2018). Assessment of forecasting models for patients arrival at emergency department. *Operations Research for Health Care*, *18*, 112-118.

Casimiro, L. M., Hall, P., Kuziemsky, C., O'Connor, M., & Varpio, L. (2015). Enhancing patient-engaged teamwork in healthcare: An observational case study. *Journal of interprofessional care*, *29*(1), 55-61.

Catchpole, K., & Alfred, M. (2018). Industrial conceptualization of health care versus the naturalistic decision-making paradigm: Work as imagined versus work as done. *Journal of Cognitive Engineering and Decision Making*, *12*(3), 222-226.

Ceglowski, R., Churilov, L., Wasserthiel, J. (2007). Combining Data Mining and Discrete Event Simulation for a value-added view of a hospital emergency department. *Journal of the Operational Research Society*, *58*(2), 246–254.

Chae, B., Yang, C., Olson, D., & Sheu, C. (2014). The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective. *Decision Support Systems*, *59*(1), 119–126.

Chahal, K., Eldabi, T., & Mandal, A. (2009). Understanding the impact of whiteboard on A&E department operations using hybrid simulation. In *Proceedings of the 27th International Conference of the System Dynamics Society*.

Chahal, K., and Eldabi, T. (2008). “Applicability of Hybrid Simulation to Different Modes of Governance in UK Healthcare.” In Proceedings of the 2008 Winter Simulation Conference, Edited by S. J. Mason, R. Hill, L. Monch, O. Rose, T. Jefferson, J. W. Fowler, (pp. 1478–1483). IEEE.

Chahal, K., Eldabi, T., & Young, T. (2013). A conceptual framework for hybrid system dynamics and discrete event simulation for healthcare. *Journal of Enterprise Information Management*, *26*(1-2), 50-74.

Chapman, B., & Turnbull, T. (2016). Health professionals’ perceptions of inappropriate use of A&E services. *British Journal of Nursing (Mark Allen Publishing)*, *25*(9), 476–483.

Chatfield, C., & Xing, H. (2019). *The analysis of time series: an introduction with R*. CRC press.

Cheek, C., Allen, P., Shires, L., Parry, D., & Ruigrok, M. (2016). Low-acuity presentations to regional emergency departments: What is the issue? *EMA - Emergency Medicine Australasia*, *28*(2), 145–152.

Chen, H., & Storey, V. C. (2012). Business intelligence and analytics : from Big Data to big impact. *36*(4), 1165–1188.

Chen, P., Lin, C., & Wu, W. (2020). International Journal of Information Management Big data management in healthcare : Adoption challenges and implications. *International Journal of Information Management*. doi.org/10.1016/j.ijinfomgt.2020.102078

Chiappe, D. L., Strybel, T. Z., & Vu, K. P. L. (2012). Mechanisms for the acquisition of situation awareness in situated agents. *Theoretical Issues in Ergonomics Science*, *13*(6), 625–647.

Chiappe, D., Strybel, T. Z., & Vu, K. P. L. (2015). A situated approach to the understanding of dynamic situations. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 33–43.

Choi, B. C., & Pak, A. W. (2005). Peer reviewed: a catalog of biases in questionnaires. *Preventing chronic disease*, *2*(1).

Choudhury, A. (2019). Hourly Forecasting of Emergency Department Arrivals: Time Series Analysis. *arXiv preprint arXiv:1901.02714*.

Chuang, S., Ou, J. C., & Ma, H. P. (2020). Measurement of resilience potentials in emergency departments: Applications of a tailored resilience assessment grid. *Safety Science*, *121*, 385-393.

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

Clegg, C. W. (2000). Sociotechnical principles for system design. *Applied ergonomics*, *31*(5), 463-477.

Codrington-Virtue, A., Chaussalet, T., Millard, P., Whittlestone, P., & Kelly, J. (2006). A system for patient management based discrete-event simulation and hierarchical clustering. *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems*, 1–5.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

Coleman, P., Irons, R., & Nicholl, J. (2001). Will alternative immediate care services reduce demands for non-urgent treatment at accident and emergency? *Emergency Medicine Journal*, *18*(6), 482–487.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the learning sciences*, *13*(1), 15-42.

Connell, N. A. D. (2001). Evaluating soft OR: some reflections on an apparently “unsuccessful” implementation using a Soft Sytems Methodology (SSM)-based approach. *Journal of the Operational Research Society*, *52*, 150–160.

Connell, N. A. D., Klein, J. H., & Powell, P. L. (2003). Its tacit knowledge but not as we know it: redirecting the search for knowledge. *Journal of the Operational Research Society*, *54*(2), 140-152.

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications.

Cornwall, A., & Jewkes, R. (1995). What is participatory research? *Social Science and Medicine*, *41*(12), 1667–1676.

Coulter, A., Locock, L., Ziebland, S., & Calabrese, J. (2014). Collecting data on patient experience is not enough: they must be used to improve care. *BMJ (Clinical research ed.)*, *348*, g2225. 1–4.

Coulter, A. (2015). Leadership for Patient Engagement. London: The King’s Fund; 2012.

Crema, M., & Verbano, C. (2019). Simulation modelling and lean management in healthcare: first evidences and research agenda. *Total Quality Management & Business Excellence*, 1-19.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Creswell, J.W and Plano Clark V.L. (2018) Designing and conducting mixed methods research (3rd ed.), Sage, Thousand Oaks, CA

Crowe, S., Brown, K., Tregay, J., Wray, J., Knowles, R., Ridout, D. A., ... & Utley, M. (2017). Combining qualitative and quantitative operational research methods to inform quality improvement in pathways that span multiple settings. *BMJ quality & safety*, *26*(8), 641-652.

Daellenbach H. G., McNickle, D.C. and Dye, S. (2012). “Management Science: Decision-Making Through Systems Thinking”. New York: Palgrave Macmillan

Dalal, M., Groel, B. & Prieditis, A. (2003). Real-Time Decision Making using simulation. *Proceedings of the 2003 Winter Simulation Conference. S. Chick, P. J. Sanchez, D. Ferrin, and D. J. Momice, eds. IEEE.*

Davenport, T. H., and Harris, J. G.(2007) Competing on analytics: the new science of winning. Boston: Harvard Business School Press.

Davenport, T. H., Harris, J. G., DeLong, D. W. & Jacobson, A. L. (2001). Data to Knowledge to Results: *California Management Review*, *43*(2), 117–138.

Davies H (2015) 'Managing doctors, doctors managing: Troubled relations at the heart of the NHS'. Nuffield Trust comment, 6 February 2015. https://www.nuffieldtrust.org.uk/news-item/managing-doctors-doctors-managing-troubled-relations-at-the-heart-of-the-nhs accessed: 02.02.20.

Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International journal of man-machine studies*, *38*(3), 475-487.

Degeling P, Zhang K, Coyle B, Xu L, Meng Q, Qu J and Hill M (2006). 'Clinicians and the governance of hospitals: A cross-cultural perspective on relations between professions and management'. Social Science & Medicine, vol 63 (3), pp 757–775.

De Gooyert, V., Rouwette, E., Van Kranenburg, H., & Freeman, E. (2017). Reviewing the role of stakeholders in operational research: A stakeholder theory perspective. *European Journal of Operational Research*, *262*(2), 402-410.

Dekker, S., Cilliers, P., & Hofmeyr, J. H. (2011). The complexity of failure: implications of complexity theory for safety investigations. *Safety science*, *49*(6), 939-945.

Delen, D., Erraguntla, M., Mayer, R. J., & Wu, C. N. (2011). Better management of blood supply-chain with GIS-based analytics. *Annals of Operations Research*, *185*(1), 181–193.

Delen, D. Demirkan, H. (2013). Data, information and analytics as services, *Decision Support Systems* 55 (1): 359–363.

Delen, D., & Zolbanin, H. M. (2018). The analytics paradigm in business research. *Journal of Business Research*, *90*(May), 186–195.

Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013, May). Addressing big data issues in scientific data infrastructure. In *2013 International Conference on Collaboration Technologies and Systems (CTS)* (pp. 48-55). IEEE.

Department of Health. (2007) NHS Information Governance: Guidance on Legal and Professional Obligations. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/200702/NHS_Information_Governance_Guidance_on_Legal_and_Professional_Obligations.pdf Last accessed: 17.01.2020

Department of Health (2018) The Government's revised mandate to NHS England for 2017-18. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/692140/NHSE_Mandate__2017-18_revised.pdf last accessed: 09.03.2020

de la Torre Diez, I., Cosgaya, H. M., Garcia-Zapirain, B., & Lopez-Coronado, M. (2016). Big Data in Health: a Literature Review from the Year 2005. *Journal of Medical Systems*, *40*(9), 1–6.

De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation research part F: traffic psychology and behaviour*, *27*, 196-217.stat de Weck, O. L., Roos, D., and Magee, C. L. (2011). *Engineering Systems: Meeting Human Need in a Complex Technological World*. Cambridge, MA: MIT Press.

Dewey, J. (1998). *The essential Dewey: Pragmatism, education, democracy* (Vol. 1). Indiana University Press.

Dinh, M. M., Russell, S. B., Bein, K. J., Chalkley, D. R., Muscatello, D., Paoloni, R., & Ivers, R. (2016). Statewide retrospective study of low acuity emergency presentations in New South Wales, Australia: who, what, where and why? *BMJ open*, *6*(5), e010964.

Dixe, M., Passadouro, R., Peralta, T., Ferreira, C., Lourenço, G., & Sousa, P. (2018). Determinants of non-urgent emergency department use. *Revista de Enfermagem Referência*, *IV Série*(Nº16), 41–52.

Dixon, D. R. (1999). The behavioral side of information technology. *International journal of medical informatics*, *56*(1-3), 117-123.

Dode, P., Greig, M., Zolfaghari, S., & Neumann, W. P. (2016). Integrating human factors into discrete event simulation: a proactive approach to simultaneously design for system performance and employees' well being. *International Journal of Production Research*, *54*(10), 3105-3117.

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of business and Psychology*, *17*(2), 245-260.

Drake, G. R., & Smith, J. S. (1996). Simulation system for real-time planning, scheduling, and control. In *Proceedings of the 1996 Winter Simulation Conference ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain. (pp. 1083 – 1090).* IEEE.

Dreyfus, S. E. (1981) Formal models vs human situational understanding: Inherent limitations on the modelling of business expertise (ORC 81-2). Berkeley, CA: University of California, Operations Research Centre.

Drory, A., & Vigoda-Gadot, E. (2010). Organizational politics and human resource management: A typology and the Israeli experience. *Human Resource Management Review*, *20*(3), 194-202.

Dunnett, S., Leigh, J., & Jackson, L. (2019). Optimising police dispatch for incident response in real time. *Journal of the Operational Research Society*, *70*(2), 269-279.

Durand, A. C., Gentile, S., Devictor, B., Palazzolo, S., Vignally, P., Gerbeaux, P., & Sambuc, R. (2011). ED patients: How nonurgent are they? Systematic review of the emergency medicine literature. *American Journal of Emergency Medicine*, *29*(3), 333–345.

Durso, F.T. , Dattel, A.R. , Banbury, S. , Tremblay, S. (2004). 'SPAM: the real-time assessment of SA'. *A Cognitive Approach to Situation Awareness: Theory and Application*. Hampshire:Ashgate Publishing Ltd. , 137-15

Easterby-Smith, M., Thorpe, R., & Jackson, P. R. (2015). *Management and business research*. Sage.

Effken, J. A. 2002. Different lenses, improved outcomes: A new approach to the analysis and design of healthcare information systems. *International Journal of Medical Informatics*, 65(1): 59–74.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of management journal*, *50*(1), 25-32.

Ekstrom, A., L. Kurland, N. Farrokhnia, M. Castren, M. Nordberg (2015) Forecasting Emergency Department Visits Using Internet Data. *Annals of Emergency Medicine* 65(4): 436-442.

Elbattah, M., & Molloy, O. (2016). *Coupling Simulation with Machine Learning : A Hybrid Approach for Elderly Discharge Planning*. 47–56.

Elbattah, M. & Molloy, O. (2017). *Clustering-Aided Approach for Predicting Patient Outcomes with Application to Elderly Healthcare in Ireland. Conference: Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, California, USA*

Eldabi, T., Balaban, M., Brailsford, S., Mustafee, N., Nance, R. E., Onggo, B. S., & Sargent, R. G. (2016). Hybrid simulation: Historical lessons, present challenges and futures. In *2016 Winter Simulation Conference (WSC)* (pp. 1388-1403). IEEE.

Eldabi, T., Paul, R. J., & Young, T. (2007). Simulation modelling in healthcare: reviewing legacies and investigating futures. *Journal of the Operational Research Society*, *58*(2), 262–270.
El Morr, C., & Ali-Hassan, H. (2019). Analytics in Healthcare A Practical Introduction. Springer: Switzerland.

Elsbach, K.D., van Knippenberg, D.: Creating high-impact literature reviews: an argument for 'integrative reviews'. J. Manag. Stud. 1–13 (2020). https://doi.org/10.1111/joms.12581

Emery, F. E., & Trist, E. L. (1960). Socio-technical systems. Management sciences, models and Techniques. Churchman CW.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32-64.

Endsley, M. R., & Garland, D. J. (2000). Theoretical underpinnings of situation awareness: A critical review. *Situation awareness analysis and measurement*, *1*(1), 3-21.

Endsley, M. R. (2016). *Designing for situation awareness: An approach to user-centered design*. CRC press.

England, W., & Roberts, S. D. (1978). *Applications of Computer Simulation to Health Care*. Institute of Electrical and Electronics Engineers (IEEE).

Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: toward a science of the structure and acquisition of expert and elite performance. *International journal of sport psychology*. 38(1), 4-34

Espinoza, C., Pascual, J., Ramis, F., & Broquez, D. (2014). Real-time simulation as a way to improve daily operations in an emergency room. In *Proceedings 2014 Winter Simulation Conference*, (pp.2600–2608). IEEE.

Evans, R. S., Lloyd, J. F., & Pierce, L. A. (2012). Clinical use of an enterprise data warehouse. In *AMIA Annual Symposium Proceedings* (Vol. 2012, p. 189). American Medical Informatics Association.

Evelson, B. (2010). The Forrester Wave™: Enterprise Business Intelligence Platforms, Q4 2010. *Intelligence (BI)*, *3*.

Faingloz, L., & Tolujew, J. (2017). Simulation Modelling Application in Real-time Service Systems: Review of the Literature. *Procedia Engineering*, *178*, 200-205.

Fakhimi, M. & Probert, J. (2013) Operations research within UK healthcare: a review. *Journal of Enterprise Information Management* 26(1/2): 21-49.

Fakhimi, M., Stergioulas, L. K. & Mustafee, N. (2015) "An investigation of Hybrid Simulation for modeling sustainability in healthcare," In *Proceedings 2015 Winter Simulation Conference (WSC)*, (pp. 1585-1596). IEEE.

Falzer, P. R. (2018). Naturalistic decision making and the practice of health care. *Journal of Cognitive Engineering and Decision Making*, *12*(3), 178-193.

Fanchao, Z., Turner, S. J., & Aydt, H. (2009). Symbiotic simulation control in supply chain of lubricant additive industry. *Proceedings - IEEE International Symposium on Distributed Simulation and Real-Time Applications, DS-RT*, 165–172.

Feldkamp, N., Bergmann, S., & Strassburger, S. (2015). Visual analytics of manufacturing simulation. In L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, & M. D. Rossetti (Eds.), In *Proceedings of the 2015 Winter Simulation Conference* (pp. 779–790), IEEE.

Feldkamp, N., Bergmann, S., & Strassburger, S. (2017). Online analysis of simulation data with stream-based data mining. Proceedings of the 2017 ACM SIGSIM- PADS Conference (pp. 241–248). New York, NY: ACM.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. (2008). Forecasting and operational research: a review. *Journal of the Operational Research Society*, *59*(1), 1150–1172.

Fioratou, E., Flin, R., Glavin, R., & Patey, R. (2010). Beyond monitoring: distributed situation awareness in anaesthesia. *British journal of anaesthesia*, *105*(1), 83-90.

Fishwick, P.A. (1995) *Simulation Model Design and Execution*. Prentice Hall, London.

Fitzgerald L, Lilley C, Ferlie E, Addicott R, McGivern G and Buchanan D (2006). Managing Change and Role Enactment in the Professionalised Organisation. London: National Coordinating Centre for NHS Service Delivery and Organisation R&D.

FitzGerald, G., Toloo, G., Aitken, P., Keijzers, G., & Scuffham, P. (2015). Public use and perceptions of emergency departments: a population survey. *Emergency Medicine Australasia*, *27*(4), 336-342.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry*, *12*(2), 219-245.

Fokkinga, B., Bleijenbergh, I., & Vennix, J. (2009). Group model building evaluation in single cases: a method to assess changes in mental models. In *27th International Conference of the System Dynamics Society, New Mexico, USA*.

Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., ... & Palmer, S. (2003). Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health*, *25*(4), 325-335.

Ford, G., Compton, M., Millett, G., & Tzortzis, A. (2017). The role of digital disruption in healthcare service innovation. In *Service business model innovation in healthcare and hospital management* (pp. 57-70). Springer, Cham.

Franco, L. A., & Montibeller, G. (2010). Facilitated modelling in operational research. *European Journal of Operational Research*, *205*(3), 489–500.

Franco, L. A., & Lord, E. (2011). Understanding multi-methodology: Evaluating the perceived impact of mixing methods for group budgetary decisions. *Omega*, *39*(3), 362–372.

Franco, L. A., & Hämäläinen, R. P. (2016). Behavioural operational research: Returning to the roots of the OR profession. *Eur. J. Oper. Res.*, *249*(3), 791-795.

Franco, L. A., & Hämäläinen, R. P. (2016). Engaging with behavioural operational research: On methods, actors and praxis. In *Behavioral Operational Research* (pp. 3-25). Palgrave Macmillan, London.

Fujimoto, R. M., Lunceford, D., Page, E. H., & Davis, P. K. (2002). Grand Challenges for Modeling and Simulation. *Dagstuhl Report*, 75.

Galetsi, P., Katsaliaki, K., & Kumar, S. (2019a). Social Science & Medicine Values, challenges and future directions of big data analytics in healthcare : A systematic review. *Social Science & Medicine*, *241*(September), 112533.

Galetsi, P., & Katsaliaki, K. (2019b). Review Article Big data analytics in health : an overview and bibliometric study of research activity. *Health Information Libraries Journal* 37: 5–25.

Galetsi, P., & Katsaliaki, K. (2019c). A review of the literature on big data analytics in healthcare. *Journal of the Operational Research Society*, *0*(0), 1–19.

Garg, L., McClean, S., Meenan, B., El-Darzi, E., & Millard, P. (2009). Clustering patient length of stay using mixtures of Gaussian models and phase type distributions. *Proceedings - IEEE Symposium on Computer-Based Medical Systems*.

Gaynor, M., Seltzer, M., & Moulton, S. (2005). A dynamic, data-driven, decision support system for emergency medical services. *Computational Science–ICCS 2005*, 61–100.

Gerring, J. (2006). *Case study research: Principles and practices*. Cambridge university press.

Ghiara, V. (2020). Disambiguating the role of paradigms in mixed methods research. *Journal of Mixed Methods Research*, *14*(1), 11-25.

Giddens, Anthony. 1984. *The Constitution of Society: Outline of a Theory of Structuration*. Berkeley, CA: University of California Press.

Gillespie, B. M., Gwinner, K., Fairweather, N., & Chaboyer, W. (2013). Building shared situational awareness in surgery through distributed dialog. *Journal of multidisciplinary healthcare*, *6*, 109.

Given, L. M. (2017). It's a New Year… So Let's Stop the Paradigm Wars. *International Journal of Qualitative Methods*, *16*, 1-2.

Glowacka, K. J., Henry, R. M., & May, J. H. (2009). A hybrid data mining/simulation approach for modelling outpatient no-shows in clinic scheduling. *Journal of the Operational Research Society*, *60*(8), 1056–1068.

Gogi, A., Tako, A. A., & Robinson, S. (2016). An experimental investigation into the role of simulation models in generating insights. *European Journal of Operational Research*, *249*(3), 931–944.

Gore, J., Banks, A., Millward, L., Kyriakidou, O., & Banks, A. (2006). *Naturalistic Decision Making and Organizations : Reviewing Pragmatic Science*. 925–942.

Gorski, P. S. (2013). What is critical realism? And why should you care? *Contemporary Sociology.* 42(5): 658-670

Graban, M. (2016). *Lean hospitals: improving quality, patient safety, and employee engagement*. CRC press.

Graham, B., R. Bond, M. Quinn, and M. Mulvenna. (2018) Using Data Mining to Predict Hospital Admissions from the Emergency Department. *IEEE Access* 6: 10458-10469.

Greasley, A and Owen, C. (2018). Modelling People's Behaviour using Discrete-Event Simulation: A Review. *International Journal of Operations and Production Management*, 38 (5), pp. 1228-1244.

Greasley, A., & Edwards, J. S. (2019). Enhancing discrete-event simulation with big data analytics : A review Enhancing discrete-event simulation with big data analytics : A review. *Journal of the Operational Research Society*, *0*(0), 1–21.

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. Association for Information Systems. Available at: https://openresearchrepository.anu.edu.au/bitstream/1885/32762/8/2007%20JAIS-ISDT-spec-220307%20Gregor_Jones_near_%20final.pdf Accessed: 21.10.19

Gregor, S. & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*. 37(2): 337-355

Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. (2005). Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med* 61(2):417-430.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., Hinder, S., ... & Shaw, S. (2017). Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of medical Internet research*, *19*(11), e367.

Greenhalgh, T., & Abimbola, S. (2019). The NASSS framework—a synthesis of multiple theories of technology implementation. *Stud. Health Technol. Inform*, *263*, 193-204.

Gruber, J., Hoe, T. & Stoye, G. (2018). Wait time targets and patient outcomes in emergency departments. VoxEU. Available at: https://voxeu.org/article/impact-waiting-time-targets-emergency-departments Accessed: 17.12.19

Gualtar, C. (2018). *Setting an industry 4.0 research and development agenda for simulation – a literature review*. *Int J Simul Model.* 17(3): 377–390.

Guise, J.-M., Hansen, M., Lambert, W., & O'Brien, K. (2017). The role of simulation in mixed-methods research: a framework & application to patient safety. *BMC Health Services Research*, *17*(1), 322.

Gul, M., & Guneri, A. F. (2015). A comprehensive review of emergency department simulation applications for normal and disaster conditions. *Computers and Industrial Engineering*, *83*, 327–344.

Gunal, M. M. (2012). A guide for building hospital simulation models. *Health Systems*, *1*(1), 17–25.

Gunal, M., & Pidd, M. (2009). Discrete Event Simulation for Performance Modelling in Healthcare: A Review of the Literature. *Management Science*, *4*(1), 1–22.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). *Journal of Strategic Information Systems Debating big data : A literature review on realizing value from big data*. *26*, 191–209.

Haas, P. J., Maglio, P. P., Selinger, P. G., & Tan, W. (2011). Data is Dead … Without What-If Models. *In Proceedings of the Very Large Database (VLDB) Endowment, 4 (12) (2011), Pp. 1486-1489*, *4*(12), 11–14.

Ham, C., Berwick, D., & Dixon, J. (2016). *Improving quality in the English NHS. A strategy for action.* (February), 38.

Ham, C. (2017). Political crisis in the NHS. *BMJ* 2017;356:j218

Hammad, A., Faridaddin, V., Cheng, Z., Mohammed, M., & Ahmad, D. (2012). Towards the Smart Construction Site: Improving Productivity and Safety of Construction Projects Using Multi-Agent Systems, Real-Time Simulation and Automated Machine Control. *Proceedings of the 2012 Winter Simulation Conference*, 0–1.

Hancké, B. (2009). *Intelligent research design: a guide for beginning researchers in the social sciences*. Oxford University Press.

Handley MA, Gorukanti A, Cattamanchi A (2016) Strategies for implementing implementation science: a methodological overview. *Emergency Medicine Journal.* 33, 660-664.

Hanisch, A., Tolujew, J., & Schulze, T. (2005). Initialization of online simulation models. In *Proceedings of the 2005 Winter Simulation Conference M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, Eds.*, (pp.1795–1803). IEEE.

Hardyman, W., Daunt, K. L., Kitchener, M., & Hardyman, W. (2015). Value Co-Creation through Patient Engagement in Health Care : A micro-level approach and research agenda. *Public Management Review*, *17*(1) 90–107.

Harenčárová, H. (2017). Managing uncertainty in paramedics' decision making. *Journal of Cognitive Engineering and Decision Making*, *11*(1), 42-62.

Harper, A., Mustafee, N., & Feeney, M. (2017). A hybrid approach using forecasting and discrete-event simulation for endoscopy services. In *2017 Winter Simulation Conference (WSC)* (pp. 1583-1594). IEEE.

Harper, A. and Mustafee, N. (2019) A Hybrid Modelling Approach using Forecasting and Real-Time Simulation to Prevent Emergency Department Overcrowding. In: *Proceedings of 2019 Winter Simulation Conference (WSC).* IEEE.

Harper, A., Mustafee, N. and Yearworth, M. (2020) Facets of Trust in Simulation Studies. *European Journal of Operational Research* https://doi.org/10.1016/j.ejor.2020.06.043

Harper, P. R., & Pitt, M. (2004). On the challenges of healthcare modelling and a proposed project life cycle for successful implementation. *Journal of the Operational Research Society*, *55*(6), 657–661.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). North-Holland.

Hau, S., Ioannidis, A., Masaoutis, P., & Verma, S. (2008). Patterns of ophthalmological complaints presenting to a dedicated ophthalmic Accident & Emergency department: Inappropriate use and patients' perspective. *Emergency Medicine Journal*, *25*(11), 740–744.

Helal, M., Rabelo, L., Sepúlveda, J., & Jones, A. (2007). A methodology for integrating and synchronizing the system dynamics and discrete event simulation paradigms. In *Proceedings of the 25th international conference of the system dynamics society* (Vol. 3, No. 3, pp. 1-24).

Henao, F., & Franco, L. A. (2016). Unpacking multimethodology: Impacts of a community development intervention. *European Journal of Operational Research*, *253*(3), 681-696.

Hersh, W. R. (2014). Healthcare Data Analytics. *Health Informatics: Practical Guide for Healthcare and Information Technology Professionals, Chapter 3* (Sixth Edition), 1–14.

Hevner, A.R., March, S.T., Park, J., Ram, S.(2004) Design Science in Information Systems Research. *MIS Quarterly* 28, 75–106

Higginson, I., & Boyle, A. (2018). What should we do about crowding in emergency departments? . *British Journal of Hospital Medicine, 79*(9), 500–503.

Hoad, K., Monks, T., & O'Brien, F. (2015). The use of search experimentation in discrete-event simulation practice. *Journal of the Operational Research Society*, *66*(7), 1155–1168.

Hodgkinson, G. P., & Starkey, K. (2011). Not simply returning to the same answer over and over again: Reframing relevance. *British Journal of Management*, *22*(3), 355–369.

Hodgkinson, G. P., & Starkey, K. (2012). Extending the Foundations and Reach of Design Science: Further Reflections on the Role of Critical Realism. *British Journal of Management*, *23*(4), 605–610.

Hollnagel, E. (2009). The four cornerstones of resilience engineering. Christopher P. Nemeth, Erik

Hollnagel, E & Dekker, S (2009). *Resilience Engineering Perspectives, Volume 2: Preparation and Restoration*, Ashgate, Chapter 6 - p. 117-134, 2009, Ashgate Studies in Resilience Engineering.

Hollnagel, E. (2010). How resilient is your organisation? An introduction to the resilience analysis grid (RAG). *Sustainable Transformation: Building a Resilient Organization*, May 2010, Toronto, Canada. ⟨hal-00613986⟩

Hollnagel, E. (2011a). Coping with complexity: past, present and future. *Cognition, Technology & Work*, *14*(3), 199-205.

Hollnagel, E. (2011b). RAG-The resilience analysis grid. *Resilience engineering in practice. A guidebook. Farnham, UK: Ashgate*.

Hollnagel, E., Braithwaite, J. & Wears, R.L. (2019). *Resilient health care*. CRC Press.

Holm, L. B., Dahl, F. A., & Barra, M. (2012). Towards a multimethodology in health care – synergies between Soft Systems Methodology and Discrete Event Simulation. *Health Systems*, *2*(1), 11–23.

Holman, M., Walker, G., Lansdown, T., & Hulme, A. (2019). Radical systems thinking and the future role of computational modelling in ergonomics: an exploration of agent-based modelling. *Ergonomics*, 1-18.

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems*, *64*, 130–141.

Hoot, N. R., LeBlanc, L. J., Jones, I., Levin, S. R., Zhou, C., Gadd, C. S., & Aronsky, D. (2008). Forecasting emergency department crowding: a discrete event simulation. *Annals of emergency medicine*, *52*(2), 116-125.

Hoot, N. R., & Aronsky, D. (2008). Systematic review of emergency department crowding: causes, effects, and solutions. *Annals of emergency medicine*, *52*(2), 126-136.

Hoot NR, Leblanc LJ, Jones I, Levin SR, Zhou C, Gadd CS, Aronsky D. (2009) Forecasting emergency department crowding: a prospective, real-time evaluation. *J Am Med Inform Assoc* 16: 338– 45.

House of Commons Health Committee (2017). Winter pressure in accident and emergency departments. Available at: https://publications.parliament.uk/pa/cm201617/cmselect/cmhealth/277/277.pdf Accessed: 01.09.19.

House of Commons Library Briefing Paper (2017) Accident and Emergency Statistics: Demand, Performance and Pressure. Available at: https://commonslibrary.parliament.uk/research-briefings/cbp-7281/ Accessed: 01.09.19.

House of Commons Library (2019) NHS pressures in England: Waiting times, demand, and capacity. Available at: https://commonslibrary.parliament.uk/insights/nhs-pressures-in-england-waiting-times-demand-and-capacity/ Accessed: 02.03.20.

House of Commons Library Briefing Paper (2020) NHS Key Statistics: England, February 2020. Available at: https://commonslibrary.parliament.uk/research-briefings/cbp-8846/ Accessed: 01.05.20.

Horbach, J., Rammer, C., & Rennings, K. (2012). Determinants of eco-innovations by type of environmental impact—The role of regulatory push/pull, technology push and market pull. *Ecological economics*, *78*, 112-122.

Howick, S. & Ackermann, F. (2011). Mixing OR methods in practice : Past, present and future directions. *European Journal of Operational Research,* 215(3): 503-511.

Hsia, R. Y., & Niedzwiecki, M. (2017). Avoidable emergency department visits: A starting point. *International Journal for Quality in Health Care*, *29*(5), 642–645.

HSJ (2020) Coronavirus insight: the cost of covid-19. Health Services Journal. Available at: https://www.hsj.co.uk/daily-insight/coronavirus-insight-the-cost-of-covid-19/7027066.article Accessed: 15.06.20

Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organization science*, *2*(1), 88-115.

Huerta, T. R., Ford, E. W., Peterson, L. T., & Brigham, K. H. (2008). Testing the hospital value proposition: An empirical analysis of efficiency and quality. *Health Care Management Review*, *33*(4), 341-349.

Hvitfeldt-Forsberg, H., Mazzocato, P., Glaser, D., Keller, C., & Unbeck, M. (2017). Staffs ' and managers ' perceptions of how and when discrete event simulation modelling can be used as a decision support in quality improvement : a focus group discussion study at two hospital settings in Sweden. *BMJ Open*.

Hyndman, R (2014) Thoughts on the Ljung-Box test. Available at: https://robjhyndman.com/hyndsight/ljung-box-test/ Accessed: 02.06.19.

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed: 06.02. 19.

Idil, H., Kilic, T. Y., Toker, İ., Dura Turan, K., & Yesilaras, M. (2018). Non-urgent adult patients in the emergency department: Causes and patient characteristics. *Turkish Journal of Emergency Medicine*, *18*(2), 71–74.

Irani, Z. (2002). Information systems evaluation: navigating through the problem domain. *Information & Management*, *40*(1), 11-24.

Isken, M. W., & Rajagopalan, B. (2002). Data mining to support simulation modeling of patient flow in hospitals. *Journal of Medical Systems*, *26*(2), 179–197.

Ivankova, N., & Wingo, N. (2018). Applying Mixed Methods in Action Research: Methodological Potentials and Advantages. *American Behavioral Scientist*, *62*(7), 978–997.

Jabbal. J. & Lewis, M. (2018). Approaches to better Value in the NHS: Improving quality and cost. The Kings Fund. Available at: https://www.kingsfund.org.uk/sites/default/files/2018-10/approaches-to-better-value-october2018_0.pdf Accessed: 06.07.19

Jackson, M. C. (1985). The itinerary of a critical approach. *Journal of the Operational Research Society*, *36*(9), 878-881.

Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L. K., & Young, T. (2010). Simulation in manufacturing and business: A review. *European Journal of Operational Research*, *203*(1), 1–13.

Jahangirian, M., Naseer, A., Stergioulas, L., Young, T., Eldabi, T., Brailsford, S., … Harper, P. (2012). Simulation in health-care: Lessons from other sectors. *Operational Research*, *12*(1), 45–55.

Jahangirian, M., Taylor, S. J., Eatock, J., Stergioulas, L. K., & Taylor, P. M. (2015). Causal study of low stakeholder engagement in healthcare simulation projects. *Journal of the Operational Research Society*, *66*(3), 369–379.

Jahangirian, M., Taylor, S. J., Young, T., & Robinson, S. (2016). Key performance indicators for successful simulation projects. . *Journal of the Operational Research Society*, 1–19

Jahangirian, M., Taylor, S. J. E., Young, T., & Robinson, S. (2017). Key performance indicators for successful simulation projects. *Journal of the Operational Research Society 1.59*, *68*(7), 747–765.

Janke, A. T., Overbeek, D. L., Kocher, K. E., & Levy, P. D. (2016). Exploring the Potential of Predictive Analytics and Big Data in Emergency Care. *Annals of Emergency Medicine*, *67*(2), 227–236.

Jayaraman, P. P., Forkan, A. R. M., Morshed, A., Haghighi, P. D., & Kang, Y. B. (2020). Healthcare 4.0: A review of frontiers in digital health. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(2), e1350.

Jeffcott, S. A., & Mackenzie, C. F. (2008). Measuring team performance in healthcare: review of research and implications for patient safety. *Journal of critical care*, *23*(2), 188-196.

Jimenez, J. I., Jahankhani, H., & Kendzierskyj, S. (2020). Health Care in the Cyberspace: Medical Cyber-Physical System and Digital Twin Challenges. In *Digital Twin Technologies and Smart Cities* (pp. 79-92). Springer, Cham.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, I. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research.* 1(2): 112-133.

Johnston, A., Abraham, L., Greenslade, J., Thom, O., Carlstrom, E., Wallis, M., & Crilly, J. (2016). Review article: Staff perception of the emergency department working environment: Integrative review of the literature. *EMA - Emergency Medicine Australasia*, *28*(1), 7–26.

Jolaoso, S., Burtner, R., & Endert, A. (2015). Toward a Deeper Understanding of Data Analysis, Sensemaking, and Signature Discovery. *International Federation for Information Processing*, *9296*, 463–478.

Jones, S. S., A. Thomas, S. Evans, S. Welch, P. Haug, & G. L. Snow. (2008). Forecasting Daily Patient Volumes in the Emergency Department. *Academic Emergency Medicine* 15: 159-170.

Jones, D. G. (2015). A practical perspective on the utility of situation awareness. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 98-100.

Jothi, N., Rashid, N. A., & Husain, W. (2015). Data Mining in Healthcare - A Review. *Procedia Computer Science*, 306–313

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

Jurishica, C. J. (2005). Emergency department simulations: Medicine for building effective models. In *Proceedings 2005 Winter Simulation Conference*, *(pp.* 2674–2680). IEEE.

Kadri, F., Chaabane, S., & Tahon, C. (2014). A simulation-based decision support system to prevent and predict strain situations in emergency department systems. *Simulation Modelling Practice and Theory*, *42*, 32–52.

Kadri, F., Chaabane, S., Bekrar, A., Tahon, C., (2015). Resilience-Based Performance Assessment of Strain Situations in Emergency Departments. Presented at the Industrial Engineering and Systems Management (IESM), 2015 International Conference on, Seville, pp. 609–618.

Kadri, F., Chaabane, S., & Tahon, C. (2016). Toward a decision support system for the management of strain situations in hospital emergency departments. In *Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference* (pp. 568-574).

Kahneman, D., & Frederick, S. (2005). A model of intuitive judgment. *The Cambridge handbook of thinking and reasoning*, 267-293.

Kang, H., Nembhard, H. B., Rafferty, C., & Deflitch, C. J. (2014). Patient flow in the emergency department: A classification and analysis of admission process policies. *Annals of Emergency Medicine*, *64*(4).

Kapoutsis, I. (2016). 2 Playing the political game at work: the roles of political will, political prudence and political skill. *Handbook of Organizational Politics: Looking Back and to the Future*, 40.

Karsh, B. T. (2004). Beyond usability: designing effective technology implementation systems to promote patient safety. *BMJ Quality & Safety*, *13*(5), 388-394.

Kao, H.-Y., Yu, M.-C., Masud, M., Wu, W.-H., Chen, L.-J., & Wu, Y.-C. J. (2016). Design and evaluation of hospital-based business intelligence system (HBIS): A foundation for design science research methodology. *Computers in Human Behavior*, *62*, 495–505.

Katsaliaki, K., & Mustafee, N. (2011). Applications of simulation within the healthcare context. *Journal of the Operational Research Society*, *62*, 1431–1451.

Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: Issues, challenges, tools and Good practices. *2013 6th International Conference on Contemporary Computing, IC3 2013*, 404–409.

Kaushik, S., Choudhury, A., Sheron, P. K., Dasgupta, N., Natarajan, S., Pickett, L. A., & Dutt, V. (2020). AI in healthcare: time-series forecasting using statistical, neural, and ensemble architectures. *Frontiers in Big Data*, *3*, 4.

Kelly, C. J., & Young, A. J. (2017). Promoting innovation in healthcare. *Future healthcare journal*, *4*(2), 121.

Kelton, W. D., R. P. Sadowski, and N. B. Zupick. (2015). Simulation with Arena. New York: McGraw-Hill.

Keogh, B. (2013). Review into the quality of care and treatment provided by 14 hospital trusts in England: overview report. NHS England.

Khalifa, M., & Zabani, I. (2016). Utilizing health analytics in improving the performance of healthcare services: A case study on a tertiary care hospital. *Journal of Infection and Public Health*, *9*(6), 757–765.

Khatri, K. L. and L. S. Tamil (2018) Early Detection of Peak Demand Days of Chronic Respiratory Diseases Emergency Department Visits Using Artificial Neural Networks. *IEEE Journal of Biomedical and Health Informatics* 22(1): 285-290.

Kiani, F., Borjali, A., & Farahbakhsh, K. (2015). *Predictive Power of Work Situation Awareness by Work Overload and Job Stress among Workers : Implication for Occupational Health*. *3*(2).

Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., & Haydock, M. (2012). Analytics: The widening divide. *MIT Sloan Management Review*, *53*(2), 1.

Klaber, R. E., & Roland, D. (2012). Engaging paediatricians in planning the training of our future workforce. *Archives of disease in childhood*, *97*(6), 517-520.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE intelligent systems*, *21*(4), 70-73.

Klein, G. (2008). Naturalistic Decision Making. *Human Factors*, *50*(3), 456–460.

Klein, J. H., & Young, T. (2015). Health care: a case of hypercomplexity? *Health Systems*, *4*(2), 104–110.

Kluth, A., Jäger, J., & Bauernhansl, T. (2014). Evaluation of complexity management systems - Systematical and Maturity-Based approach. *Procedia CIRP*, *17*, 224–229.

Knowles, E., O'Cathain, A., & Nicholl, J. (2012). Patients' experiences and views of an emergency and urgent care system. *Health Expectations*, *15*(1), 78–86.

Koh, H. C., & Tan, G. (2005). Data mining applications in healthcare. *J Healthc Inf Manag*, *19*(2), 64–72.

Komashie, A; Mousavi, A; Clarkson, P J; Young, T. (2015). An Integrated Model of Patient and Staff Satisfaction Using Queuing Theory - Journal of Translational Engineering in Health and Medicine. *IEEE Journal of Translational Engineering in Health and Medicine*, *3*: 1–10.

Kosnik, L. (2013). Breakthrough demand–capacity management strategies to improve hospital flow, safety, and satisfaction. *International Series in Operations Research and Management Science*, *206*, 153–171.

Koster, J., Stewart, E., & Kolker, E. (2016). Health Care Transformation. *Academic Medicine*, *91*(2), 165–167.

Kotiadis, K., & Mingers, J. (2006). Combining PSMs with hard OR methods: the philosophical and practical challenges. *Journal of the Operational Research Society*, *57*(7), 856-867.

Kotiadis, K., Tako, A. A., & Vasilakis, C. (2013). A participative and facilitative conceptual modelling framework for discrete event simulation studies in healthcare. *Journal of the Operational Research Society*, *65*(2), 197–213.

Kotiadis, K., Tako, A. A., & Vasilakis, C. (2014). A participative and facilitative conceptual modelling framework for discrete event simulation studies in healthcare. *Journal of the Operational Research Society*, *65*(2), 197–213.

Kotiadis, K., & Tako, A. A. (2016). A facilitation workshop for the implementation stage: A case study in health care. *Proceedings of the Operational Research Society Simulation Workshop 2016, SW 2016*, 165–174.

Kotiadis, K., & Tako, A. A. (2018). Facilitated post-model coding in discrete event simulation (DES): A case study in healthcare. *European Journal of Operational Research*, *266*(3), 1120–1133.

Kraaijvanger, N., Rijpsma, D., Roovers, L., Van Leeuwen, H., Kaasjager, K., Van Den Brand, L., … Edwards, M. (2018). Development and validation of an admission prediction tool for emergency departments in the Netherlands. *Emergency Medicine Journal*, *35*(8).

Krebs, L. D., Kirkland, S. W., Chetram, R., Nikel, T., Voaklander, B., Davidson, A., … Rowe, B. H. (2017). Low-acuity presentations to the emergency department in Canada: Exploring the alternative attempts to avoid presentation. *Emergency Medicine Journal*, *34*(4), 249–255.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, *51*(11), 1016-1022.

Kuiler, E. W. (2014). From big data to knowledge: An ontological approach to big data analytics. *Review of Policy Research*, *31*(4), 311–318.

Kukafka, R., Johnson, S. B., Linfante, A., & Allegrante, J. P. (2003). Grounding a new information technology implementation framework in behavioral science: a systematic analysis of the literature on IT use. *Journal of biomedical informatics*, *36*(3), 218-227.

Kumar, A. Krishnamurthi, R.. Nayyar, A Sharma, K. Grover V.& Hossain, E. (2020) A Novel Smart Healthcare Design, Simulation, and Implementation Using Healthcare 4.0 Processes. In *IEEE Access*, 8: 118433-118471.

Kuo, YH., Rado, O., Lupia, B. *et al.* (2016) Improving the efficiency of a hospital emergency department: a simulation study with indirectly imputed service-time distributions. *Flex Serv Manuf J* 28**,** 120–147

Lampaki, A., & Papadakis, V. (2018). The impact of organisational politics and trust in the top management team on strategic decision implementation success: A middle-manager's perspective. *European Management Journal*, *36*(5), 627-637.

Land, L., & Meredith, N. (2013). An evaluation of the reasons why patients attend a hospital emergency department. *International Emergency Nursing*, *21*(1), 35–41.

Lane, D.C., Monefield, C. and Rosenhead, J.V. (2000) Looking in the wrong place for health care improvements: a system dynamics study of an accident and emergency department, Journal of the Operational Research Society, 51: 518-531.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, Vol. 36.

Lattimer, V., Brailsford, S., Turnbull, J., Tarnaras, P., Smith, H., George, S., … Maslin-Prothero, S. (2004). Reviewing emergency care systems I: insights from system dynamics modelling. *Emergency Medicine Journal : EMJ*, *21*(6), 685–691.

Law, A. M. (2009). How to build valid and credible simulation models. *Proceedings - Winter Simulation Conference*, 24–33.

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & quantity*, *43*(2), 265-275.

Lega, F., & Mengoni, A. (2008). Why non-urgent patients choose emergency over primary care services? Empirical evidence and managerial implications. *Health Policy*, *88*(2–3), 326–338.

Levin, S., Sauer, L., Kelen, G., Kirsch, T., Pham, J., Desai, S., & France, D. (2012). Situation awareness in emergency medicine. *IIE Transactions on Healthcare Systems Engineering*, *2*(2), 172-180.

Liberati, E. G., Ruggiero, F., Galuppo, L., Gorli, M., González-Lorenzo, M., Maraldi, M., ... & Vespignani, R. (2017). What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implementation Science*, *12*(1), 113.

Lin, W. D., & Chia, L. (2017). Combined Forecasting of Patient Arrivals and Doctor Rostering Simulation Modelling for Hospital Emergency Department. *Proceedings of the 2017 IEEE IEEM*, 2391–2395.

Linares, M. P., Montero, L., Barceló, J., & Carmona, C. (2016, December). A simulation framework for real-time assessment of dynamic ride sharing demand responsive transportation models. In *2016 Winter Simulation Conference (WSC)* (pp. 2216-2227). IEEE.

Liscott, E. (2016). What causes inappropriate or avoidable use of urgent and emergency care services , and how far can the ambulance service reduce this through admission avoidance ? *Brighton Journal of Research in Health Science*, (2013), 1–11.

Lee, K., & Fishwick, P. A. (1999). OOPM/RT: A multimodeling methodology for real-time simulation. *ACM Trans. Model. Comput. Simul.*, *9*(2), 141–170.

Lee, A. C. K., Blank, L., Payne, N., McShane, M., & Goyder, E. (2013). Demand management: misguided solutions?. *British Journal of General Practice*, *63*(613), e589-e591.

Loebbecke, C., van Fenema, P. C., & Powell, P. (2016). Managing inter-organizational knowledge sharing. *Journal of Strategic Information Systems*, *25*(1), 4–14.

Long, K. M., & Meadows, G. N. (2018). Simulation modelling in mental health : A systematic review Simulation modelling in mental health : A systematic review. *Journal of Simulation*, *12*(1), 1–10.

Long, K. M., McDermott, F., & Meadows, G. N. (2019). Factors affecting the implementation of simulation modelling in healthcare: A longitudinal case study evaluation. *Journal of the Operational Research Society*, 1–13.

Lorenz, T. (2009). Abductive fallacies with agent-based modeling and system dynamics. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5466 LNAI*, 141–152.

Low, M. Y. H., Turner, S. J., Ling, D., Peng, H. L., Chan, L. P., Lendermann, P., & Buckley, S. (2007). Symbiotic simulation for business process reengineering in high tech manufacturing and service networks. In *Proceedings of the 2007 Winter Simulation Conference*, (pp. 568–576). IEEE.

Lynch, C., Padilla, J., Diallo, S., Sokolowski, J., & Banks, C. (2014). A multi-paradigm modeling framework for modeling and simulating problem situations. In *Proceedings of the 2014 Winter Simulation Conference*, (pp.1688–1699). IEEE.

Lynch, C. J., & Diallo, S. Y. (2015). A taxonomy for classifying terminologies that describe simulations with multiple models. In *Proceedings 2015 Winter Simulation Conference*, (pp.1621–1632). IEEE.

Lyons, G. J., & Duggan, J. (2015). System dynamics modelling to support policy analysis for sustainable health care. *Journal of Simulation*. *9*(2), 129-139.

Ma, S., Zheng, Y., & Wolfson, O. (2014). Real-time city-scale taxi ridesharing. *IEEE Transactions on Knowledge and Data Engineering*, *27*(7), 1782-1795.

Macal, C. M. (2016). Everything You Need to Know about Agent-based Modelling and Simulation. *Journal of Simulation*, *10*(2): 144-156.

MacArthur, H., Phillips, C., Simpson, H., Bowling, M., Brace, A., Gozzard, D., ... & Matthias, J. (2012). Improving Quality Reduces Costs: Quality as the Business Strategy. Available at: https://qi.elft.nhs.uk/wp-content/uploads/2014/04/1000-lives-plus-improving-quality-reduces-costs1.pdf Accessed: 03.04.2020

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS ONE*, *13*(3).

Mane, K. K., Schmitt, C., Owen, P., Gersing, K., Ahalt, S. C., & Wilhelmsen, K. (2012). Data-driven approaches to augment clinical decision in EMR Era. In *2012 3rd International Workshop on Cognitive Information Processing (CIP)* (pp. 1-5). IEEE.

Marchand, C., & Peckham, S. (2017). Addressing the crisis of GP recruitment and retention: a systematic review. *Br J Gen Pract*, *67*(657), e227-e237.

Marmor, Y. N., Wasserkrug, S., Zeltyn, S., Mesika, Y., Greenshpan, O., Carmeli, B., … Mandelbaum, A. (2009). Toward simulation-based real-time decision-support systems for emergency departments. *Proceedings - Winter Simulation Conference*, 2042–2053.

Marshall, D. A., Burgos-liz, L., Eng, I., Ijzerman, M. J., Crown, W., Padula, W. V, … Osgood, N. D. (2015). Selecting a Dynamic Simulation Modeling Method for Health Care Delivery Research — Part 2 : Report of the ISPOR Dynamic Simulation Modeling Emerging Good Practices Task Force. *Value in Health*, *18*(2), 147–160.

Marshall, D. A., Burgos-Liz, L., Pasupathy, K. S., Padula, W. V., IJzerman, M. J., Wong, P. K., … Osgood, N. D. (2016). Transforming Healthcare Delivery: Integrating Dynamic Simulation Modelling and Big Data in Health Economics and Outcomes Research. *PharmacoEconomics*, *34*(2), 115–126.

Martin, N., Depaire, B., & Caris, A. (2018). A synthesized method for conducting a business process simulation study. In *2018 Winter Simulation Conference (WSC)* (pp. 276-290). IEEE.

Mason, S., Jacques, R., Rimmer, M., & Ablard, S. (2017). *Perspectives on the reasons for Emergency Department attendances across Yorkshire and the Humber Final Report* | *March 2017*.

Mbizvo, G. K., Bennett, K. H., Schnier, C., Simpson, C. R., Duncan, S. E., & Chin, R. F. (2020). The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies. *Epilepsia*.

McGeorge, N., Hegde, S., Berg, R. L., Guarrera-Schick, T. K., LaVergne, D. T., Casucci, S. N., ... & Benda, N. C. (2015). Assessment of innovative emergency department information displays in a clinical simulation center. *Journal of cognitive engineering and decision making*, *9*(4), 329-346.

McGuigan, T., & Watson, P. (2010). Non-urgent attendance at emergency departments. *Emergency Nurse : The Journal of the RCN Accident and Emergency Nursing Association*, *18*(6), 34–38.

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research*. Routledge.

Mehta, N., Pandit, A., & Shukla, S. (2019). Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. *Journal of biomedical informatics*, 103311.

Menke, N. B., Caputo, N., Fraser, R., Haber, J., Shields, C., & Menke, M. N. (2014). A retrospective analysis of the utility of an artificial neural network to predict ED volume. *The American journal of emergency medicine*, *32*(6), 614-617.

Meyners, M., & Castura, J. (2014). Check-All-That-Apply Questions. *Novel Techniques in Sensory Characterization and Consumer Profiling*, (September), 271–306.

Middleton, B., Bloomrosen, M., Dente, M. A., Hashmat, B., Koppel, R., Overhage, J. M., ... & Zhang, J. (2013). Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. *Journal of the American Medical Informatics Association*, *20*(e1), e2-e8.

Mielczarek, B. & Uziałko-Mydlikowska, J. (2012). Application of computer simulation modeling in the health care sector: A survey*. Simulation*, 88(2), 197–216.

Mielczarek, B. and Zabawa, J. (2016). Modeling healthcare demand using a hybrid simulation approach. In *2016 Winter Simulation Conference* (WSC), (pp. 1535-1546). IEEE.

Huberman, A. M., Miles, M., & Saldana, J. (2014). Qualitative data analysis: A methods sourcebook. *The United States of America: SAGE publications*.

Mingers, John (1997). Multi-Paradigm Multimethodology. In: Mingers, John and Gill, Anthony, eds. *Multimethodology: Theory and Practice of Combining Management Science Methodologies*. Wiley, Chichester, pp. 1-20.

Mingers, J. (2000). The contribution of critical realism as an underpinning philosophy for OR/MS and systems, *Journal of the Operational Research Society*, 51:11, 1256-1270.

Mingers, J. (2001). Combining IS Research Methods : Towards a Pluralist Methodology. *Information Systems Research*, *12*(3), 240–259.

Mingers, John (2008) Reaching the Problems that Traditional OR/MS Methods Cannot Reach. Working paper. University of Kent Canterbury, Canterbury

Mingers, J. (2011) Soft OR comes of age-but not everywhere! *Omega* 39 (6) 729–741.

Mingers, J. (2015). Helping business schools engage with real problems: The contribution of critical realism and systems thinking. *European Journal of Operational Research*, *242*(1).

Mingers, J., & Brocklesby, J. (1997). *Multimethodology : Towards a Framework for Mixing Methodologies*. *25*(5), 489–509.

Mingers, J., & Rosenhead, J. (2004). Problem structuring methods in action. *European journal of operational research*, *152*(3), 530-554.

Mohiuddin, S., Busby, J., Savovi, J., Richards, A., Northstone, K., Hollingworth, W., … Vasilakis, C. (2017). Patient flow within UK emergency departments : a systematic review of the use of computer simulation modelling methods. *BMJ Open*, 7.

Molloy, O., Onggo, B. S., Juan, A. A., Mustafee, N., & Smart, A. (2019). Symbiotic Simulation System: Hybrid Systems Model Meets Big Data Analytics. In *2018 Winter Simulation Conference (WSC)*, (pp. 1358–1369). IEEE.

Monks, T. (2015). Operational research as implementation science: definitions, challenges and research priorities. *Implementation Science*, *11*(1), 81.

Monks, T., & Meskarian, R. (2017). Using simulation to help hospitals reduce emergency department waiting times: examples and impact. In E. W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N.
Mustafee, G. Wainer, and E. Page (Ed.), *Proceedings of the 2017 Winter Simulation Conference*. IEEE.

Monitor (2014) Walk-in centre review: final report and recommendations. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/283778/WalkInCentreFinalReportFeb14.pdf Accessed: 15.09.2019

Moon, Y. B. (2017). Simulation modelling for sustainability : a review of the literature. *International Journal of Sustainable Engineering*, *7038*, 0.

Morecroft, J., & Robinson, S. (2005). Explaining puzzling dynamics: comparing the use of system dynamics and discrete-event simulation. In *Proceedings of the 23rd International Conference of the System Dynamics Society* (pp. 17-21). Boston, MA: System Dynamics Society.

Morgan, J. S., Howick, S., & Belton, V. (2017). A toolkit of designs for mixing Discrete Event Simulation and System Dynamics. *European Journal of Operational Research*, *257*(3), 907–918.

Morley, C., Unwin, M., Peterson, G. M., Stankovich, J., & Kinsman, L. (2018). Emergency department crowding: A systematic review of causes, consequences and solutions. In *PLoS ONE* (Vol. 13).

Morris, Z. S., Boyle, A., Beniuk, K., & Robinson, S. (2012). Emergency department crowding: Towards an agenda for evidence-based intervention. *Emergency Medicine Journal*, *29*(6), 460–466.

Morrow, R. A., & Brown, D. D. (1994). *Critical theory and methodology* (Vol. 3). Sage.

Mortenson, M. J., Doherty, N. F., & Robinson, S. (2015). Operational research from Taylorism to Terabytes: A research agenda for the analytics age. *European Journal of Operational Research*, *241*(3), 583-595.

Morzuch, B. J., & Allen, P. G. (2006). Forecasting hospital emergency department arrivals. In *26th Annual Symposium on Forecasting, Santander, Spain* (pp. 11-14).

Mowen, J. C., Licata, J. W., & McPhail, J. (1993). Waiting in the emergency room: How to improve patient satisfaction. *Journal of Health Care Marketing*, 13(2):26.

Mousavi, A., Komashie, A., & Tavakoli, S. (2011). SIMulation-based real-time performance MONitoring (SIMMON): A platform for manufacturing and healthcare systems. *Proceedings - Winter Simulation Conference*, 600–611.

Moxley, J. H., Ericsson, A. K., Charness, N., & Krampe, R. T. (2012). The role of intuition and deliberative thinking in experts' superior tactical decision-making. *Cognition*, *124*(1), 72–78.
Mullins, L. J. (2007). *Management and organisational behaviour*. Pearson education.

Munton, T., Martin, A., Marrero, I., Llewellyn, A., Gibson, K. & Gomersall, A. (2011) Evidence: getting out of hospital? The Health Foundation. Available at: https://www.health.org.uk/sites/default/files/GettingOutOfHospital_fullversion.pdf Accessed: 15.03.20

Murphy, D. R., Satterly, T., Giardina, T. D., Sittig, D. F., & Singh, H. (2019). Practicing clinicians' recommendations to reduce burden from the electronic health record inbox: a mixed-methods study. *Journal of general internal medicine*, *34*(9), 1825-1832.

Murray, R., Jabbal, J., Maguire, D. & Ward, D. (2018) How is the NHS performing? March 2018 quarterly monitoring report. The King's Fund. Available at: https://www.kingsfund.org.uk/publications/how-nhs-performing-march-2018 Accessed: 23.09.19

Mustafee, N., Katsaliaki, K., & Taylor, S. J. (2010). Profiling literature in healthcare simulation. *Simulation*, *86*(8-9), 543-558.

Mustafee, N., Powell, J., Brailsford, S. C., Diallo, S., Padilla, J., & Tolk, A. (2015a). Hybrid simulation studies and hybrid simulation systems: definitions, challenges, and benefits. In E. L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti (Ed.), *Proceedings of the 2015 Winter Simulation Conference*, IEEE

Mustafee, N., Brailsford, S. C., Djanatliev, A., Eldabi, T., Kunc, M., & Tolk, A. (2017a). Purpose and benefits and hybrid simulation: contributing to the convergence of its definition. In *2017 Winter Simulation Conference*. IEEE

Mustafee, N., Powell, J. H., Martin, S., Fordyce, A., & Harper, A. (2017b). Investigating the use of real-time data in nudging patients' emergency department (ED) attendance behaviour. In: *Proceedings of the* 2017*Simulation Workshop*. IEEE.

Mustafee, N., Powell, J. H., & Harper, A. (2018). RH-RT: a data analytics framework for reducing wait time at emergency departments and centres for urgent care. In *2018 Winter Simulation Conference (WSC)* (pp. 100-110). IEEE.

Mustafee, N., & Powell, J. H. (2018). From hybrid simulation to hybrid systems modelling. In *Proceedings of the 2018 Winter Simulation Conference*, (pp.1430–1439). IEEE.

Mustafee, N., & Katsaliaki, K. (2020). Classification of the existing knowledge base of OR/MS research and practice (1990-2019) using a proposed classification scheme. *Computers & Operations Research*, 104920

Naderpour, M., Lu, J., & Zhang, G. (2016). A safety-critical decision support system evaluation using situation awareness and workload measures. *Reliability Engineering and System Safety*, *150*, 147–159.

Næss, P. (2015). Critical realism, urban planning and urban research. *European Planning Studies*, *23*(6), 1228-1244.

Nagree, Y., Camarda, V. J., Fatovich, D. M., Cameron, P. A., Dey, I., Gosbell, A. D., … Mountain, D. (2013). Quantifying the proportion of general practice and low-acuity patients in the emergency department. *Medical Journal of Australia*, *198*(11), 612–615.

Nahhas, A., Awaldi, A. and Reggelin, T. (2017). Simulation and the Emergency Department Overcrowding Problem. In: *16th Conference on Reliability and Statistics in Transportation and Communication,* RelStat'2016, 19-22 October, 2016, Latvia.

Nählinder, S., & Berggren, P. (2002). Dynamic Assessment of Operator Status. Paper presented at the Proceedings of the HFES 2002 Conference., Baltimore, MD, USA

Nählinder, S., Berggren, P., & Svensson, E. (2004). Reoccurring LISREL patterns describing mental workload, situation awareness and performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*, 2485–2488.

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1)

Nau, R. (2019) Statistical forecasting: notes on regression and time series analysis. Fuqua School of Business, Duke University. Available at: https://people.duke.edu/~rnau/411home.htm Accessed: 23.01.20

Nelson, J. (2011). Why patients visit emergency units rather than use primary care services. *Emergency Nurse*, *19*(1), 32–36.

NHS Operational Planning and Contracting Guidance 2019/20. Available at: https://www.england.nhs.uk/wp-content/uploads/2018/12/nhs-operational-planning-and-contracting-guidance.pdf Accessed: 24.01.20

NHSDigital (2019a) Non-urgent A&E attendances. Available at: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/innovative-uses-of-data/demand-on-healthcare/unnecessary-a-and-e-attendances Accessed: 26.09.19

NHSDigital (2019b) Hospital Accident & Emergency Activity 2018-19. Available at: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-accident--emergency-activity/2018-19 Accessed: 14.02.20

NHSEngland (2019) Data Services. Available at: https://www.england.nhs.uk/data-services/validate/ Accessed: 14.02.20

NHSEngland (2020a) Urgent and Emergency Care. Available at: https://www.england.nhs.uk/urgent-emergency-care/ Accessed: 14.02.20

NHSEngland (2020b) Advanced Forecasting Techniques. Available at: https://www.england.nhs.uk/publication/advanced-forecasting-techniques/ Accessed: 16.07.20

NHS Improvement (2019) Quarterly performance of the NHS provider sector: quarter 4 2018/19. Available at: https://improvement.nhs.uk/resources/quarterly-performance-nhs-provider-sector-quarter-4-201819/ Accessed: 03.03.20

NHS Improvement (2018) Emergency Care Workforce Programme. Available at: https://improvement.nhs.uk/resources/emergency-care-workforce-programme/ Accessed: 23.10.19

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In T. Plomp, & N. Nieveen (Eds.), Educational design research – part A: An introduction (pp. 152-169). Enschede, the Netherlands: SLO.

Oakley, D., Onggo, B. S., & Worthington, D. (2020). Symbiotic simulation for the operational management of inpatient beds: model development and validation using Δ-method. *Health care management science*, 23, 153-169.

Ocloo, J., & Matthews, R. (2016). From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. *BMJ Qual Saf*, *25*(8), 626-632.

Offermann, P., Blom, S., & Bub, U. (2011). Strategies for creating, generalising and transferring design science knowledge–a methodological discussion and case analysis. In *10th International Conference on Wirtschaftsinformatik* (pp. 1187-1196). Zurich, Switzerland.

O'keefe, R. (2014). Design Science, the design of systems and Operational Research: back to the future?. *Journal of the Operational Research Society*, *65*(5), 673-684.

Oliva, R., & Sterman, J. D. (2001). Cutting corners and working overtime: Quality erosion in the service industry. *Management Science*, *47*(7), 894-914.

Onggo, B. S. S. and J. Hill. (2014). Data Identification and Data Collection Methods in Simulation : A Case Study at ORH Ltd. Journal of Simulation 8(3):195-205.

Onggo, B. S., Mustafee, N., Smart, A., Juan, A. A., & Molloy, O. (2018). Symbiotic simulation system: hybrid systems model meets big data analytics. In *Proceedings of the 2018 winter simulation conference* (pp. 1358-1369). IEEE.

Onggo, B. S. (2019). Symbiotic simulation system (S3) for industry 4.0. In *Simulation for Industry 4.0* (pp. 153-165). Springer, Cham.

Onggo, B. S., Corlu, C. G., Juan, A. A., Monks, T., & de la Torre, R. (2020). Combining symbiotic simulation systems with enterprise data storage systems for real-time decision-making. *Enterprise Information Systems*, 1-18.

Onggo, B. S., and M. Karatas. (2016). Test-driven simulation modelling: a case study using agent-based maritime search-operation simulation. *European Journal of Operational Research* 254 (2): 517–531

Ormerod, R. J. (2001). The Success and Failure of Methodologies-A Comment on Connell (2001): Evaluating Soft OR. *The Journal of the Operational Research Society*, *52*(10), 1176–1179.

Ostrowski, Ł., Helfert, M., & Hossain, F. (2011). A Conceptual Framework for Design Science Research. *Perspectives in Business Informatics Research*, 345-354.

Padilla, J. J., Tolk, A., & Diallo, S. Y. (2013). M&S methodological challenges. In *Proceedings of the Emerging M&S Applications in Industry & Academia/Modeling and Humanities Symposium* (pp. 1-9).

Palmer, R., Fulop, N. J., & Utley, M. (2018). A systematic literature review of operational research methods for modelling patient flow and outcomes within community healthcare and other settings. *Health Systems*, *7*(1), 29-50.

Pandas (2014) pandas.DataFrame.resample Available at: https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.resample.html Accessed: 03.03.20

Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of workload. In P. A. Hancock & P. E. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 305–320). Mahwah, NJ: Erlbaum.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation Awareness , Mental Workload , and Trust in Automation : Viable , Empirically Supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making 2(2), 140–160.*

Park, E. H., Park, J., Ntuen, C., Kim, D., & Johnson, K. (2008). Forecast Driven Simulation Model for Service Quality Improvement of the Emergency Department in the Moses H. Cone Memorial Hospital. *Asian Journal on Quality*, *9*(3), 1–14.

Patton, G. G., & Thakore, S. (2012). Reducing inappropriate emergency department attendances:a review of ambulance service attendances at a regional teaching hospital in Scotland. *Emergency Medicine Journal*, 30(6), 459-461.

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

Pearson, M., Monks, T., Gibson, A., Allen, M., Komashie, A., Fordyce, A., ... & Stein, K. (2013). Involving patients and the public in healthcare operational research—The challenges and opportunities. *Operations Research for Health Care*, *2*(4), 86-89.

Pearson, C., Kim, D. S., Mika, V. H., Imran Ayaz, S., Millis, S. R., Dunne, R., & Levy, P. D. (2018). Emergency department visits in patients with low acuity conditions: Factors associated with resource utilization. *American Journal of Emergency Medicine*, *36*(8), 1327–1331.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, *24*(3), 45-77.

Peffers, K., Tuunanen, T., & Niehaves, B. (2018). Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. *European Journal of Information Systems*, *27*(2).
Pennathur, P. R., Cao, D., Bisantz, A. M., Lin, L., Fairbanks, R. J., Wears, R. L., ... & Sui, Z. (2011). Emergency department patient-tracking system evaluation. *International Journal of Industrial Ergonomics*, *41*(4), 360-369.

Penson, R., Coleman, P., Mason, S., & Nicholl, J. (2012). Why do patients with minor or moderate conditions that could be managed in other settings attend the emergency department? *Emergency Medicine Journal*, *29*(6), 487–491.

Peral, J., Maté, A., & Marco, M. (2017). Application of Data Mining techniques to identify relevant Key Performance Indicators. *Computer Standards & Interfaces*, *50*(April), 55–64.

Perez, J., De Looze, M. P., Bosch, T., & Neumann, W. P. (2014). Discrete event simulation as an ergonomic tool to predict workload exposures during systems design. *International Journal of Industrial Ergonomics*, *44*(2), 298-306.

Pessôa, L. A. M., Lins, M. P. E., Da Silva, A. C. M., & Fiszman, R. (2015). Integrating soft and hard operational research to improve surgical centre management at a university hospital. *European Journal of Operational Research*, *245*(3), 851–861.

Peute, L. W., Driest, K. F., Marcilly, R., Da Costa, S. B., Beuscart-Zephir, M. C., & Jaspers, M. W. (2013). A framework for reporting on human factor/usability studies of health information technologies. In *CSHI* (pp. 54-60).

Pfannstiel, M. A., & Rasche, C. (2017). *Service Business Model Innovation in Healthcare and Hospital Management*. Springer.

Pitt, M., Monks, T., Crowe, S., & Vasilakis, C. (2016). Systems modelling and simulation in health service design, delivery and decision making. *BMJ Quality and Safety*, *25*(1), 38–45.

Pleasants, N. (2002). *Wittgenstein and the idea of a critical social theory: A critique of Giddens, Habermas and Bhaskar*. Routledge.

Plomp, T., & Nieveen, N. (2007). An introduction to educational design research. In *Proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China)* (Vol. 23).

Pope, C., McKenna, G., Turnbull, J., Prichard, J., & Rogers, A. (2019). Navigating and making sense of urgent and emergency care processes and provision. *Health Expectations*, *22*(3), 435-443.

Porter ME (2010). What is value in health care? *New England Journal of Medicine*, 363(26), 2477–81.

Powell, J. H., & Bradford, J. P. (2000). Targeting intelligence gathering in a dynamic competitive environment. *International Journal of Information Management*, *20*(3), 181-195.

Powell, J., & Mustafee, N. (2014). Soft OR Approaches in Problem Formulation Stage of a hybrid M&S Study. In E. A. Tolk, S. D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller (Ed.), *Proceedings of the 2014 Winter Simulation Conference.* IEEE.

Powell, J. H., & Mustafee, N. (2016). Widening requirements capture with soft methods: an investigation of hybrid M&S studies in health care. *Journal of the Operational Research Society*, *68*(10), 1211-1222.

Powell, A. & Davies, H. (2016) Managing Doctors, Doctors Managing. Research Report, Nuffield Trust. Available at: https://www.nuffieldtrust.org.uk/files/2017-01/doctors-managers-lit-review-web-final.pdf Accessed: 12.03.20

Proudlove, N. C., Bisogno, S., Onggo, B. S. S., Calabrese, A., & Ghiron, N. L. (2017). *Towards fully-facilitated discrete event simulation modelling : Addressing the model coding stage*. *263*, 583–595.

Quan SD, Morra D, Lau FY, et al. (2013). Perceptions of urgency: defining the gap between what physicians and nurses perceive to be an urgent issue. *Int J Med Inform*. 82:378-386

Rabe, M., S. Spieckermann, and S. Wenzel. (2008). A New Procedure Model for Verification and Validation in Production and Logistics Simulation. In *Proceedings of the 2008 Winter Simulation Conference*, edited by S.J. Mason et al., 1717-1726. Piscataway, New Jersey: IEEE.

Rabelo, L., Cruz, L., Bhide, S., Joledo, O., Pastrana, J., & Xanthopoulos, P. (2014). Analysis of the expansion of the panama canal using simulation modeling and artificial intelligence. In *Proceedings of the 2014 Winter Simulation Conference* (pp. 910-921). IEEE.

Rafferty, L. A., Stanton, N. A., & Walker, G. H. (2013). Great expectations: A thematic analysis of situation awareness in fratricide. Safety Science, 56, 63–71.

Raghupathi, W., & Raghupathi, V. (2013). An Overview of Health Analytics. *J Health Med Informat*, *4*(3).

Ranawat, A. S., Koenig, J. H., Thomas, A. J., Krna, C. D., & Shapiro, L. A. (2009). Aligning physician and hospital incentives: the approach at hospital for special surgery. *Clinical Orthopaedics and Related Research®*, *467*(10), 2535-2541.

Ranyard, J. C., Fildes, R., & Hu, T. I. (2015). Reassessing the scope of OR practice: The influences of problem structuring methods and the analytics movement. *European Journal of Operational Research*, *245*(1), 1-13.

Rasheed, M. B., Javaid, N., Malik, M. S. A., Asif, M., Hanif, M. K., & Chaudary, M. H. (2019). Intelligent multi-agent based multilayered control system for opportunistic load scheduling in smart buildings. *IEEE Access*, *7*, 23990-24006.

Rashwan, W., Abo-Hamad, W., & Arisha, A. (2015). A system dynamics view of the acute bed blockage problem in the Irish healthcare system. *European Journal of Operational Research*, *247*(1), 276-293.

Rasmussen, T., & Ulrich, D. (2015). Learning from practice: How HR analytics avoids being a management fad. *Organizational Dynamics*, *44*(3), 236–242.

RCEM (2019) Improving Quality Indicators and System Metrics for Emergency Departments in England. Royal College of Emergency Medicine Position Statement. Available at: https://www.rcem.ac.uk/docs/Policy/RCEM_position_statement_Improving_quality_indicators_and_system_metrics.pdf Accessed: 02.02.20

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), 491-497.

Richey, R. C., & Klein, J. D. (2014). Design and development research. In *Handbook of research on educational communications and technology* (pp. 141-150). Springer, New York, NY.

Ridder, H. G. (2017). The theory contribution of case study research designs. *Business Research*, *10*(2), 281-305.

Risser, D. T., Rice, M. M., Salisbury, M. L., Simon, R., Jay, G. D., Berns, S. D., & MedTeams Research Consortium. (1999). The potential for improved teamwork to reduce medical errors in the emergency department. *Annals of emergency medicine*, *34*(3), 373-383.

Rivas, A., Chamoso, P., & Rodríguez, S. (2017). An agent-based internet of things platform for distributed real time machine control. In *2017 IEEE 17th International Conference on Ubiquitous Wireless Broadband (ICUWB)* (pp. 1-5). IEEE.

Riveiro, M., Falkman, G., & Ziemke, T. (2008). Improving Maritime Anomaly Detection and Situation Awareness through Interactive Visualization. *2008 11th International Conference on Information Fusion*, 47–54.

Robert, G., Cornwell, J., Locock, L., Purushotham, A., Sturmey, G., & Gager, M. (2015). Patients and staff as codesigners of healthcare services. *BMJ*, *350*, g7714.

Robert G, Greenhalgh T, MacFarlane F, Peacock R. (2010). Adopting and assimilating new non-pharmaceutical technologies into health care: a systematic review. *J Health Serv Res Policy* 15(4):243-250.

Robertson, R. (2019). Why is public satisfaction with the NHS still falling? The Kings Fund. Available at: https://www.kingsfund.org.uk/blog/2019/03/why-public-satisfaction-nhs-still-falling Accessed: 15.06.19

Rodič, B. (2017). Industry 4.0 and the new simulation modelling paradigm. *Organizacija*, *50*(3), 193-207.

Robinson, S. (2002). General concepts of quality for discrete-event simulation. *European Journal of Operational Research*, *138*(1), 103–117.

Robinson, S. (2004). *Simulation*. John Wiley & Sons, Ltd.

Robinson, S. (2006). Conceptual modeling for simulation: issues and research requirements. *Proceedings of the 38th Conference on Winter Simulation*, (1994), 800.

Robinson, S. (2008). Conceptual modelling for simulation Part II : a framework for conceptual modelling. *Journal of the Operational Research Society* 59, 291 --304

Robinson, A., Levis, J., & Bennett, G. (2010). INFORMS to officially join analytics movement. *OR/MS Today*, *37*(5), 59.

Robinson, S., Radnor, Z. J., Burgess, N., & Worthington, C. (2012). SimLean: Utilising simulation in the implementation of lean in healthcare. *European Journal of Operational Research*, *219*(1), 188–197.

Robinson, S. (2013). Conceptual modeling for simulation. *Proceedings of the 2013 Winter Simulation Conference R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds*, 377–388. IEEE.

Rogers, P., & Gordon, R. J. (1993). Simulation for real-time decision making in manufacturing systems. In *Proceedings of the 25th conference on winter simulation* (pp. 866-874). IEEE.

Rojas, E., Munoz-gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare : A literature review. *Journal of Biomedical Informatics*, *61*, 224–236.

Ross, J., Stevenson, F., Lau, R., & Murray, E. (2016). Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). *Implementation science*, *11*(1), 146.

Rowe, B., Cook, C., Wootton, R., & Brown, T. (2015). *A&E: Studying parental decision makıng around non-urgent attendance among under 5s*. (June). Department of Health.

Rowlands, T., Waddell, N., & McKenna, B. (2016). Are we there yet? A technique to determine theoretical saturation. *Journal of Computer Information Systems*, *56*(1), 40-47.

Royston, G. (2009). One Hundred Years of Operational Research in Health: UK 1948-2048. *Journal of the Operational Research Society*, s169-s179.

Royston, G. (2011). Meeting global health challenges through operational research and management science. *Bulletin of the World Health Organization*, *89*, 683-688.

Royston, G. (2013). Operational Research for the Real World: big questions from a small island. *Journal of the Operational Research Society*, *64*(6), 793-804.

Rudin, C., and Wagstaff, K. L. (2014). Machine learning for science and society. *Mach.Learn.* 95, 1–9.

Ruiz-Torres, A. J., & Nakatani, K. (1998). Application of real-time simulation to assign due dates on logistic-manufacturing networks. *Proceedings of the 1998 Winter Simulation Conference D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, eds.,* 1205–1210

Rutberg, M. H., Wenczel, S., Devaney, J., Goldlust, E. J., & Day, T. E. (2015). Incorporating Discrete Event Simulation Into Quality Improvement Efforts in Health Care Systems. *American Journal of Medical Quality*, *30*(1), 31–35.

Sachdeva, R., Williams, T., & Quigley, J. (2007). Mixing methodologies to enhance the implementation of healthcare operational research. *Journal of the Operational Research Society*, *58*(2), 159–167.

Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human factors*, *37*(1), 123-136.

Salas, E., Rosen, M. a., & DiazGranados, D. (2010). Expertise-Based Intuition and Decision Making in Organizations. *Journal of Management*, *36*(4), 941–973.

Salehi, V., & McMahon, C. (2009). Action research into the use of parametric associative CAD systems in an industrial context. In *DS 58-5: Proceedings of ICED 09, the 17th International Conference on Engineering Design, Vol. 5, Design Methods and Tools (pt. 1), Palo Alto, CA, USA, 24.-27.08. 2009*.

Salleh, S., Thokala, P., Brennan, A., Hughes, R., & Booth, A. (2017). Simulation Modelling in Healthcare : An Umbrella Review of Systematic Literature Reviews. *PharmacoEconomics*, *35*(9), 937–949.

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297-323.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D. P., & Rafferty, L. (2010). Is it really better to share? Distributed situation awareness and its implications for collaborative system design. *Theoretical Issues in Ergonomics Science*, *11*(1–2), 58–83.

Salmon, A., Rachuba, S., Briscoe, S., & Pitt, M. (2018). A structured literature review of simulation modelling applied to Emergency Departments: Current patterns and emerging trends. *Operations Research for Health Care*, 1–13.

Sampson, R., Barbour, R., & Wilson, P. (2016). The relationship between GPs and hospital consultants and the implications for patient care: a qualitative study. *BMC family practice*, *17*(1), 45.

Sancton, K., Sloss, L., Ccfp, M. D., Berkowitz, J., Strydom, N., Chb, M. B., & Mccracken, R. (2018). Low-acuity presentations to the emergency department. *Canadian Family Physician*, *64*(August), e354-60.

Sanjay, M., & Allamma, B. H. (2016). An Insight into Big Data Analytics – Methods and Application. *Proceedings of the International Conference on Inventive Computation Technologies, ICICT 2016, 1 , Art. No. 7823269*.

Sato, A., & Huang, R. (2015). From data to knowledge: A cognitive approach to retail business intelligence. In *2015 IEEE International Conference on Data Science and Data Intensive Systems* (pp. 210-217). IEEE.

Saxena, R., & Srinivasan, A. (2013). Business intelligence. In *Business Analytics* (pp. 85-99). Springer, New York, NY.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Routledge.

Schulz, C. M., Endsley, M. R., Kochs, E. F., Gelb, A. W., & Wagner, K. J. (2013). Situation Awareness in AnesthesiaConcept and Research. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, *118*(3), 729-742.

Seekhao, N., Shung, C., JaJa, J., Mongeau, L., & Li-Jessen, N. Y. (2016). Real-time agent-based modeling simulation with in-situ visualization of complex biological systems: A case study on vocal fold inflammation and healing. In *2016 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 463-472). IEEE.

Seila, A. F., & Brailsford, S. (2009). Opportunities and challenges in health care simulation. In *Advancing the Frontiers of Simulation* (pp. 195-229). Springer, Boston, MA.

Shachak, A., Kuziemsky, C., & Petersen, C. (2019). Beyond TAM and UTAUT: Future directions for HIT implementation research. *Journal of biomedical informatics*, *100*, 103315.

Shannon, R. E. (1998). Introduction to the art and science of simulation. *Proceedings of the 1998 Winter Simulation Conference*, 835–840.

Shanthikumar, J. G., & Sargent, R. G. (1983). A Unifying View of Hybrid Simulation / Analytic Models and Modeling *Operations Research, 31*(6), 1030–1052.

Shao, G., Shin, S. J., & Jain, S. (2014). Data analytics using simulation for smart manufacturing. In *Proceedings of the Winter Simulation Conference 2014* (pp. 2192-2203). IEEE.

Sharda, R., Delen, D., & Turban, E. (2016). *Business intelligence, analytics, and data science: a managerial perspective*. Pearson.

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, *23*(4), 433–441.

Shaw, E. K., Howard, J., Clark, E. C., Etz, R. S., Arya, R., & Tallia, A. F. (2013). Decision-making processes of patients who use the emergency department for primary care needs. *Journal of Health Care for the Poor and Underserved*, *24*(3), 1288–1305.

Simpao, A. F., Ahumada, L. M., & Rehman, M. A. (2015). Big data and visual analytics in anaesthesia and health care. *British Journal of Anaesthesia*, Vol. 115, pp. 350–356.

Sinreich,D. and Jabali,O. (2007) Staggered work shifts: away to downsize and restructure an emergency department workforce yet maintain current operational performance. *Health Care Management Science*, 10(3), 293–308.

Smith, D. P., & McNally, A. (2014). Delivering enhanced safety, productivity and experience: early results from a frequent caller management system. *Journal of Paramedic Practice*, *6*(12), 634-641.

Son Y, Wysk RA. (2001) Automatic simulation model generation for simulation-based, real-time shop floor control. Comput Industry;45(3):291–308.

Son, C., Sasangohar, F., & Peres, S. C. (2017). Redefining and Measuring Resilience in Emergency Management Systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1651-1652). Sage CA: Los Angeles, CA: SAGE Publications.

Song, L., Ramos, F., & Arnold, K. (2008). A framework for real-time simulation of heavy construction operations. In J. W. F. eds. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson (Ed.), *Proceedings of the 2008 Winter Simulation Conference* (pp. 2387–2395). IEEE.

Soremekun, O. A., Takayesu, J. K., and Bohan, S. J. (2011). Framework for analyzing wait times and other factors that impact patient satisfaction in the emergency department. *The Journal of Emergency Medicine*, 41(6):686 – 692.

Southon, G., Sauer, C., & Dampney, K. (1999). Lessons from a failed information systems initiative: issues for complex organisations. *International journal of medical informatics*, *55*(1), 33-46.

Sovacool, B. K., Axsen, J., & Sorrell, S. (2018). Promoting novelty, rigor, and style in energy social science: towards codes of practice for appropriate methods and research design. *Energy Research & Social Science*, *45*, 12-42.

Soyiri, I. N., & Reidpath, D. D. (2013). An overview of health forecasting. *Environmental Health and Preventive Medicine*, *18*(1), 1–9.

Spruit, M., Vroon, R., & Batenburg, R. (2014). Towards healthcare business intelligence in long-term care: an explorative case study in the Netherlands. *Computers in Human Behavior*, *30*, 698-707.

Stake, R. E. (1995). *The art of case study research*. Sage.

Stanton, N. A., Stewart, R., Harris, D., Houghton, R. J., Baber, C., McMaster, R., … Green, D. (2006). Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics*, *49*(12–13), 1288–1311.

Stanton, N. A., Salmon, P. M., & Walker, G. H. (2015). Let the reader decide: A paradigm shift for situation awareness in sociotechnical systems. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 44-50.

Stanton, N. A (2016) Distributed situation awareness, *Theoretical Issues in Ergonomics Science*, 17:1, 1-7

Stanton, N. A., Salmon, P. M., Walker, G. H., Salas, E., & Hancock, P. A. (2017). State-of-science: situation awareness in individuals, teams and systems. *Ergonomics*, *60*(4), 449–466.

Statsmodels (2019a) statsmodels.tsa.statespace.sarimax.SARIMAX. Available at: statsmodels.tsa.statespace.sarimax.SARIMAX Accessed: 12.10.19

Statsmodels (2019b) statsmodels.tsa.stattols.adfuller. Available at: https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html Accessed: 12.10.19

Stewart, B. (2016) Simple Linear Regression. Available at: https://scholar.princeton.edu/sites/default/files/bstewart/files/lecture5handout.pdf Accessed: 12.10.19

Stewart J, Williams R. (2005) The wrong trousers? Beyond the design fallacy: social learning and the user. In: Rohracher H, editor. User Involvement in Innovation Processes. Strategies and Limitations From a Socio-Technical Perspective. Munich, Germany: Profil-Verlag

Sukanya, J. M. C. A., Phil, M., Sc, S. V. K. M., & Phil, M. (2017). *Applications of Big Data Analytics and Machine Learning Techniques in Health Care Sectors*. *6*(7), 21963–21967.

Sun BC, Hsia RY, Weiss RE, et al. (2013) Effect of emergency department crowding on outcomes of admitted patients. Ann Emerg Med. 2013;61:605–11 e606.

Svensson, E., & Wilson, G., F. (2002). Psychological and Psychophysiological Models of Pilot Performance for Systems Development and Mission Evaluation. The International Journal of Aviation Psychology, 12(1), 95-110.

Syed, J., & Mingers, J. (2018). A Critical Realist Pathway to Relevant and Ethical Research. In *The Palgrave Handbook of Knowledge Management* (pp. 667-690). Palgrave Macmillan, Cham.

Tako, A. A., & Robinson, S. (2015). Is simulation in health different and is it more difficult? *Proceedings of the 6th Operational Research Society Simulation Workshop 2012 (SW 2012)*, 92–101.

Tako, A. A., & Kotiadis, K. (2015). PartiSim : A multi-methodology framework to support facilitated simulation modelling in healthcare. *European Journal of Operational Research*, *244*(2), 555–564.

Tan, K. W., Tan, W. H., & Lau, H. C. (2013). *Improving patient length-of-stay in emergency department through dynamic resource allocation policies. IEEE International Conference on Automation Science and Engineering (CASE)* 984–989.

Taylor, S. J. E., Khan, A., Morse, K. L., Tolk, A., Yilmaz, L., Zander, J., & Mosterman, P. J. (2015). *Grand challenges for modeling and simulation : simulation everywhere — from cyberinfrastructure to clouds to citizens*. *91*(7), 648–665.

Taylor, S. J. (2019). Distributed simulation: state-of-the-art and potential for operational research. *European Journal of Operational Research*, *273*(1), 1-19.

Tavakoli, S., Mousavi, A., & Komashie, A. (2008). A generic framework for real-time discrete event simulation (DES) modelling. *Proceedings of the 2008 Winter Simulation Conference*, 1931–1938.

Tomar, D., & Agarwal, S. (2013). A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio-Technology*, *5*(5), 241–266.

Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi journal of anaesthesia*, *11*(Suppl 1), S80.

Turner, J., O'Cathain, A., Knowles, E., & Nicholl, J. (2013). Impact of the urgent care telephone service NHS 111 pilot sites: a controlled before and after study. *Bmj Open*, *3*(11).

Turner, S. J. (2011). Symbiotic Simulation and Its Application to Complex Adaptive Systems. *2011 IEEE/ACM 15th International Symposium on Distributed Simulation and Real Time Applications*, 3–3.

The Kings Fund (2010). Approaches to demand management: commissioning in a cold climate. London: The King's Fund. Available at: http://www.kingsfund.org.uk/current_projects/quality_in_a_cold_climate/approaches_to_demand.html Accessed: 12.10.19.

The Kings Fund (2019). The NHS budget and how it has changed. Available at: https://www.kingsfund.org.uk/projects/nhs-in-a-nutshell/nhs-budget Accessed: 12.02.20.

Thistlethwaite P (2011). Integrating Health and Social Care in Torbay. Improving Care for Mrs Smith. The King's Fund, London. http://bit.ly/1ndPtkt
Thompson, D. A., Yarnold, P. R., Williams, D. R., and Adams, S. L. (1996). Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. *Annals of Emergency Medicine*, 28(6):657 – 665.

Thornton, J. (2017). Government is warned over" serious" risks to NHS this winter. *BMJ: British Medical Journal (Online)*, *359*.

Thuemmler, C. & Bai, C. (2017). Health 4.0: How Virtualization and Big Data are Revolutionizing Healthcare. Springer.

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human resource development review*, *4*(3), 356-367.

Trkman, P., Mccormack, K., Paulo, M., Oliveira, V. De, & Bronzo, M. (2010). The impact of business analytics on supply chain performance. *Decision Support Systems*, *49*(3), 318–327.

Tsai, J. C. H., Liang, Y. W., & Pearson, W. S. (2010). Utilization of emergency department in patients with non-urgent medical problems: Patient preference and emergency department convenience. *Journal of the Formosan Medical Association*, *109*(7), 533–542.

Turner, C. J., Tiwari, A., Olaiya, R., & Xu, Y. (2012). Process mining: from theory to practice. *Business Process Management Journal*. *18*(3), 493.

Uhlemann, T. H. J., Lehmann, C., & Steinhilper, R. (2017). The digital twin: Realizing the cyber-physical production system for industry 4.0. *Procedia Cirp*, *61*, 335-340.

Unwin, M., Kinsman, L., & Rigby, S. (2016). Why are we waiting? Patients' perspectives for accessing emergency department services with non-urgent complaints. *International Emergency Nursing*, *29*, 3–8.

Uriarte, A. G., Zúñiga, E. R., Moris, M. U., & Ng, A. H. (2017). How can decision makers be supported in the improvement of an emergency department? A simulation, optimization and data mining approach. *Operations Research for Health Care*, *15*, 102-122.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. Routledge.

van den Berg, M. J., van Loenen, T., Westert, G.P (2016). Accessible and continuous primary care may help reduce rates of emergency department use. An international survey in 34 countries, *Family Practice*, 33(1): 42–50.

Van Lent, W. A., Vanberkel, P., & Van Harten, W. H. (2012). A review on the relation between simulation and improvement in hospitals. *BMC Medical Informatics and Decision Making*, *12*(1).

Varshney, L. R. (2016). Fundamental limits of data analytics in sociotechnical systems. *Frontiers in ICT*, *3*, 2.

Veal, A. J. (2006). Research methods for leisure and tourism: a practical guide. Essex: Pearson Education

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European journal of information systems*, *25*(1), 77-89.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), 425-478.

Viana, J. (2014). Reflections on two approaches to hybrid simulation in healthcare. In *Proceedings of the 2014 Winter Simulation Conference* (pp. 1585-1596). IEEE.

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639.

Vincent, C., & Amalberti, R. (2016). *Safer healthcare: strategies for the real world* (p. 157). Springer Nature.

Visser, W. (2010). Schön: Design as a reflective practice. Collection, Parsons Paris School of art and design, 2010, *Art + Design & Psychology*, (2), 21-25.

Vom Brocke, J., Winter, R., Hevner, A., & Maedche, A. (2020). Special Issue Editorial– Accumulation and Evolution of Design Knowledge in Design Science Research: A Journey Through Time and Space. *Journal of the Association for Information Systems*, *21*(3), 9.

Wagner, G., Seck, M., & McKenzie, F. (2016). Process modeling for simulation: observations and open issues. In E. T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick (Ed.), *Proceedings of the 2016 Winter Simulation Conference* (pp. 1072–1083).

Wagstaff, K. L. (2012). “Machine learning that matters,” in *Proceedings of the 29th International Conference on Machine Learning (ICML 2012)*, 529–536.

Waks, L. J. (2001). Donald Schon's philosophy of design and design education. *International Journal of Technology and Design Education*, *11*(1), 37-51.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wan, J., Zou, C., Ullah, S., Lai, C.-F., Zhou, M., & Wang, X. (2013). Cloud-enabled wireless body area networks for pervasive healthcare. *IEEE Network Magazine*, *27*(5), 56–61.

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70*, 287–299.

Wang, Y., Kung, L., Gupta, S., & Ozdemir, S. (2019). Leveraging big data analytics to improve quality of care in healthcare organizations: A configurational perspective. *British Journal of Management*, *30*(2), 362-388.

Wang, J., Zuo, W., Rhode-barbarigos, L., Lu, X., Wang, J., & Lin, Y. (2020). Literature review on modeling and simulation of energy infrastructures from a resilience perspective. *Reliability Engineering and System Safety*, *183*: 360–373.

Waterson, P., Robertson, M. M., Cooke, N. J., Militello, L., Roth, E., & Stanton, N. A. (2015). Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety. *Ergonomics*, *58*(4), 565-599.

Weber, E. J., Hirst, E., & Marsh, M. (2017). The patient’s dilemma: attending the emergency department with a minor illness. *BMJ (Clinical Research Ed.)*, *357*, j1941.

Weick K. (1995) Sensemaking in Organizations. Thousand Oaks, CA: Sage

Weick, K. E., Sutcliffe, K. M., Obstfeld, D., Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2016). *Organization Science inf 3M1 Organizing and the Process of Sensemaking*. *16*(4), 409–421.

Weinick, R. M., Burns, R. M., & Mehrotra, A. (2010). Many emergency department visits could be managed at urgent care centers and retail clinics. *Health affairs*, *29*(9), 1630-1636.

Weiner, J., Tanniru, M., Khuntia, J., Bobryk, D., Naik, M., & Lepage, K. (2016). *Digital leadership in action in a hospital through a real time dashboard system implementation and experience*. *5*(4).

Weyer, S., Meyer, T., Ohmer, M., Dominic, S., Meyer, T., Ohmer, M., … Gorecky, D. (2016). Future Modeling and Simulation of CPS-based Factories : Future Factories : Future Factories : Example from the Automotive Industry Future an Factories : an Example from the Automotive Example from the Automotive an Example from the Automotive Industry. *IFAC PapersOnLine*, *49*(31), 97–102.

Wezel, V., & Winterfeldt, V. (2016). *Behavioural operational research : Returning to the roots of the OR profession*. *249*, 791–795.

White, L. (2009). Understanding problem structuring methods interventions. *European Journal of Operational Research*, *199*(3), 823–833.

White, L., Burger, K., & Yearworth, M. (2016). Smart cities: big data and behavioral operational research. In *Behavioral Operational Research* (pp. 303-318). Palgrave Macmillan, London.

Williams, A., O'Rourke, P., & Keogh, S. (2009). Making choices: Why parents present to the emergency department for non-urgent care. *Archives of Disease in Childhood*, *94*(10), 817–820.

Wong, B. W., & Blandford, A. E. (2002). Analysing ambulance dispatcher decision making: Trailing emergent themes analysis. In *Proceedings of the HF2002 Human Factors Conference: Design for the whole person: integrating physical, cognitive and social aspects*. Ergonomics Society of Australia.

Woods, D. D., & Branlat, M. (2011). Basic patterns in how adaptive systems fail. *Resilience engineering in practice*, *2*, 1-21.

Wong, J., M. M. Horwitz, L. Zhou, and S. Toh. 2018. "Using Machine Learning to Identify Health Outcomes from Electronic Health Record Data". *Current Epidemiology Reports* *5*(4):331-342.

Wright, M. C., & Endsley, M. R. (2017). Building shared situation awareness in healthcare settings. In *Improving healthcare team communication* (pp. 97-114). CRC Press.

Wu, W., Gao, J., Yuan, Y., Zhao, H., & Chang, K. (2011). Water distribution network real-time simulation based on SCADA system using OPC communication. In *2011 International Conference on Networking, Sensing and Control* (pp. 329-334). IEEE.

Xu, K., & Chan, C. W. (2016). Using future information to reduce waiting times in the emergency department. *Manufacturing & Service Operations Management*, *18*(3), 314–331.

Yin, R. K. (1994). Discovering the future of the case study. Method in evaluation research. *Evaluation practice*, *15*(3), 283-290.

Yin, R. K. (2006). Mixed methods research: Are the methods genuinely integrated or merely parallel. *Research in the Schools*, *13*(1), 41-47.

Yin, R. K. (2014). Getting started: How to know whether and when to use the case study as a research method. *Case study research: design and methods*, *5*, 2-25.

Yin, R. K. (2015). Causality, generalizability, and the future of mixed methods research. In *The Oxford handbook of multimethod and mixed methods research inquiry*.

Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage publications.

Young, T., Brailsford, S., Connell, C., Davies, R., Harper, P., & Klein, J. H. (2004). Using industrial processes to improve patient care. *British Medical Journal*, *328*(7432), 162–164.

Young, T., Eatock, J., Jahangirian, M., Naseer, A., & Lilford, R. (2009). Three critical challenges for modeling and simulation in healthcare. In *Proceedings of the 2009 Winter Simulation Conference*, (pp.1823–1830). IEEE.

Yousefi, M., Yousefi, M., Fathi, M., & Fogliatto, F. S. (2019). Patient visit forecasting in an emergency department using a deep neural network approach. *Kybernetes*.

Zachariadis, M., Scott, S. & Barrett, M. (2013). Methodological Implications of Critical Realism for Mixed-Methods Research. *MIS Quarterly*. 37(3): 855-879.

Zajic, A. (2019). Introduction to AIC – Akaike Information Criterion. Available at: https://towardsdatascience.com/introduction-to-aic-akaike-information-criterion-9c9ba1c96ced Accessed: 03.04.20

Zhang, F., Rio, M., Allais, R., Zwolinski, P., Carrillo, T. R., Roucoules, L., ... & Buclet, N. (2013). Toward a systemic navigation framework to integrate sustainable development into the company. *Journal of cleaner production*, *54*, 199-214.

Zhang, X., Liang, Y., & Liu, W. (2017). Pricing model for the charging of electric vehicles based on system dynamics in Beijing. *Energy*, *119*, 218-234.

Zhang, X., Lhachimi, S. K., & Rogowski, W. H. (2020). Reporting Quality of Discrete Event Simulations in Healthcare—Results from a Generic Reporting Checklist. *Value in Health*, 1–9.

Zhao, Y., Peng, Q., Strome, T., Weldon, E., Zhang, M., & Chochinov, A. (2015). Bottleneck detection for improvement of Emergency Department efficiency. *Process Management Journal*, *21*(3), 564–585.

Zhou, Z., Wang, Y., & Li, L. (2014). Process Mining Based Modeling and Analysis of W orkflows in Clinical Care - A Case Study in a Chicago Outpatient Clinic. *Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control*, 590–595.

Zhou, L., Zhao, P., Wu, D., Cheng, C., & Huang, H. (2018). Time series model for forecasting the number of new admission inpatients. *BMC Medical Informatics and Decision Making*, *18*(1).

Zsambok, C. E., & Klein, G. (Eds.). (2014). *Naturalistic decision making*. Psychology Press.

Zulkepli, J., & Eldabi, T. (2015). Towards a framework for conceptual model hybridization in healthcare. In *2015 Winter Simulation Conference (WSC)* (pp. 1597-1608). IEEE.