Cardiff University

School of Mathematics

Modelling trauma hip fracture hospital activities

By

Cheryl Voake

B.Sc. Mathematics, Operational Research and Statistics, Cardiff University

A thesis presented for the degree of Doctor of Philosophy

August 2012

SUMMARY

Hip fracture is the most common reason for an elderly person to be admitted to an acute orthopaedic ward. The main aim of this research is to provide a statistical evaluation of a hip fracture database, and then to use Operational Research (OR) techniques, using the statistical output, to model activities associated with the care of hip fracture patients. OR techniques employed in this thesis include simulation and queuing theory.

This research focuses on hip fracture admissions to the University Hospital of Wales in Cardiff, with a primary aim of ascertaining whether the time between admission and surgical intervention has any impact upon patient outcome. Outcome is considered in terms of mortality, hospital length of stay and discharge destination.

Statistical analyses are performed, via regression and CART analysis, to investigate length of stay and mortality variables. The results from these statistical tests are compiled, compared and investigated in more depth. Additionally, a principal component analysis is performed to investigate whether it would be feasible to reduce the dimensionality of the dataset, and subsequently principal component regression methodology is used to complement the output.

Simulation is used to model activities in both the hip fracture ward and the trauma theatre. These models incorporate output from the statistical analysis and encompass complexities within the patient group and theatre process. The models are then used to test a number of 'what-if' type scenarios, including the future anticipated increase in demand.

Finally, results from queuing theory are applied to the trauma theatre in order to determine a desired daily theatre allocation for these patients. Specifically, the M | G | 1 queuing system and results from queues with vacations are utilised.

The thesis concludes with some discussion of how this research could be further expanded. In particular, two areas are considered; risk scoring systems and the Fenton-Wilkinson approximation.

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …………………………………   Date ……………………

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed …………………………………   Date ……………………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …………………………………   Date ……………………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …………………………………   Date ……………………

ACKNOWLEDGEMENTS

PUBLICATIONS AND PRESENTATIONS

Publications

Griffiths J, Turley A, Voake C and Williams J (2012) Modelling the effect of operative delay on a trauma hip fracture ward at a large teaching hospital. *GSTF International Journal of Mathematics, Statistics and Operations Research* 1 (1) 61-66. doi: 10.5176/2251-3388_1.1.12

Baboolal K, Griffiths J, Knight V, Nelson A, Voake C and Williams J (2012) How efficient can an emergency unit be? A perfect world model. *Emergency Medicine Journal* (to appear, published Online First February 2012). doi: 10.1136/emermed-2011-200101.

Voake C (2010) A National Taught Course Centre in Operational Research (NATCOR): a student's perspective. *MSOR Connections* 10 (1) 19-20.

Voake C (2009) Attending OR50. *Inside OR* January 2009 15.

Griffiths J, Turley A, Voake C and Williams J. A simulation model of a trauma operating theatre – how can we reduce cancellations? [in draft format].

Conference contributions

Mathematics Teacher Fellow and student shadowing project – Wilson R, Powell P and Voake C.

National HE STEM Programme Conference, University of Birmingham, September 2012.

Modelling trauma hip fracture patients: the impact on operating theatres and the orthopaedic ward – Voake C, Griffiths J and Williams J.

Operational Research Applied to Health Services (ORAHS); Genoa, Italy; July 2010.

Modelling the spread of MRSA around a hospital ward – Voake C.

Speaking of Science; Cardiff University, May 2010.

Statistics workshop: An introduction to some techniques available to Operational Researchers – Voake C.

Student Conference in Operational Research (SCOR); University of Nottingham; April 2010.

Modelling orthopaedic patient hospital activities – Voake C, Gedgaudaite R, Griffiths J, Williams J, Turley A and Johansen A.

Velindre NHS Trust Research & Development Conference; Cardiff; November 2009.

Modelling trauma hip fracture hospital activities – Voake C, Griffiths J and Williams J.

EUROXXIII; Bonn, Germany; July 2009.

SCOR; Lancaster University; March 2009.

LANCS PhD Symposium; Cardiff; January 2009.

OR50; York; September 2008.

Investigating trauma hip fracture hospital activities – Voake C, Griffiths J and Williams J.

Physical Sciences & Engineering Graduate School Annual Lecture; Cardiff University; November 2009 [poster].

ORAHS; Leuven, Belgium; July 2009.

Health Solutions Wales and Cardiff School of Mathematics joint seminar; Cardiff School of Mathematics; July 2009.

An investigation into transmission models for MRSA – Voake C, Griffiths J and Williams J.

ORAHS; Toronto, Canada; August 2008.

Awards

Best Talk – *Speaking of Science*; Cardiff University, May 2010.

Best Newcomer Oral Presentation – *Velindre NHS Trust Research & Development Conference*; Cardiff; November 2009.

Best Poster – *Physical Sciences & Engineering Graduate School Annual Lecture*; Cardiff University; November 2009.

CONTENTS

CHAPTER 9: A THEORETICAL APPROACH TO MODELLING THE TRAUMA THEATRE 272

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1 Trauma fracture neck of femur

1.1.1 Overview

The most common reason for an elderly person to be admitted to an acute orthopaedic ward is because of fractured neck of femur (NoF), also known as hip fracture. The average age of patients who suffer from a hip fracture is over 80, and 80% of these are women. The incidence of hip fractures in the United Kingdom (UK) is approximately 86,000 per year, and 95% of these are the result of a fall (Parker and Johansen 2006).

Approximately 30% of people aged over 65 years and living in the community fall each year, increasing to 50% of people aged 80 years and over (RCN 2004). A fifth of all fall incidents require medical attention (Gillespie et al. 2003). These numbers are around three times higher amongst those living in institutions (Parker and Johansen 2006). Incidence rates within institutions are also increased in the first months after admission to a nursing home and with increasing age (Rapp et al. 2008).

Due to the world's ageing population, most areas are seeing a 1-3% increase in the number of hip fractures each year, but this varies widely by region (Cummings and Melton III 2002). The worldwide prevalence of hip fracture was estimated as 1.3m in 1990; this was estimated to double by 2025 to 2.6m, with a greater percentage increase seen in men compared with women (Gullberg et al. 1997). By 2050, there are expected to 6.3m hip fractures annually across the globe. Shifts have also been seen over time in average patient age and type and locality of fracture (Kannus et al. 1996), but further research is required into the effects of other determinants, such as socioeconomic status for example, on changes in and impact upon hip fracture prevalence (Marks 2010).

The incidence of hip fracture in the UK is expected to be approximately 101,000 by 2020 (BOA 2007). The increase can be largely explained by the ageing population; a 28% increase is estimated in the over 50s population between 2004 and 2031, with the largest percentage increases seen in the over 80s population. By 2031, 45% of fractures will occur in patients aged 85 years or greater, an increase from 34% in 2004 (Holt et al. 2009).

The substantial and increasing burden that this injury has on healthcare systems worldwide is evident. Due to the great expense and sometimes inefficient rehabilitation of treating these patients, the best way to relieve financial and social pressures of hip fracture is suggested to be via effective preventative medicine (Melton III 1993).

The group of patients of interest here are those who incur a trauma hip fracture. While many patients undergo treatment for hip fracture electively and surgery may be booked weeks or even months in advance, here the concern is with those patients who are admitted as emergency cases. Trauma cases require more urgent medical treatment than elective patients.

1.1.2 Costs

Inpatient costs account for 50% of the total cost of a fall related injury. Additional costs are accrued through outpatient appointments, rehabilitation, loss of earnings to carers and general practice appointments (WHO 2007).

It has been estimated that hip fracture patients occupy one in five orthopaedic beds in England and Wales (Lindsay 1995) and account for more than two million hospital bed days per year in England alone (DoH 2004). The cost of hip fractures to the National Health Service (NHS) and social care services is estimated to be around £1.73 billion per year (Torgerson et al. 2001), while the charity Age UK has warned that falls among the elderly, the commonest result from which is hip fracture, may be costing the English NHS up to £4.6m per day (BBC News 2010).

Reported costs per case are variable and are substantially higher than those reported for other injuries such as vertebral or wrist fractures (Dolan and Torgerson 1998). Variation in costs and cost-effectiveness is dependent upon type and efficacy of treatment, fracture risk, patient age and patient compliance (Johnell 1997).

In 2005, a team in Nottingham reported that the average hospital expenditure per hip fracture patient was £12,163 (Lawrence et al. 2005). 84% of this cost was attributable to ward costs, where the mean length of stay was 23 days. Surgical costs accounted for 9% of the total costs and the remainder (7%) was due to medical investigations. Three years later a team in Dublin reported lower costs and a notably different expenditure breakdown (Azhar et al.

2008). Average cost per patient was €9,236, which equates to approximately £7,296 using a historical exchange rate from mid-2008 (X-Rates© 2012), just 60% of the cost reported by the team in Nottingham. Ward costs accounted for 55% of the total expenditure, surgical costs accounted for 40% and 5% was due to medical investigations. Mean length of stay was eleven days. Ward costs in the earlier paper are therefore higher per patient per day, £444 compared with approximately £365 according to Azhar et al, but operative costs were much lower on average, £1,095 compared with approximately £2,918.

Thakar et al compared the costs of treating patients by whether they had a medical complication (Thakar et al. 2010). Patients with a complication were shown to incur more than double the cost than those without, £18,709 compared with £8,610. This difference is largely attributable to the additional time spent in hospital for medically complicated patients, but the average cost per patient per day was still slightly greater, £298 compared with £263 for medically uncomplicated patients. Medical complications have been shown to double hospital stay for hip fracture patients, but in this case cost rates were not higher; non-medically complicated patients cost 63% of what medically complicated patients cost (Khasraghi et al. 2003).

One objective of this thesis is investigating whether surgical delay has any impact upon patient outcome and hospital stay. Surgical delay was the main focus of one cost study, where it was shown that reallocating budget to ensure that surgery is performed within 48 hours is more cost-effective than allowing surgery to be delayed for longer than this (Shabat et al. 2003).

Considerable costs are also generated post-discharge. Care for hip fracture patients is approximately three times as much as for matched uninjured controls and around 40% of this excess cost is incurred in the first three months after discharge (Haentjens et al. 2005). It is estimated that a hip fracture patient will spend approximately 17% of their life in a nursing facility, with expenses relating to these facilities accounting for 44% of the total cost of patient treatment (Braithwaite et al. 2003).

1.1.3   Osteoporosis

Osteoporosis literally means 'porous bones', which results from the rate of bone renewal not matching the rate of breakdown, eventually leading to weak and brittle bones, see Figure D1.1.3a of Appendix D (BoneMatters® 2011). This weakening means that should an osteoporosis sufferer experience a fall, their bones are more likely to break.

Many sufferers do not realise they have this condition until they break a bone (Glenville 2005), despite the possibility of identifying the existence of osteoporosis via a heel quantitative ultrasound scan, which has shown to be a reliable indicator of both bone density and hip fracture risk (Hans et al. 1996, Khaw et al. 2004), while bone mineral density itself has also been shown to be of substantial importance in predicting the risk of hip fracture (Johnell et al. 2005). The cost of treating fractures related to osteoporosis in Britain is estimated to rise to £2.1bn by 2020, an increase from £750m in 1995 (Gorman 1996).

The lifetime risk of any osteoporotic fracture is high; 40-50% for women and 13-22% for men (Johnell and Kanis 2005), while it has been stated that in the UK, after the age of 50, one in two-three women and one in five-12 men will sustain a fracture, the majority of which will be due to osteoporosis (NOS 2007). These are more conservative figures than previously given, where it was estimated that half of all women and one in five men will suffer a fracture due to osteoporosis after the age of 50 (van Staa et al. 2001).

The World Health Organisation suggests that it is crucial to contain the effects of osteoporosis via health promotion and preventative measures, and identifying the risk factors of this condition are vital in order to do this (WHO 2008). Reports indicate that risk factors for osteoporosis include, but are not limited to; low body weight, female sex, Caucasian ethnic origin, sedentary individuals, dementia, high alcohol consumption, diseases of malabsorption (Crohn's disease, for example), history of falls and older age (Cluett 2010, Hannan et al. 2000, Pugh 2011, Siris et al. 2001).

Paget's disease is a medical condition which can also cause weakening of the bone. Sufferers of this condition are at a three times greater risk of requiring a total hip replacement than non-sufferers (van Staa et al. 2002).

1.2   Locality

Cardiff and Vale University Health Board (UHB) is the operational name of Cardiff and Vale University Local Health Board (LHB). It is one of the largest NHS organisations in the UK and provides day-to-day services to the populations in the two regions of Cardiff and The Vale of Glamorgan, a total of approximately 445,000 people. Other responsibilities include the delivery of NHS primary care services across this area as well as some services in some specialties across the wider population of Mid and South Wales. Cardiff and Vale UHB manages nine hospitals and 17 health centres in total; the focus of this thesis is the largest of these hospitals, the University Hospital of Wales, where the Board headquarters are based. In 2009/10, the approximate income for the UHB was estimated at £940m (NHS Wales 2010).

Figure 1.2i shows the proportion of Wales that is covered by the regions of Cardiff and The Vale of Glamorgan. It was estimated by the Welsh Assembly Government (WAG) that the population of Wales stood at approximately 3m people in mid 2009, 60% of whom were of working age (WAG 2010a), while the total Welsh population accounted for 1 in 20 of the total UK population (WAG 2009b). In 2008 there were an estimated 1.3m households countrywide (WAG 2010g).
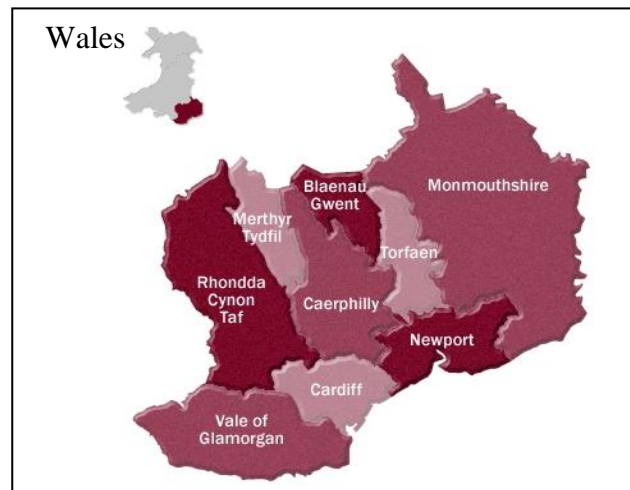
Figure 1.2i: Map of Wales, with detailed local administrative regions in South-East Wales

It can be seen that these two regions have a high population density in comparison with the rest of the country. Indeed, there are a total of 22 unitary authorities in Wales, and while the two authorities of Cardiff and The Vale of Glamorgan cover just 2.3% of the land space, they accommodate around one sixth of the Welsh population. It is additionally interesting to note

that Wales has a higher proportion of people aged 60 years and over than the rest of the UK and also a higher proportion of the population with a limiting long-term illness (ONS 2010).

The University Hospital of Wales (UHW) is a large teaching hospital, indeed it is the largest hospital in Wales and the third largest university hospital in the UK. It is located in Cardiff, the capital city of Wales. According to figures published by the Welsh Assembly Government, in 2008/09 there was an average of 1042.2 daily beds available at the UHW with an average of 83.6% of these beds occupied (WAG 2010f).

Any person sustaining a trauma hip fracture while under the care of Cardiff and Vale UHB will be admitted to the UHW and cared for by the Hip Fracture Service (HFS). This consists of a Clinical Nurse Specialist (CNS) for elderly trauma rehabilitation, two hip fracture nurse specialists and a consultant orthogeriatrician and his medical team, with the aim of striving to improve the holistic care provided to patients admitted with a hip fracture. The HFS was developed to improve the care and outcome of these patients, which it has achieved by developing effective pathways to improve management pre-, peri- and post-operation.

All patients arrive from Accident & Emergency (A&E); once a hip fracture is diagnosed there then the patient is transferred to a trauma ward, where day admissions are clerked by a Trauma House Officer and night admissions by a Senior House Officer. Some patients have to wait in the Emergency Unit until a bed is available. UHB guidelines state that patients should be seen within two hours and bedded within four, but commonly these are not met.

At the UHW, there is no dedicated theatre for these patients, nor is there any dedicated theatre time. (Elective surgery takes place elsewhere.) Each evening, the hip fracture team will nominate patients who are fit for surgery that they would like to send to theatre the following day; usually two patients per day are nominated. However, depending on the demand for surgical treatment from elsewhere in the hospital, patients may or not be scheduled for theatre the following day. Patients with a hip fracture are operated on in the same theatre as other emergencies arriving from A&E, such as road traffic accidents, so often even scheduled patients will not be seen. A member of the team will typically contact the theatres at around 3:30pm to check on the list to gain an idea of whether any nominated patients who are still waiting are likely to go to theatre that day; there is usually no advanced warning that a patient is going to theatre until a porter arrives to collect them.

1.3   Objectives and outline of this thesis

Recall that trauma hip fracture is a common injury in the elderly population. It is an unfortunate reality that, due to a number of reasons, patients do not often undergo surgery promptly. This thesis proposes several analyses which address trauma hip fracture patient flow and predicts improvements based on varying parameters.

The main aim of this research is to employ statistical and Operational Research techniques in order to fully analyse and explore the data available relating to trauma hip fracture patients at the UHW. By doing this, the systems can be fully scrutinised in order to identify areas and methods of improvement. Specifically, the following objectives can be identified:

• Objective 1: Understand the factors, if any, which affect length of stay and outcome for trauma hip fracture patients at the UHW. In particular, investigate whether a delay between admission and surgery has any impact on these variables.

• Objective 2: Use the insight gained from Objective 1 to build a simulation model to represent the hip fracture ward, and investigate the changes that would occur should any alterations be made to the system or patient management.

• Objective 3: Build a separate simulation model of the trauma theatre in order to investigate different management policies on key outputs of the theatre, showing in particular how to reduce cancellations.

• Objective 4: Identify and use appropriate queuing systems to model trauma theatre activities.

Objective 1 is addressed in Chapters 2 to 5, where a database of patients is analysed in considerable detail using a variety of statistical techniques. In particular, length of stay and mortality are investigated in Chapters 3 and 4 respectively, using CART and regression. Factors found to be influential are discussed and analysed in greater depth. In Chapter 5, principal components analysis is used in an attempt to collapse the dimensionality of the dataset.

Chapter 6 presents two discrete event simulation models, built in Visual Basic for Applications (VBA), of the hip fracture ward, and thus addresses the second objective. Results and conclusions from previous chapters are drawn upon to decide inputs to the model. A number of scenarios are tested to investigate the impact of making adjustments to the system.

Data relating to the trauma theatre is analysed in Chapter 7, where detailed results are presented. Appropriate ways to segregate patients are discussed and results are used to inform the building of an appropriate simulation model of the trauma theatre at the UHW, presented in Chapter 8 and thus addressing Objective 3. A variety of policies relating to the organisation of this operating theatre are considered in order to investigate more effective ways to manage the running of the theatre.

Objective 4 is addressed in Chapters 9 and 10. Scrutiny of the data and consideration of the arrival and service processes mean that appropriate queuing theory results can be applied to the trauma theatre, and then parameters amended to determine the impact of making changes to the system. In particular, a novel and bespoke queuing system is formulated in Chapter 10 which considers two types of arrival, two types of service and a limit on the number of patients who may join the system.

Finally, Chapter 11 provides a conclusion to the research and presents some ideas for expanding on the work presented in this thesis.

A general outline of this research has thus been presented; note that appendices are also included which provide some additional key information and results. Figures and Tables are labelled sequentially (i, ii, …) according to the section in which they feature.

An overview of all probability distributions used throughout this work is given in Appendix A, including formulae for the mean and standard deviation of these distributions. Appendix B gives more information on the variables used, while a medical glossary is given in Appendix C. Terms in the glossary were collected from a variety of sources throughout the course of this research and checked in mid-2012 (MedlinePlus 2012, Merriam-Webster 2012, NHS 2012a). The thesis concludes with Appendix D, which provides some supplementary material and results. Items in the appendices are referenced throughout this thesis where necessary. Figures in the appendices are labelled firstly with the letter indicating to which

appendix they belong, followed by the section to which they relate and then sequentially alphabetically.

1.3.1   Data

A variety of data sources were made available for this work. Specifically, separate databases were provided by each of the Hip Fracture Service and the theatres team.

Data from the HFS was provided on two occasions, some time apart; these two databases had some disparities due to a change in the data recording system for trauma hip fracture patients (see Section 6.3.1).   These databases store information at a patient level, including demographic information, pre-fracture status, date and type of surgery and discharge destination.

Theatre data was also provided on two occasions and again considerable time had elapsed between these instances.   The first database provided exclusively included hip fracture patients, while the second included all patients operated on in the trauma theatre at the UHW. The databases were similar in that they allowed for a patient's journey through the surgical process to be mapped (see Section 7.2).

Consequentially the data available for certain analyses changed throughout this research and unfortunately meant that consistency was not always possible.   This is explained and discussed in more detail where relevant.

1.4   Literature review

A broad overview of several relevant topics covered in this thesis is now given. Other specific topics are introduced and referenced where necessary throughout the course of this research. A more detailed literature review is also given on some of the following topics where they arise in this thesis, where necessary.

1.4.1   Treatment

After arrival in hospital, fast-tracking patients through A&E is beneficial not only to patients, but also to A&E, ward and orthopaedic staff, and achieving this is a key first phase of the treatment of hip fracture patients (Ryan et al. 1996).

The majority of trauma hip fracture patients are treated surgically. Deciding between a surgical or conservative approach is the initial phase of any treatment plan; however, due to prolonged hospital stay and inferior rehabilitation, conservative treatment is now rarely used (Parker and Johansen 2006). Where surgical treatment is unavailable or inappropriate, non-surgical treatment such as analgesics is prescribed. While this will evidently remove the risk of surgical complications, rehabilitation is likely to be slower and limb deformity is more common (Handoll and Parker 2008). Despite this, no differences in mobility, mortality or residence have been shown between patients treated surgically or non-surgically after hip fracture by one study (Hossain et al. 2009).

The choice of operation is dependent upon a number of factors and can be partially dependent upon clinician subjectivity. Hip fractures may be fixed via internal implants or replaced via arthroplasties; many of the implants and arthroplasties currently in use have been around for over 50 years but improvements in surgical technique have led to fewer complications and reoperations (Parker and Gurusamy 2005). The choice of surgical procedure will be influenced by the type and location of fracture. Fractures may be intracapsular or extracapsular, see Figure D1.4.1a for a pictorial classification of fracture type (Parker and Johansen 2006). They can be further classified as displaced or undisplaced; whether or not the bone has moved from its usual place. Fracture type and choice of surgical implant have been reported to have no impact upon patient outcome in one recent systematic review paper (Butler et al. 2011).

Two types of hip operation using replacement are total hip replacement (THR) and hemiarthroplasty; the hemiarthroplasty replaces just the ball portion of the hip joint, while with a total hip replacement the socket is also replaced. These types of operation are typically used to treat intracapsular fractures; see Figure D1.4.1b for radiographs of a fitted prosthesis after hemiarthroplasty surgery (Parker and Johansen 2006), and total hip replacement surgery (ONSMD.com 2012). Given these two options, most orthopaedic surgeons advocate hemiarthroplasty even though good, and sometimes better, results are achieved for THR (Blomfeldt et al. 2007). Despite longer surgery duration, THR has been shown to have better results with regard to hip function and health-related quality of life. Hip function in this case was measured by the Harris score (Harris 1969), a popular means of evaluating hip function post-surgery. Better short-term clinical results and fewer complications have also been reported for THR, when compared with hemiarthroplasty (Baker et al. 2006). Various types of screws, plates and nails are used to fix extracapsular fractures, also see Figure D1.4.1b for a radiograph of a fitted intramedullary nail prosthesis, an option which is increasingly being used for this fracture type (Parker and Johansen 2006).

An operative and supervision algorithm, the Hvidovre algorithm, was created by a Danish team which specified treatment choice for hip fracture patients (Palm et al. 2012). The choice of surgical procedure had to follow the algorithm post-implementation and was based solely on fracture type, patient age and whether the patient was bedridden pre-fracture. Clinician subjectivity was therefore removed. A decline in the number of required reoperations was seen after implementation; this held true for junior surgeons operating with or without supervision. It was estimated that extra bed days consumed by reoperations was reduced from 24% to 18% of all bed days.

Clearly anaesthesia will be required for an operation as invasive as hip surgery. A review of 15 research articles came to the conclusion that the use of regional anaesthesia was marginally advantageous to general anaesthesia, in terms of reducing early mortality and the risk of deep venous thrombosis (Urwin et al. 2000). A later review of 56 articles on anaesthesia for geriatric hip fracture patients also concluded that spinal anaesthesia is better, stating that it holds a number of advantages over general anaesthesia, including lower early mortality rates, less post-operative confusion and fewer cases of pneumonia. However, it was also recommended that more research is required in this area (particularly with respect to mortality) and that the method of anaesthesia should be based on several factors, including

patient preference and the clinical experience of the anaesthesiologist (Luger et al. 2010). Another study concluded that differences in clinical outcomes was unclear between the two techniques, but showed that despite taking slightly longer to administer, spinal anaesthesia was significantly cheaper per patient than general anaesthesia; £194 compared with £271 respectively (Chakladar and White 2010).

A 2005 review paper of best practices for the care of elderly hip fractures found that, in addition to spinal anaesthesia as stated previously, the use of peri-operative antibiotics and pressure-relieving mattresses were consistently beneficial (Beaupre et al. 2005).

Lean thinking techniques were applied to a large hospital in Birmingham in an attempt to improve outcome following hip fracture (Yousri et al. 2011). Lean thinking in healthcare is "about getting the right things to the right place, at the right time, in the right quantities, while minimising waste and being flexible and open to change" (NHS 2012b). A significant reduction in mortality was observed post-implementation, while improvements were also made in surgical delay, trauma bed usage and early hospital discharge.

An integrated care pathway (ICP) is "a document that describes a process within Health and Social Care", while the purpose is to put patients at the centre of care. It has a similar remit to lean thinking in that it aims to have the right people, in the right order, in the right place, with these people doing the right thing, in the right time, with the right outcomes. An ICP is a multidisciplinary best practice outline of anticipated care, which is evidence-based and reflects a patient-centred approach (NLIAH 2005).

ICPs are developed locally, but it has been recommended that the development of a national validated ICP for hip fractures may be important in order to avoid unnecessary local deviations from national guidelines (Smith et al. 2008). However, there is controversy surrounding the effectiveness of ICPs for hip fracture patient management (Parker 2004) and the use of ICPs for treating hip fracture patients is an area which requires further research (SIGN 2002). Some reports focus on the effects of a multidisciplinary approach to treating these patients, thus while ICPs may not be referenced specifically, the treatment methodologies are largely comparable.

An ICP implemented at a hospital in Yorkshire concluded that this approach has potential benefits. Length of stay was significantly reduced and, while statistical significance was not

reached, improvements in surgical delay and one month mortality rates were also seen (Gholve et al. 2005).

Shorter length of stay was also reported by Choong et al after implementation of a clinical pathway for hip fracture patients. This was achieved without increasing complication or readmission rates, though these measures were not improved upon either (Choong et al. 2000).

Six hospitals, two of which had clinical pathways for hip fracture, were compared across various outcomes (March et al. 2000). Length of stay was significantly reduced for nursing home patients; a reduction was also seen for other patients but results did not reach statistical significance. There was also a non-significant decrease in nursing home admission rates and no difference in mortality rates at four months.

Successful results were also reported after an ICP was implemented in Southampton. However, in this case length of stay was shown to increase, but this did lead to an improvement in clinical outcome; better ambulation on discharge and a reduction in long term care admissions were also both seen (Roberts et al. 2004).

One review concluded that multidisciplinary rehabilitation led to better outcomes, but that differences were not significant. However, since a multidisciplinary approach is not harmful, it was suggested that it is still preferable (Handoll et al. 2009). This approach is also advocated by another review article which states that surgeons cannot accept sole responsibility for these patients, but that geriatric care should encompass holistic patient management (Leung et al. 2010a).

The British Geriatrics Society (BGS) believes that joint care between geriatricians and orthopaedic surgeons delivers the best patient care amongst a list of orthogeriatric care models. Care on a dedicated orthogeriatric ward is also advised (Aylett et al. 2007). This is consistent with the National Service Framework for Older People which also recommends that hospitals should have at least one ward developed as a centre of excellence for the care of older people with fractures (DoH 2001).

One study compared three outcomes (mortality, length of stay and discharge destination) between two patient groups, one of which was managed jointly between a consultant geriatrician and orthopaedic surgeons, while for the other there was no geriatrician. No

differences in any of the three measures were found between the two groups and it was thus concluded that combined orthogeriatric care had no impact on patient outcome (Khan et al. 2002). A contrasting study compared treatment via consultation by geriatricians with joint care provided by geriatricians and orthopaedic surgeons (González-Montalvo et al. 2010). The latter group had earlier assessment, earlier surgery and a reduced acute and total hospital stay. This was achieved without compromising clinical or functional outcome.

Through collaboration with endocrinologists, the orthopaedic service at one hospital found that adding vitamin supplements and an endocrinology appointment effectively improved treatment of hip fracture patients. Patient compliance was also increased (Piziak and Rajab 2011).

A reduction in post-operative morbidity, specifically for post-operative heart failure, cardiac arrhythmias and delirium, was found after the implementation of a care pathway for hip fracture patients (Beaupre et al. 2006). Importantly, this was done without any negative impact on resources. Hospital length of stay increased for patients following the care pathway, but rehabilitation length of stay decreased; overall impact was no differences in length of stay dependent upon whether a patient followed the pathway. In-hospital mortality also remained unaffected.

The Sheba model is based upon the concept that a hip fracture represents a geriatric disease and not an orthopaedic disease, and is implemented via treatment through a comprehensive orthogeriatric unit, which covers all aspects of care for hip fracture patients. Evidence shows that applying this model results in short length of stay, acceptable functional outcome and low mortality and morbidity rates (Adunsky et al. 2005).

A useful review of best practice management for hip fracture patients was published by Bruyere et al, but concluded further study is needed. Proper nutrition was highlighted as a key area in which care should be focussed, while overall appropriate management of these patients can prevent, or at least minimise the risk of, further fractures and health deterioration (Bruyere et al. 2008). Other areas emphasised as key factors for hip fracture patient care include urinary tract management and the prevention of deep venous thrombosis (Huddleston and Whitford 2001).

Finally, the benefits of using mapping methods to model hip fracture care have been shown. Mapping care pathways, via drawing on a variety of information sources, is the first step in planning for future health services and system improvements (Vasilakis et al. 2008).

1.4.2   Outcomes

It is estimated that 20% of people who fracture their hip die within one year (Cummings and Melton III 2002). There were no changes in UK six month or one year mortality rates over a time period of 40 years (1959-1998) (Haleem et al. 2008), although an increase over time in mortality following hip fracture has been reported elsewhere (Vestergaard et al. 2007).

Sustaining this injury can be detrimental to the subsequent life of the sufferer for those who do survive. Half of survivors can no longer live independently and a quarter are no longer able to prepare their own meals, while almost half of patients who could previously walk unaided are no longer able to do so (Osnes et al. 2004).

Another study estimated that 25-50% of those who survive a hip fracture regain their previous level of functionality and ability to perform activities of daily living (ADLs) (Isenberg et al. 2004). The index of independence of activities of daily living measures adequacy of performing six functions: bathing, dressing, toileting, transferring, continence, and feeding (Katz and Stroud 1989), so any reduction in the ability to perform these tasks would be considerably detrimental to quality of life. It is also likely that elderly women will continue to suffer from a loss in quality of life and experience substantial functional impairment, even if they show significant signs of recovery in the first year post-fracture and after adjusting for age and comorbidities (Boonen et al. 2004).

The change in quality of life post-fracture has been shown to be dependent on the type of fracture, with displaced fractures resulting in lower quality of life than undisplaced fractures. Treatment type was also indicative of complications and reoperations, with internal fixation having significantly poorer results by these two measures compared with total hip replacements (Tidermark 2003).

Many patients who previously lived at home are discharged to a nursing home. After rehabilitation, some of them may then return to the community; the likelihood of this is

shown to depend upon the utilisation (admissions to beds ratio) of the nursing home and whether the patient achieves ambulation prior to discharge from hospital (Fitzgerald and Dittus 1990).

Comparison of outcome other than mortality is difficult between studies since there is no standardised, validated scale used to measure outcome; a review which initially looked at over 4000 papers found that those relating to ADLs proved to be the most popular (Hutchings et al. 2011).

Time to ambulation post-surgery was measured in a group of hip fracture patients to assess whether it is related to, or impacts upon, a range of medical and demographical variables. No relation was found between time to ambulation and a variety of other factors, including sex, age and, interestingly, the functional status of the patient prior to admission. A longer hospital stay, however, was related to a longer time between surgery and ambulation (Kamel et al. 2003).

Six functional independence measures were used to measure recovery in older hip fracture patients to assess the influence of impaired cognition on long-term care requirements. It was found that cognitively impaired patients scored worse across all six measures and required more assistance than those not impaired, leading to the suggestion that planning the long-term care of these patients is required to impede or prevent admission to a nursing home (Young et al. 2011).

The association between depression, apathy and cognitive impairment with functional improvement for hip fracture patients in two discharge destination was examined, namely inpatient rehabilitation facilities (IRFs) and skilled nursing facilities (SNFs). It was concluded that patients suffering from one of the previously mentioned conditions had significantly better outcomes (as measured by functional improvement) if treated at an IRF compared to a SNF (Lenze et al. 2007).

A review paper investigated the merits of pre-operative education for patients undergoing hip or knee replacement surgery, but found that while there may be a beneficial impact on patient anxiety pre-operation, albeit modest, there was little evidence to support the use of education to improve patient outcome. Patients educated via written information, watching a video or discussion with a healthcare professional did not show an improvement with respect to length

of stay, pain or functioning, compared with those receiving no education (McDonald et al. 2004).

Post-operative delirium was the main focus of one study of hip fracture patients aged 65 or over. Predictors of the development of delirium post-surgery were male sex, surgery under general anaesthesia and a history of mild dementia, while the effects included longer length of stay and higher mortality at one year (Edelstein et al. 2004)

1.4.3   Infections

At the UHW, all patients over the age of 65 are screened for the deadly infection Methicillin-resistant *Staphylococcus aureus* (MRSA) on arrival at the ward and are bedded in an isolated cubicle if a positive test result is received. Risk factors of contracting MRSA while in hospital are carrying MRSA at admission (one can be a carrier of MRSA without being infected), increasing age and, interestingly, hip fracture (Shukla et al. 2009). A study by a team from the UHW found that rates of MRSA colonisation were higher in patients admitted from a nursing home (17.4%) than those admitted from their own home (3.6%) (Thyagarajan et al. 2009). The additional cost of treating a patient who contracts MRSA while on an orthopaedic trauma ward was shown to be £13,972; this is for additional medicines, therapy and a considerably longer length of stay (50 extra days). The cost of preventing an infection is much lower at £3,200 (Nixon et al. 2006). It has been suggested that due to the high prevalence of MRSA colonisation on orthopaedic wards, all patients should be screened for MRSA on arrival (Walley et al. 2009). However, the majority of hip fracture patients admitted to the UHW are over the age of 65 (see Section 2.2.1) and thus almost all patients are currently screened.

Other infections are common in this vulnerable group of patients and can have a devastating effect on outcome. For example, mortality is significantly higher in hip fracture patients who are infected with *Clostridium difficile* (Gulihar et al. 2009). Surgical site infection (SSI) is another problem, while the risk of SSI has been shown to be significantly greater for reoperations than first operation, and there is great variation in rates between hospitals (Wilson et al. 2008). MRSA was found to be the commonest pathogen which caused SSI, particularly for hemiarthroplasty patients (Ridgeway et al. 2005).

1.4.4   Surgical delay

Many trauma hip fracture patients do not undergo surgery promptly. A delay to operation may sometimes be caused by medical reasons, but it is an unfortunate actuality that it is often due to system limitations, such as space, staffing, equipment and other resource constraints. An important focus of this thesis is the impact of delay upon patient outcome, focussing on in-hospital mortality and hospital length of stay in particular, to assess whether or not it has any deleterious effects.

The impact of a delay to operation for hip fracture patients is very well-documented in the literature, thus a complete overview is infeasible and so is not given here or elsewhere in this thesis. Individual key published results are reported later (see Sections 3.4.2 and 4.6.7 in particular) while some conclusions drawn from some useful review papers are included next.

(a) Impact on length of stay and mortality

A review of 52 papers, which involved 291,413 patients, in order to assess the timing of hip fracture surgery was published in 2009 (Khan et al. 2009). Papers were rated for methodological quality using a validated checklist specifically designed to evaluate healthcare studies (Downs and Black 1998), in order to ensure that any consensus made was according to conclusions drawn in the higher quality papers. The main conclusion drawn was that early surgery, within 48 hours, is beneficial in terms of a shorter length of stay and possible benefits in relation to a reduction in complications and mortality. The authors additionally conclude that a large randomised trial is required to fully resolve the issue of the timing of surgery but suggest that the actuality of this is unlikely due to ethical issues.

Of the papers which reported a conclusion on mortality (all but three), there was an almost even split on whether delay did matter; 51% reported no effect, 45% reported a reduction in mortality for early surgery and 4% (two papers) reported an increase in mortality for early surgery. 18 papers reported on the impact of delay on medical complications, with an equal split of nine papers each concluding whether delay did or did not matter. A total of 19 papers investigated whether delay affected length of stay; 68% concluded that length of stay was reduced for early surgery, the remainder concluding no effects. Finally, just four papers reported on the likelihood of patients returning home post-injury and again there was an even

split on whether or not delay did matter. The interested reader is referred to Khan et al's paper for a full breakdown of conclusions reported by the 52 studies, many of which are reported elsewhere in this thesis.

A separate systematic review was performed one year earlier (Shiga et al. 2008), and also assessed the quality of previous studies according to the checklist developed by Downs and Black. In total, 16 studies involving 257,367 patients were identified for further scrutiny and results were pooled to calculate overall findings. The review itself was performed according to the guidelines of the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Statement (Stroup et al. 2000). The main definition of operative delay was again a wait of greater than 48 hours, but this was relaxed if other cut-off times were found. Using a cut-off of 48 hours, operative delay was shown to increase the odds of 30-day mortality by 41% and one-year mortality by 32%.

In another article, Simunovic et al surveyed a total of 66 papers in order to review evidence of the effect of surgical timing on various outcomes, including mortality, post-operative complications and length of stay in hospital (Simunovic et al. 2011). Based on a pooled estimate using five papers, it was shown that earlier surgery was associated with a 19% reduction in mortality and that the effect of a delay on mortality was seen irrespective of the delay definition (24, 48 or 72 hours). The reasons for a delay to surgery were discussed and it was concluded that there is no theoretical benefit (in terms of mortality risk) for healthy patients to wait for surgery, while in medically unfit patients the effect of a delay was unclear. The majority of the studies looked at which included investigation into the effect of delay on length of stay concluded that as delay increased, so did hospital stay. A similar paper published one year previously also quoted pooled results for the impact of earlier surgery on post-operative complications, reducing the risk of pressure sores by 52% and the risk of in-hospital pneumonia by 41% (Simunovic et al. 2010).

42 articles were identified by Leung et al in another review article, where the authors concluded that surgeons should treat patients "as soon as their bodies meet the basic anaesthetic requirements" (Leung et al. 2010b). It is highlighted that, while this will inevitably vary between patients, setting a goal of surgery within 24 hours would greatly help to provide a timely and effective treatment. Despite this, there is no definitive conclusion given on whether operative delay has any effect upon mortality, instead it is stated that the

evidence is conflicting. It is concluded, however, that evidence on the whole suggests that delay impacts negatively on morbidity, the incidence of pressure sores and length of hospital stay.

The final paper considered is not included due to its type (it is an observational study and not a review article), but due to its coverage (Bottle and Aylin 2006). Data on a total of 129,522 hip fracture admissions was collected from 151 NHS Trusts in England, covering a three year period. Two definitions of delay, more than one day and more than two days, were considered and huge variation in the proportion of delayed patients between Trusts was shown. It was additionally shown that operative delay was significantly associated with the risk of in-hospital death and that this persisted as the delay increased. A decline in an increased mortality risk was only seen after a delay of 12 days.

Clearly the definition of what constitutes "a delay" is inconsistent across studies. A thorough review of the literature has indicated, however, that the most commonly used cut-off is two days, or 48 hours where data would allow this level of precision. In the review paper by Khan et al, 14 of the 52 papers assessed used this as their primary delay definition, while it was investigated within wider definitions by some other papers, more than any other definition used. Thus for this reason, alongside advice given by clinicians involved in this project, the main definition used for a significant delay to operation is a wait longer than two days, or 48 hours where data would allow.

(b) Impact on other outcomes

The effect of delay on functional outcome and avascular necrosis was investigated for patients aged 60 years old or less. It was shown that delayed surgery, classified as a wait greater than 12 hours, was associated with an increased rate of avascular necrosis (/osteonecrosis), but that this did not lead to an adverse effect on functional outcome (Jain et al. 2002). The relationship between surgical delay and osteonecrosis has also been shown specifically for paediatric hip fracture patients (Varshney et al. 2009).

While mortality and length of stay were shown to be uninfluenced by delay, a one week delay was shown to increase the incidence of postoperative complications at a Spanish hospital.

Three-month and one-year functional recovery were also shown to be unaffected (Rodriguez-Fernandez et al. 2011).

A direct correlation between operative delay and the incidence of thromboembolism has also been shown, leading to the suggestion that all patients delayed longer than 24 hours should undergo ultrasound prior to surgery to examine for the presence of deep venous thrombosis (Smith et al. 2011).

A particular problem reported in the literature is that a delay to surgery can increase the risk of pressure sores, which has been shown in several studies (Grimes et al. 2002, Parker and Pryor 1992, Pathak et al. 1997). This additional complication may then in turn lead to the requirement of additional medical care, and thus a longer stay in hospital. This relationship has, however, shown to be insignificant elsewhere, but it was instead suggested that patients undergoing earlier treatment were less susceptible to contracting a urinary tract infection (Davis et al. 1988).

A study of elderly female patients showed that delay had a detrimental effect on progress in terms of social circumstance; that is, whether their social circumstances at three months post-surgery were similar to pre-fracture, or whether their circumstances had deteriorated (including death). This was shown to be true regardless of pre-fracture social status (Villar et al. 1986).

1.4.5   Principal components analysis

Principal components analysis (PCA) is a statistical technique used to convert a set of observations, usually measured by many possibly correlated variables, into a set of values of uncorrelated variables. More information on this technique and how it was used in this study is presented in Chapter 5, while an overview of where this procedure has been used in previous research is now given. Due to the nature of this thesis, this review concentrates on PCA in healthcare and hip research. It should be noted that due to the plethora of literature available on this subject this will not cover every area of interest.

A recent study investigated the feasibility of using multivariate analysis to derive summary factors to predict hip fracture. While it was found that PCA did result in composite factors

appropriate for this, neither of the factors found were better than the individual measurements used as inputs to predict hip fracture risk (LaCroix et al. 2010). However, a separate study concluded that PCA was successful in identifying contributions to the risk of hip fracture, indeed finding this better than other methods (Gregory et al. 2004). A review paper also presents the reasons why PCA is an appropriate method for comparing femoral form and structure, which may be then used to assess the risk of hip fracture with respect to femoral geometry (Gregory and Aspden 2008). Certain structural features of the femoral head are related to an increase in fracture risk (Black et al. 2008, Kaptoge et al. 2008), as well as influencing the location and type of fracture sustained (Szulc et al. 2006), so accurately modelling these features allows for better assessment of osteoporotic hip fracture risk. These benefits have been documented (Bryan et al. 2009), with a particular focus on the development of a statistical model, also constructed using PCA, of the whole femur to include inter-patient variability. This model was then used, as previously, to assess the risk of femoral neck fracture.

PCA was used to quantify and summarise gait data on patients suffering from a hip disease in order to obtain a simple evaluation criteria for quantitative gait evaluation (Yamamoto et al. 1983). This was achieved by reducing ten original items into three principal components.

Sexual dimorphism patterns in hip bones were investigated using PCA and it was found that a large amount of dimorphism is accounted for by size differences (Arsuaga and Carretero 1994). A particularly interesting result found was that female hip bones are different in traits associated with a larger pelvic inlet; these physical skeletal differences may go some way to explaining differences in fracture prevalence and fracture types between genders.

The genetic factors which contribute to variability in bone mass, and thus fracture risk, were investigated by Karasik et al. PCA proved to be a successful approach and conclusions regarding genetic influences were reached (Karasik et al. 2004).

As a final example of where PCA has been used in studies relating to (hip) fractures, Sipilä et al used this technique in order to condense muscle strength results into one score, then later to investigate the relationship of this with the risk of all fall-related fractures (Sipilä et al. 2006).

Other fields within the medical literature where PCA has been used include the development a household socioeconomic status index in order to reach useful conclusions regarding care-

seeking behaviour (Schellenberg et al. 2003), summarising scientific information for assessing health claims for foods and supplements, in particular relating to coronary heart disease (Castro et al. 2005), neuroanatomy (Dien et al. 2003), psychiatry (Robertson et al. 2008, Stewart et al. 2007), and identifying unusual lung function in males as well as providing a means for defining and quantifying different aspects of lung function (Cowie et al. 1985).

1.4.6   Operating theatre management

Four key management issues have been identified that need to be addressed in order for operating theatres to achieve high levels of efficiency; the system's rewards, ineffective logistical design, reluctance to accept responsibility and lack of effective teamwork (Calmes and Shusterich 1992). The first of these is dependent upon costs, and the last two must be tackled by softer methods; changing opinions and better communication.   The main organisational changes that can be made, arguably giving the greatest impact, come under the second issue.   In Chapter 8, a simulation model of the trauma theatre at the UHW is presented. This model was built to investigate the impact of different management strategies on the performance of the theatre, in particular with regard to effective approaches to reduce the number of cancellations made due to running out of scheduled theatre time. There is an abundance of literature in this field and an overview is now presented. Further information focussing specifically on the scheduling of operations is given in Chapter 11; this topic is less pertinent for this research since due to the unpredictability of emergency arrivals, the creation of a theatre schedule in not really feasible.

Lemos et al compared retrospective results for pre- and post-implementation of a dedicated orthopaedic trauma theatre. No differences in mortality were found, but morbidity rates were significantly reduced after the establishment of the dedicated trauma theatre, despite an increase in surgical delay (Lemos et al. 2009). Some similar results were reported in an earlier study, which found that a system using a dedicated orthopaedic trauma theatre had roughly half of the post-operative morbidity of a regular system with no dedicated time for orthopaedics. However, in this case surgical delay was also shown to decrease, again by approximately half (Elder et al. 2005).

A comparable approach was used by an American hospital which trialled the use of an unbooked orthopaedic trauma theatre, in which no elective cases may be scheduled, with a primary aim of a reduction in the growing trend of orthopaedic cases being done at night. Improvements reported by using the unbooked system include a reduction in hip fracture surgeries performed at night, a reduction in theatre overutilisation and a shorter surgery time (Bhattacharyya et al. 2006). Similar results, also reporting fewer cases done at night after the implementation of a dedicated emergency theatre, have also been reported elsewhere (Calder et al. 1998, Sweetnam et al. 1994). Benefits of reducing surgeries at night include greater job satisfaction for surgeons (Ostrum 2003) and avoiding errors attributable to tiredness; for example, one survey found that 33% of all self-reported medical errors were associated with fatigue (Gawande et al. 2003). This topic is covered in more detail in Section 7.3.2.

However, Wullink et al found that reserving capacity in elective theatres for emergency cases was preferable to having a dedicated emergency theatre. Their policy means that an emergency arrival can be operated on as soon as any elective case currently occupying the theatre has finished, instead of reserving costly theatre capacity for whenever an emergency case may arrive. Waiting times, theatre utilisation and staff overtime all showed an improvement in results (Wullink et al. 2007). A study in Sweden also found that reserving some capacity for emergency cases, coupled with a policy to increase staff on standby during this time, significantly improved the performance of the operating theatre department (Persson and Persson 2010).

A compromise between these approaches is to include some 'deferrable' elective patients in an emergency session. This is a patient who may be offered earlier treatment, if emergency demand is low, in return for accepting the possibility of postponement, if the emergency demand on the day of their appointment is high enough for elective patients to be cancelled. It has been suggested that an elective list equivalent to 90 minutes per trauma session may be beneficial in reducing theatre underutilisation during dedicated orthopaedic sessions (Bowers and Mould 2002). The relationship between elective inpatient services and emergency admissions has also been studied analytically, where the introduction of a booked admissions policy for elective patients was considered. Emergency admissions were included in calculations for total demand and different booking systems were modelled with a key result of showing the variance in demand for beds (Utley et al. 2003b). Variation in bed requirements is something discussed in more detail later in this thesis.

Tardiness relating to start time, theatre turnover time (the time between consecutive operations) and resource underutilisation are important matters to consider with respect to theatre efficiency. These topics are discussed later in Chapters 7 and 8 but a review of policies used to tackle these issues is now presented.

One suggested strategy to avoid early finishes (and hence reduce underutilisation of the theatre) involves moving patients to the trauma theatre from other lists which are likely to finish early, or to schedule elective orthopaedic patients within the sessions which could improve the overall utilisation of the trauma theatre (Bowers and Mould 2004). Dexter developed a strategy to decide whether to move the last scheduled case of the day to another empty operating room, in order to reduce overtime costs, with positive results (Dexter 2000).

Another approach which involved moving cases to a different theatre, specifically when theatres are running late, was shown to reduce tardiness in those moved cases by 50-70%. However, since few cases were moved, overall tardiness was only decreased by 6-9% (Wachtel and Dexter 2009). The main cause of tardiness relating to start time can be attributed to staff, thus the main way to combat this tardiness is by changing human behaviour (Lapierre et al. 1999).

A change in human behaviour has also been shown to contribute towards a reduction in turnover times (Overdyk et al. 1998). A decrease in turnover times between operations has also been shown to significantly decrease via the use of a second anaesthetist to commence induction of a patient just as the previous case is being completed (Sokolovic et al. 2002, Torkki et al. 2005). A team from Finland used simulation to compare four different parallel workflow models to the traditional model where patients are operated on in sequence, one at a time, without any overlap. It was found that each of the parallel models gave better cost-efficiency than the traditionally sequenced working pattern (Marjamaa et al. 2009).

An interesting approach to investigating theatre efficiency was employed by Stepaniak et al, who investigated whether the personality of the Operating Room Coordinator (ORC), who is responsible for filling gaps in the theatre schedule, had any effect. It was shown that less risk-averse ORCs created significantly less unused theatre time, without an increase in the probability of running operating theatres after regular working hours or the number of cancelled cases (Stepaniak et al. 2009b).

1.4.7 Simulation in healthcare

Computer simulation is a popular methodology in the healthcare field and is used to analyse, understand and investigate the workings of a health system. This technique is employed twice in this thesis, see Chapters 6 and 8, where the reasoning for choosing simulation over other techniques is also discussed.

There is an abundance of existing literature in this field and some excellent and thorough survey papers are available (Günal and Pidd 2010, Jun et al. 1999, Mielczarek and Uziałko-Mydlikowska 2012). Brailsford et al used a systematic heuristic to produce a final dataset of 342 healthcare papers and present a useful breakdown of a number of variables, including methods, level of implementation and functional area. This study was conducted with a particular, but not exclusive, interest in applications of simulation (Brailsford et al. 2009). An earlier systematic review, which did focus specifically on the use of discrete event simulation in healthcare, had 182 papers which met the authors' inclusion criteria (Fone et al. 2003). Clearly, simulation in healthcare is a prolific research area.

Whole hospital simulation models are uncommon. Reasons for this include the level of complexity (and therefore the data) that these models would require and the resources required to build such a model (Jun et al. 1999). Despite these challenges, there are examples reported in the literature; these include a study by Van der Meer et al which models every treatment phase of a subgroup of elective orthopaedic patients (Van der Meer et al. 2005), and a whole-hospital model which covers all bed-related activities (Cochran and Bharti 2006).

Modelling A&E departments using simulation is a well-researched area. These models tend to be exclusive to particular departments, but a model of a generic A&E department has been developed for use in UK hospitals. The original intention was to gain a better understanding of patient flow, but was later developed so that it could be applied locally by individual hospitals (Fletcher et al. 2007). Hospital-specific models usually require a high level of detailed information to be used as inputs, but if this can be achieved then helpful insights can be gained into prospective strategies to improve throughput (Duguay and Chetouane 2007, Huang et al. 1995). Baboolal et al used simulation to develop a 'perfect world' model of an A&E department, modelling the unit not how it currently is, but how it could be, with particular attention paid to different staffing configurations (Baboolal et al. 2012). This

'perfect world' approach is employed in Section 10.5 with regard to modelling the trauma theatre, where results are instead generated using a mathematical approach.

Simulation is also widely used to model outpatient departments. An extensive review of papers which deal with methods of solving scheduling problems in outpatient clinics has been reported by Cayirli and Veral, and was later extended to use simulation to evaluate the performance of a variety of appointment systems for ambulatory care (Cayirli and Veral 2003, Cayirli et al. 2006). It was concluded that patient sequencing (for example, first-come-first-served) had the greatest effect on care performance than the method used to determine appointment times.

There are also a multitude of research papers reporting using simulation methods to model inpatient departments. Examples include a model designed specifically to simulate geriatric patients in a North London Health District, with the intention that results could be used to evaluate the effectiveness of the department and to demonstrate the effects of changes to the current system (El-Darzi et al. 1998).

Simulation models have also been developed to predict future demand on healthcare services by incorporating demographic changes, in particular with regard to long-term care of the elderly (Lagergren 1994) and future requirements of healthcare provision caused by fractures relating to osteoporosis (Bleibler et al. 2012).

Briefly, other examples include using simulation to investigate means of reducing the spread of MRSA (Barnes et al. 2010), modelling bed occupancy in a critical care unit with the aim of finding a suitable strategy to minimise elective surgery cancellations (Griffiths et al. 2010) and modelling the effect of HIV (human immunodeficiency virus) treatment on transmission rates (Gray et al. 2003).

1.4.8 Queuing theory in healthcare

In Chapters 9 and 10, queuing theory is used to analytically model the trauma theatre at the UHW, by utilising existing results and developing a bespoke queuing system. Bailey is widely credited as pioneering the use of queuing theory in healthcare several decades ago (Bailey 1952), and since then queuing theory has been extensively used in healthcare and

some examples of published work are now given. The interested reader is referred to some survey papers on this topic for a fuller summary (Creemers et al. 2007, Fomundam and Herrmann 2007, Green 2006).

Finding ways to minimise waiting times for patients is often a primary focus of queuing theory studies, while also fulfilling the discordant objective of maximising the utilisation of resources. It has been shown that increasing capacity will not necessarily reduce waiting times, since arrival of the users (the patients) can be reactive to the system state; as users realise that waiting times are decreasing, more of them will arrive, thus increasing waiting time and queue size once more (Worthington 1991). In 1970, a different approach was taken for the first time whereby queuing theory was used to establish an index of quality of care based on service waiting times (Haussmann 1970).

Phase-type distributions are used to describe the time until absorption for a chain of n finite transient states, where each phase can be considered to be a Poisson process. Fackrell presents a thorough and recent overview of the use of phase-type distributions in healthcare (Fackrell 2009). A specific case which lends itself well to the healthcare setting is the Coxian phase-type distribution, where all arrivals (patients) start in phase 1 and sequential transition is possible between any state i ($i = 1, ..., n$) to the next state $i + 1$. Exit from the system to the absorption state ($n + 1$), usually discharge or death, can occur from any phase. Recent applications include modelling emergency services by using Coxian phase-type distributions to represent overall service time, split by case urgency, as well as times for sub-stages of service (Knight and Harper 2012). Coxian phase-type distributions are commonly used to model heavily-skewed data and another recent paper found that this methodology was useful in not only modelling length of stay, in this case for patients who have suffered from a myocardial infarction, but also for other skewed distributions such as censored data and inpatient costs (Tang et al. 2012).

Length of stay for geriatric patients is often shown to be heavily-skewed and has been modelled using Coxian phase-type distributions, where it was also shown how this methodology can be used to include influencing variables such as patient age (Faddy and McClean 1999). The survival of patients after admission due to a hip fracture has been modelled using conditional phase-type distributions (Marshall and Shaw 2008). In particular, this methodology was used to identify system delays and how addressing them could reduce

length of stay, and has also been used to include clinical and demographic data to predict length of stay and discharge destination for geriatric patients (Marshall et al. 2002).

The Erlang distribution is a particular case of the Coxian phase-type distribution where all phases must be completed sequentially before transfer to the absorption state can occur, and is used later in this thesis to model the service time for the trauma theatre.

Effective management of expensive hospital resources is another area where queuing theory has been used and a particular relevant example is optimising the number of beds at a care facility. Gorunescu et al used a phase-type queuing model to minimise the number of beds in order to achieve at most a certain pre-determined probability of delay at a geriatric hospital department (Gorunescu et al. 2002). Utley et al used an analytical approach to investigate the possibility of introducing an intermediate care facility, so that patients occupying acute care beds who do not require acute care would instead be cared for at such a facility. Results were used to suggest the proportion of all beds at the facility that should remain designated for acute care (Utley et al. 2003a). Results have also been applied to intensive care units, where beds, equipment and staff are very costly (Costa et al. 2003, Griffiths et al. 2006).

Indeed, queuing theory has been used to determine staffing levels in healthcare for some time. An early application involved applying a queuing model to a hospital messenger service (several servers) to find an appropriate trade-off between cost and efficiency measures such as waiting time and queue length (Gupta et al. 1971). More recent applications include using queuing theory to reallocate staff in order to increase throughput in an emergency department, despite an increase in demand (Green et al. 2006), while Lucas et al used mathematical modelling to calculate the probability of certain types of admission requiring urgent surgery for different arrival rates. Results were then used to find the probability that two operating rooms would be simultaneously occupied and decisions could then be made regarding staffing levels. It was concluded that an on-call team provide sufficient staffing to achieve immediate operating room availability for centres admitting fewer than 500 cases per year (Lucas et al. 2001). This methodology was later repeated to determine staffing requirements for paediatric operating theatres (Tuggle et al. 2010).

As a final example in healthcare, queuing theory has also been applied to pharmacy applications and Nosek and Wilson give a thorough survey of research in this area, with particular attention to improving customer satisfaction (Nosek and Wilson 2001).

CHAPTER 2: WARD DATA AND INITIAL ANALYSES

2.1   Introduction

This chapter is primarily an introduction to one of the databases made available for this study, namely the Cardiff Hip Fracture Survey database. Many different variables are recorded into this database for each patient admitted to the University Hospital of Wales with a trauma fractured neck of femur.

2.1.1   Cardiff Hip Fracture Survey

The Cardiff Hip Fracture Survey database includes information relating to patient's personal information, their medical condition, their treatment and their discharge. In some cases follow-up data was also available but unfortunately was not reliable or complete enough to use here.

This data is collected and recorded by the hip fracture team and initially recorded for each patient onto a data capture sheet, then entered into an SPSS database whenever time permits.

Before more sophisticated statistical analysis is undertaken, an overview of some of the data available is now presented.

The extract of data used here covers all patients admitted between October 2003 and mid-February 2008, a total of 2182 observations. Missing values and errors are inevitable for manually entered data. All data was checked and validated prior to the following analysis being undertaken. Any erroneous values (a negative age, for example) were removed and regarded as a missing value if they could not be inferred from other available fields.

2.2 Data

2.2.1 Age and gender

The majority of patients admitted with a trauma hip fracture are elderly. Another important characteristic of this injury is that they usually occur within females. The reason for this is two-fold; firstly, since hip fractures are more likely to occur in the older population due to the onset of osteoporosis, the issue of average life expectancy is considered in order to gain an insight into the age distribution of the population. It is a general conception that women live longer than men, hence the elderly population would be expected to have a higher proportion of women than men. For that reason it is important to check whether or not this is true for the areas of interest here. The generally accepted definition of an 'elderly' person being defined as one who is the age of 65 or above is used here.

Life expectancy is calculated in two ways in the United Kingdom; at birth and at age 65. Life expectancy at a given age for an area, in the specified time period, is an estimate of the average number of years a person of that age would survive if he or she experienced that area's age-specific mortality rates for that time period, throughout the rest of their life. The figure therefore reflects mortality among the population living in a certain area, rather than those born in the area. Life expectancy at age 65 is the number of further years a person who reaches the age of 65 in the specified time period could expect to live.

Life expectancy figures for the local authorities of Cardiff and The Vale of Glamorgan are given in Table 2.2.1i. These values are calculated and disseminated by the Office for National Statistics (ONS 2009).

Table 2.2.1i: Life expectancy at birth and at age 65 for Cardiff and The Vale of Glamorgan, 2006-2008

Local authority	Life expectancy at birth [95% confidence interval]		Life expectancy at age 65 [95% confidence interval]	
	Males	Females	Males	Females
Cardiff	76.6 [76.2, 77.0]	81.7 [81.4, 82.1]	17.0 [16.7, 17.3]	20.3 [20.0, 20.6]
The Vale of Glamorgan	77.8 [77.1, 78.4]	82.3 [81.8, 82.8]	17.6 [17.2, 18.0]	20.3 [19.9, 20.6]

It can be seen for the most part that the life expectancy is greater for The Vale of Glamorgan than for Cardiff, while in all cases women have a higher life expectancy than men. National figures for Wales show that the two authorities studied in detail here are roughly on a par with the rest of the country. The notion that women live longer than men is therefore justified and in part explain why there are more women in the dataset than there are men.

The second issue to consider is that it is known women's bones tend to deteriorate sooner than men's bones and a bone is more likely to break under pressure, from a fall for example, if it has weakened over time. One explanation for this is that the onset of the menopause accelerates bone loss in women (Wei 2004); the average age at which women reach the menopause in the United Kingdom is 52 (NHS 2010).

Table 2.2.1ii shows some summary statistics on age and gender at the time of admission (S.D. – standard deviation). It can be seen that for every male admitted, approximately three females were admitted to the hospital over the same time period. Additionally, the average age of female patients is around five years greater than the average age of male patients. The variation in ages is higher in males, but this is in part due to the more extreme outliers seen amongst the male patients. Splitting into age groups of ten years allows the spread within age groups to be seen in Figure 2.2.1iii.

Table 2.2.1ii: Summary statistics of age (years), classified by gender

Gender	*n*	Mean	S.D.	Minimum	Maximum
Males	580	76.29	13.25	14	100
Females	1598	81.33	10.12	31	101
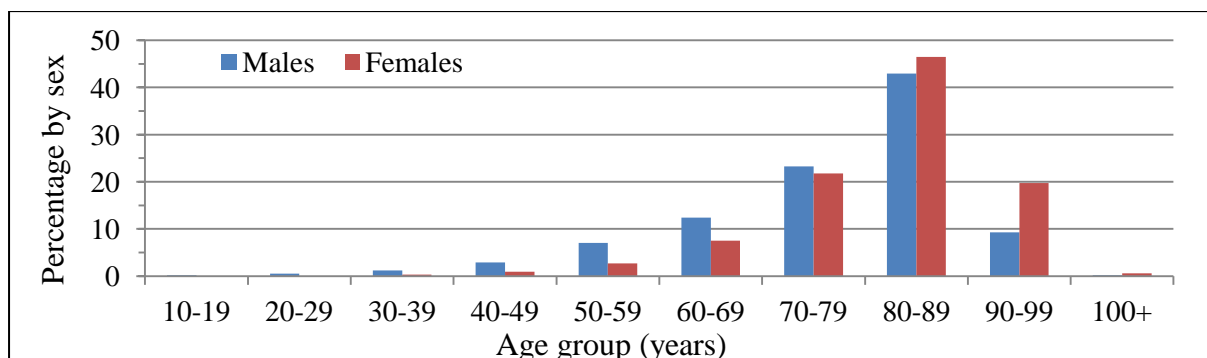
Figure 2.2.1iii: Percentage of patients in each age group, classified by sex

It can immediately be seen that women are more likely to suffer from a trauma hip fracture later in life, while men are shown to have higher percentages in the younger age groups. Of course, the issue that the elderly population is made up of more women than men must be considered here.

Using population estimates provided by the Welsh Assembly Government (WAG 2008a) for the calendar year 2007, alongside the Hip Fracture Survey data from the same year, this aspect can be investigated. Since the University Hospital of Wales takes all cases of trauma hip fracture from across the local authority regions of Cardiff and The Vale of Glamorgan, population figures for these regions were aggregated. Younger age groups with very small numbers are excluded and patients aged 100 years or over are combined with the 90-99 age group.
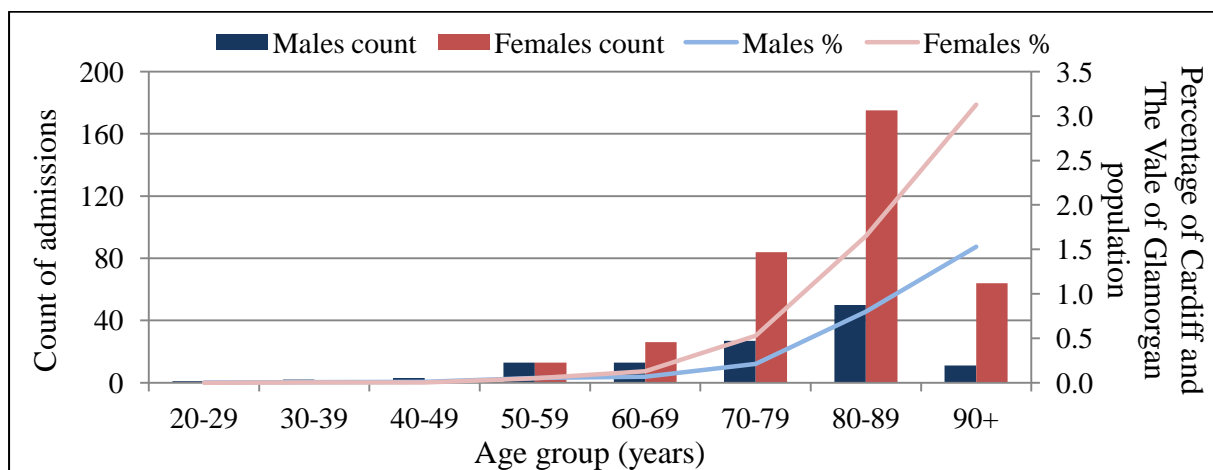
Figure 2.2.1iv: Count of admissions and the percentage of Cardiff and The Vale of Glamorgan population admitted, classified by age group and gender; 2007

For both males and females it can be seen that the number of admissions as a percentage of the relevant population steadily increases across the age groups. Indeed, 3.13% of females in the *90 and over* age category incurred a trauma hip fracture in 2007. As a proportion of the population, women are also more likely to be admitted with this injury.

As a reference point, fracture incidence rates were calculated in a Cardiff study published in 1997 (Johansen et al. 1997). This was across all age groups and for all fracture types. The

overall fracture incidence was 2.11%, 2.35% for males and 1.88% for females, thus highlighting the increased hip fracture risk for the older population.

2.2.2 Admission source

There are eight distinct places from which a patient may have been admitted. Note that this is their current residency, not necessarily the place where the injury was incurred. A count within each group is displayed in Figure 2.2.2i. This value was missing in three cases.
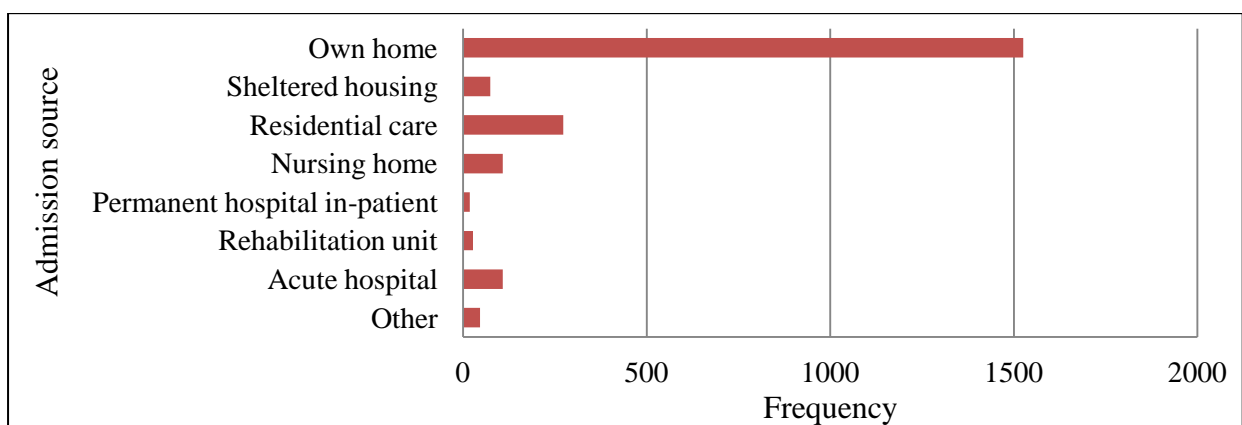
Figure 2.2.2i: Frequency of patients arriving from each admission source

A high proportion (70.0%) of all patients were admitted from their own home. A study in Cardiff previously showed that hip fracture rates are considerably higher amongst care home residents, compared with sheltered accommodation and community dwelling residents (Brennan née Saunders et al. 2003).

2.2.3 Living alone status

Until March 2007, another element of the residential status of each patient was recorded; whether they lived alone or not. This variable could take three values: the patient lived alone, did not live alone, or lived in institutional care. The latter of these options is covered by the previous variable of admission source, so it is only really of interest where the patient is not in institutional care; that is, they lived in their own home or sheltered housing, leaving 1195

and 61 patients respectively, for data items where both living alone status and admission source was available.

Within patients admitted from home, almost equal proportions were seen; 48.5% lived alone and 51.5% did not. However, marked differences can be seen within the smaller group of patients admitted from sheltered housing, where just 6.5% of patients did not live alone.

2.2.4   ASA grade

The American Society of Anesthesiologists (ASA) grade is used to denote a patient's medical fitness. It is a physical status classification system scored on a scale of I (one) to VI (six), which is often used to assess patients prior to surgery. The official definitions of the six grades are given in Table 2.2.4i and were obtained from documentation published by the ASA (American Society of Anesthesiologists 2010). The observed prevalence of each grade are also given and correspond to the value of all patients for this value was known; it was missing for 115 observations (5.2%) in the dataset. Note that no patient with an ASA grade of VI would be admitted to the hip fracture ward.

Table 2.2.4i: ASA grade descriptions and percentages

ASA Grade	Description	Percentage
I	A normal healthy patient	4.4%
II	A patient with mild systemic disease	32.5%
III	A patient with severe systemic disease	52.2%
IV	A patient with severe systemic disease that is a constant threat to life	10.7%
V	A moribund patient who is not expected to survive without the operation	0.2%
VI	A declared brain-dead patient whose organs are being removed for donor purposes	-

Immediately it is noticeable that over half of all patients have an ASA grade of III and around one third have an ASA grade of II, while fewer than five percent have the most desired ASA

grade of I. These results are not especially remarkable considering the high age of the majority of these patients.

It will be seen later in this thesis that grades I and II are combined into one group and grades V and VI are not used. It was advised by the clinicians involved in the project that grades I and II could be combined since there is no difference in the way these patients are treated. It was also advised to ignore patients with an ASA grade of V; these patients are likely to have several comorbidities and it is probable that their hip fracture injury is not the main concern of any treatment plan. While it would be useful to investigate the effect of comorbidities on all patients, unfortunately this information was unavailable. Additionally on inspection it was found that there were very few patients recorded with an ASA grade of V and thus excluding these patients would have negligible impact on later work.

2.2.5 Mental state

Each patient's mental state on admission is recorded as one of three values. This variable was missing in ten cases. Table 2.2.5i gives the frequency and percentage of each of the mental states, listed in order of severity.

Table 2.2.5i: Count and percentage of observations by mental state

Mental state	Frequency	Percentage
Normal	1460	67.2%
Known dementia	396	18.2%
Confusion	316	14.5%

Just over two-thirds of all patients do not have any mental problems as recorded by these definitions and are classified as 'normal'. It would be expected to have reasonable amounts in the two other categories, due again the high age profile of these patients (Jorm and Jolley 1998, Nicholl 2009).

2.2.6 Walking aids and ability

The walking aids and walking ability pre-fracture of each patient are recorded as part of the Hip Fracture Survey. These are both measured on an ordinal scale of one to five:

Table 2.2.6i: Description of walking aids and ability values

Value	Walking aids definition	Walking ability definition
1	None	Outside, alone
2	One (stick, crutch)	Outside, with someone
3	Two	Inside, alone
4	Frame	Inside, with someone
5	Wheelchair / bed-bound	Wheelchair / bed-bound

Both of these data items were available in all but seven cases. A useful way to view this data is via a so-called 'bubble' plot, as seen in Figure 2.2.6ii. The axes correspond to the values which the two variables, walking aids and walking ability pre-fracture, may take. The size of the bubble (its area) corresponds to the frequency in the dataset which take these values. For reference, the bubble at (1, 1) represents 572 patients, while the bubble at (5, 5) represents 48 patients.
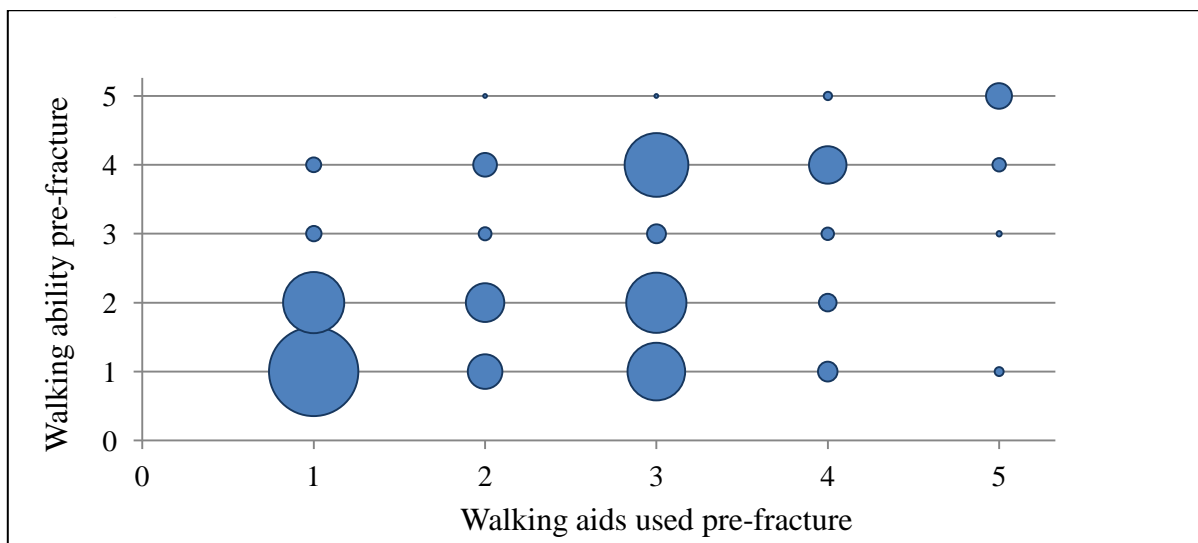
Figure 2.2.6ii: Bubble plot displaying the frequencies between walking aids used and walking ability pre-fracture

Clearly there are some discrepancies, since if a patient has a value of five recorded for one of these variables, then it is sensible that they will also have the same value recorded for the other variable, which is not always the case. However, this occurs in a rather minimal amount of cases so do not lead to much cause for concern.

These results are unsurprising considering the age profile of this patient group; an increasing age is expected to be associated with a decline in mobility (Troosters et al. 1999). However, they do highlight the difficulties faced by a large amount of these patients even before incurring their hip fracture injury.

While modern methods of hip replacement surgery can provide highly functional outcome for elderly patients with a fractured hip (Schmidt et al. 2009), one study showed that only half of hip fracture patients managed to regain their pre-fracture functional status, as measured by walking ability and the need for walking aids (Sernbo and Johnell 1993) so some impact would clearly be expected. It has additionally been shown that men experience better functional recovery than women after a trauma hip fracture (Arinzon et al. 2010). Unfortunately the follow-up information relating to walking aids and ability was scarcely populated.

2.2.7 Mobility

A similar variable is the mobility score of a patient. This is measured on an ordinal scale of one to three, as defined in Table 2.2.7i alongside observed frequencies and percentages. There were eight observations for which this variable was missing. Almost 40% of all patients in this dataset are housebound, while an almost exact amount has the 'best' mobility score of *1*.

Table 2.2.7i: Count and percentage of observations by mobility score

Mobility score	Description	Frequency	Percentage
1	Able to shop	858	39.5%
2	Able to get outside but unable to shop	463	21.3%
3	Housebound	853	39.2%

Note that while a higher level of mobility may indicate that an individual is generally medically fitter and more able, in actuality this may mean that they are more at risk from fracture due to a more active lifestyle (Porter et al. 1990).

2.2.8   WAASP score

The WAASP score is a nutritional screening tool used within the Cardiff and Vale University Health Board. It is an acronym for Weight, Appetite, Ability to eat, Stress factors, Pressure sores / wounds, so encompasses various factors into one score. A lower score is more desirable.

The actual score is not recorded for hip fracture patients but is converted to a category of *1* for WAASP scores with a value of one or two, *2* for WAASP scores between three and six and *3* for a WAASP score of seven or greater. Inevitably this means that a certain amount of information is lost. 24.5%, 42.3% and 33.2% of patients had WAASP categories of *1*, *2* and *3* respectively. This value was not recorded for 18 patients.

The highest proportion of patients thus falls within the middle category, leaving a third of patients in the top ('worst') category and a quarter of patients in the bottom ('best') category. Evidence shows that this group of patients are likely to be malnourished on admission and show a rapid deterioration in nutritional status during admission (Nematy et al. 2006). This deterioration may be due to a variety of factors including oral problems, mental difficulties, anorexia and catering limitations (Patel and Martin 2008).

2.2.9   Pathological fracture

A pathological fracture occurs when a bone breaks in an area that is weakened by another disease process. Causes of weakened bone include tumours, infection and certain inherited bone disorders. There are six codes used to specify whether each patient has a pathological fracture and, if so, what the nature of the pathological fracture is; descriptions and results are given in Table 2.2.9i. More detailed medical definitions of these descriptions can be found in Appendix C. This data item was missing for 47 observations.

Table 2.2.9i: Pathological fracture codes, their associated descriptions and percentages

Code	Description	Percentage
1	None	95.55%
2	Malignant secondary bony tumour	2.72%
3	Malignant primary bony tumour	0.23%
4	Bone cyst	0.14%
5	Paget's disease	0.70%
6	Other	0.66%

The vast majority of patients do not have a pathological fracture; they have incurred a fracture which was not caused by bone weakness due to another disease. Within the remainder, the most common type of pathological fracture is a malignant secondary bony tumour, but this still accounts for less than 3% of all patients.

2.2.10 Fracture type

There are six different classifications used for how the fracture is described clinically, which are detailed with counts in Table 2.2.10i. This value was missing for 29 observations and again there is more detailed information relating to these categories available in Appendix C. The most common type of fracture seen was displaced intrascapular, with 43.6% of all values. It would have been useful to see whether the type had any relation to whether or not it is a pathological fracture, but analysis is difficult since few fractures were pathological.

Table 2.2.10i: Type of fracture codes and their associated descriptions

Code	Description	Count
1	Undisplaced intracapsular	227
2	Displaced intracapsular	939
3	Basocervical	108
4	Trochanteric, two fragment	370
5	Trochanteric, multi fragment	380
6	Subtrochanteric	127

2.2.11 Type of operation

The classification system for the recording of operation type was amended in March 2007. With the assistance of a clinician at the hospital, new categories were defined in order to reclassify the existing data which would suit both of the previous systems, but unfortunately some information was lost in forming the new definitions. Operation type classifications and percentages are given in Table 2.2.11i and more information is available in Appendix C. This information was available or could be inferred in all but eight cases.

Dynamic hip screw and hemiarthroplasty are the most common operation types and the data shows almost equal numbers of patients in these two groups. Just 3.4% of patients do not undergo surgery, either because they are given non-surgical (conservative) treatment or because they die prior to the operation being performed.

Table 2.2.11i: Operation codes, their associated descriptions and prevalence

Code	Description	Percentage
A	No operation; died pre-operation or given conservative treatment	3.4%
B	Dynamic hip screw	35.3%
C	Screws	11.5%
D	Intramedullary nail	9.7%
E	Hemiarthroplasty	35.4%
F	Total hip arthroplasty	3.7%
G	Other	0.9%

2.2.12 Side of fracture

The side of the body, left or right, on which the fracture was incurred is another recorded data item. This value was missing in just eight cases and as expected there is no noteworthy difference in proportions between these two groups. 1163 (53.5%) patients incurred a left-sided fracture and 1011 (46.5%) incurred a right-sided fracture.

2.2.13 Acute discharge destination

Ten different values are recorded for acute discharge destination. It is also important to consider the admission source of a patient, as well as their discharge destination, as it can then be seen if there has been any change in the residential status of a patient which could be attributed to the hip fracture. However, in this brief overview just the destinations are given. More detailed analysis is seen later in Chapter 6. This value was missing for 27 observations in the dataset and counts are presented in Figure 2.2.13i.
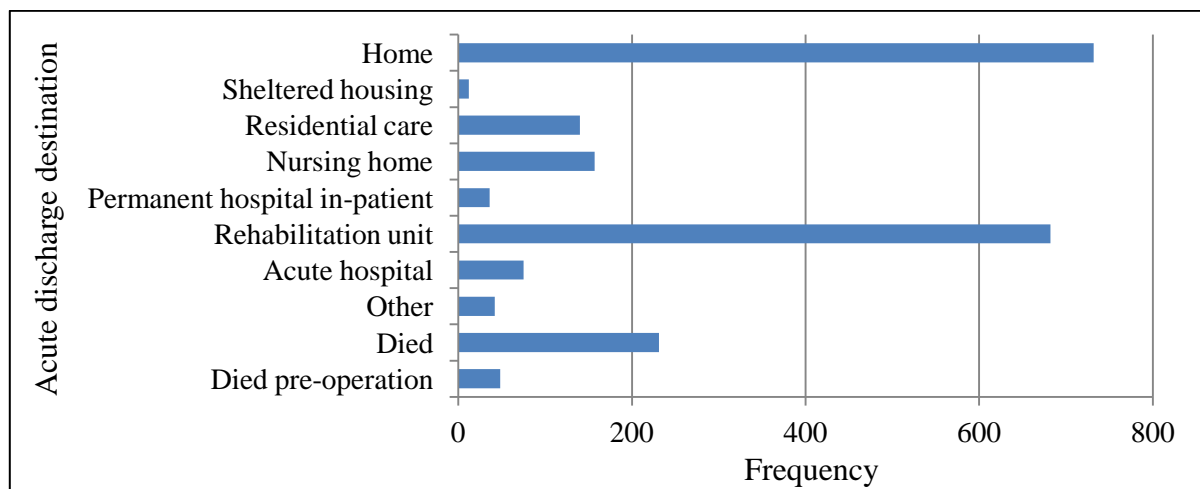
Figure 2.2.13i: Count of observations by acute discharge destination

The two most prevalent acute discharge destinations are home and rehabilitation unit, with 34.0% and 31.6% of all observations respectively. Many patients therefore do not actually finish their treatment once they leave the hip fracture ward, but remain under the care of the University Health Board (UHB) for further treatment. 12.9% of patients do not survive their stay in hospital, and 17.2% of these die pre-operation.

Some information is available on the final discharge destination of each patient; their ultimate discharge destination when they leave the UHB. For many patients the acute and final discharge destinations are the same. This is not discussed in more detail here but is analysed further with respect to mortality in Chapter 4.

2.2.14 Length of stay

Patient length of stay is discussed in more detail in Chapter 3, but a brief introduction is given here for completeness. There are two length of stay values recorded; ward length of stay and UHB length of stay. Many patients are not discharged completely when they leave the ward at the UHW, but are transferred elsewhere within the UHB (see Section 2.2.13). Summary statistics are given in Table 2.2.14i. Data was available in 2157 and 1914 cases for ward and UHB length of stay respectively.

Table 2.2.14i: Summary statistics for length of stay (days)

Length of stay category	Mean	S.D.	Minimum	Maximum
Ward	28.72	29.26	0	354
University Health Board	39.93	41.21	0	534

Length of stay is shown to be almost one month, on average, for trauma hip fracture patients, with approximately eleven more days if they are not discharged from the UHB when they leave the ward. One patient actually spent almost a year and a half within the UHB after being admitted with this injury.

2.2.15 Other collected data items

A number of other data items are collected in the Cardiff Hip Fracture Survey which are not discussed in more detail. Many of these are follow-up items relating to patient condition or status after they are discharged and is therefore difficult to collect, resulting in sparse data.

Other routinely collected information includes the ward which the patient was on. A small amount of ad-hoc work was completed with regards to ward, where the focus can be summarised as analysing whether patients receive the same treatment on each ward (measured crudely as outcome for similar patient groups). It was shown that there were no significant differences; for this reason and with the agreement of a clinician involved with this work it was decided to ignore this variable.

2.3 Bed occupancy

Section 2.2 presents simple analyses of the raw data collected by the Cardiff Hip Fracture Survey. This data represents the characteristics of this patient cohort, as well as some information on the injury sustained and the treatment given. The progress of patients as they are admitted and subsequently discharged from the ward is also given. However, no data regarding all patients as a whole is routinely collected and recorded. This makes analysis of the utilisation of the resources within the ward difficult to undertake, which would have knock-on effects in terms of capacity, resource and patient management.

One example of interest is the number of beds occupied at any time within the hospital by this patient group. By using the arrival and discharge dates of each individual patient, a daily bed occupancy profile can be developed. However, this will always provide an overestimate of the number of the number of beds occupied. Consider a patient discharged in the morning on any given day. The bed that this patient has just vacated thus becomes empty and available for another patient admitted later in the day. The method used here records this bed twice, once for each patient. However, while the admission time is available for patients admitted under the new recording system, there is no information available for the time a patient is discharged. This is therefore an unavoidable obstacle in calculating bed occupancy. One justification of 'ignoring' this issue is that there will always be some unknown turnaround time for the bed to become available, so any assumption that a bed becomes immediately available after a patient is discharged is inaccurate.

A program was written in Visual Basic that reads in the admission date and the discharge date at an individual patient level and then calculates the number of ward beds occupied on each day over a given time period. The results were output to an Excel spreadsheet and sensible cut-offs were made with respect to the start and end points of the time period which could be used later. This resulted in the number of beds occupied on each day for just over three years (May 2004 to July 2007) being available for further analysis. Summary statistics of the number of beds occupied over this time period are given in Table 2.3i. Clearly the number of beds occupied fluctuates considerably, dropping to a minimum of 22 beds and rising to a maximum of 71 beds. This variation is displayed on a day-by-day basis in Figure 2.3ii. Finally, a frequency histogram of beds occupied is presented in Figure 2.3iii. While it

appears that the number of occupied beds may follow a Normal distribution, a formal test showed that this was not the case.

Table 2.3i: Summary statistics for bed occupancy, trauma hip fracture patients

Mean	Standard deviation	Minimum	Maximum	Mode	Median
42.10	8.07	22	71	37	41
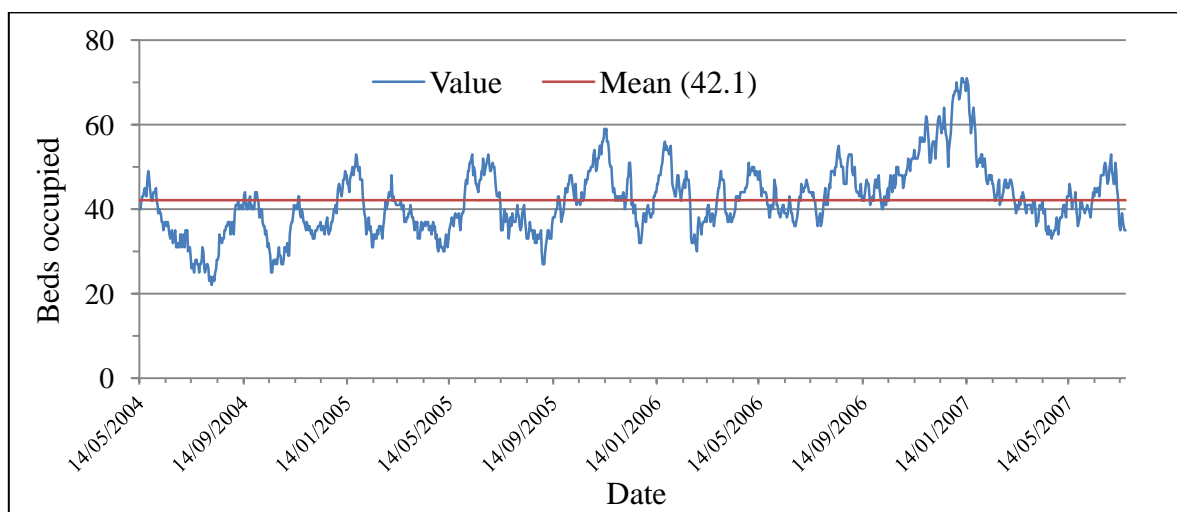
Figure 2.3ii: Daily beds occupied by trauma hip fracture patients, May 2004 – July 2007
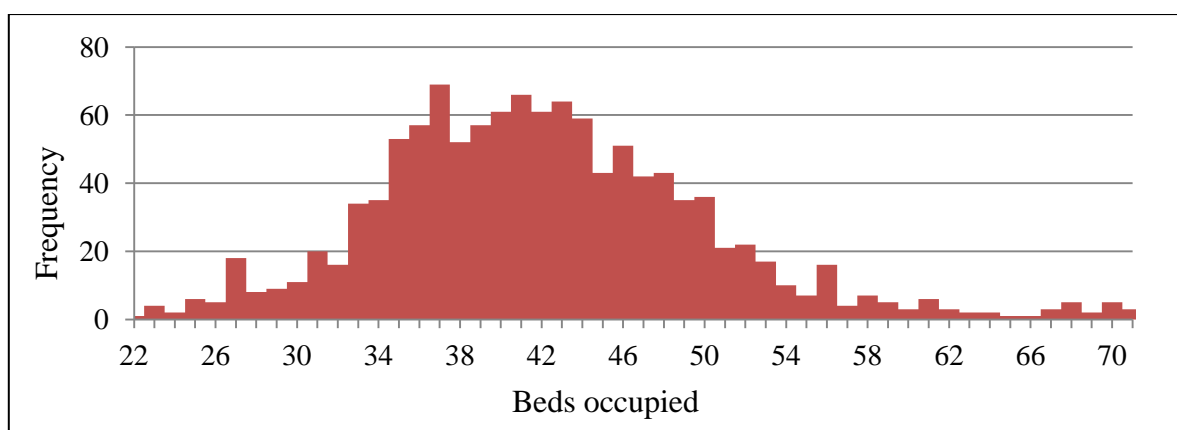
Figure 2.3iii: Histogram of beds occupied by trauma hip fracture patients, May 2004 – July 2007

While the variation seen here would be evident to the hip fracture team from daily observations within the ward, it is unlikely that the extent of this variation or indeed the values of bed occupancy are known within the ward. A greater knowledge and comprehension of areas such as bed occupancy will aid resource and patient management.

Using the values of final length of stay would also mean that the number of beds occupied within the University Health Board (UHB) could be evaluated. However, this value was missing in many cases and if a patient did not leave the UHB when they were discharged from the ward then it could not be inferred from other data items. Any calculations would therefore be an underestimate of the true value and since there is no way of knowing the extent of this underestimate precisely then these calculations are not included here.

2.4 Time dependency

It may be supposed that admissions are time-dependent and day of the week, monthly and yearly investigations into time dependency are now presented.

The first time measurement that is considered here is the day of admission. If one day had a particularly higher or lower number of admissions in comparison to the other days then this may be something that needs to be taken into account later on. The distribution of the number admissions according to day of the week is given in Figure 2.4i. Friday is shown to have the most admissions, while Sunday has the least, but it can be seen that each of the seven segments is relatively evenly-sized.
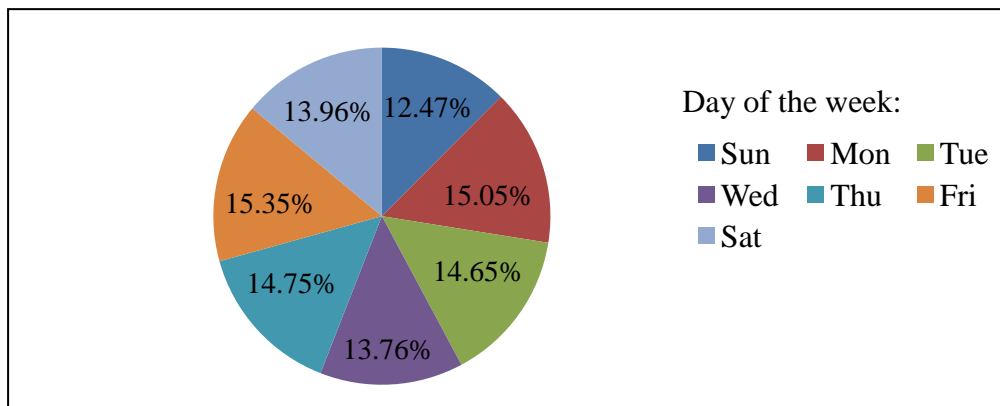
Figure 2.4i: Percentage of admissions by day of the week, 2004-2007

It may also be supposed that the number of beds occupied would increase over the winter months; the majority of trauma hip fractures come as the result of a fall and icy conditions may lead to a rise in this. While the highest peak on the graph shown in Figure 2.3.1ii is seen to occur during January, there is another peak during the summer months of 2005, for example. These peaks and troughs are therefore attributed to randomness alone.

Monthly analysis was completed where a complete calendar year of data was available, namely for the years 2004 to 2007. Counts of admissions over these months are given in Figure 2.4ii. While the highest total count is seen for the month of December, the assumption of no seasonal trends is affirmed here. The second highest count is seen in May, while January and August had exactly the same number of admissions over this four year period.
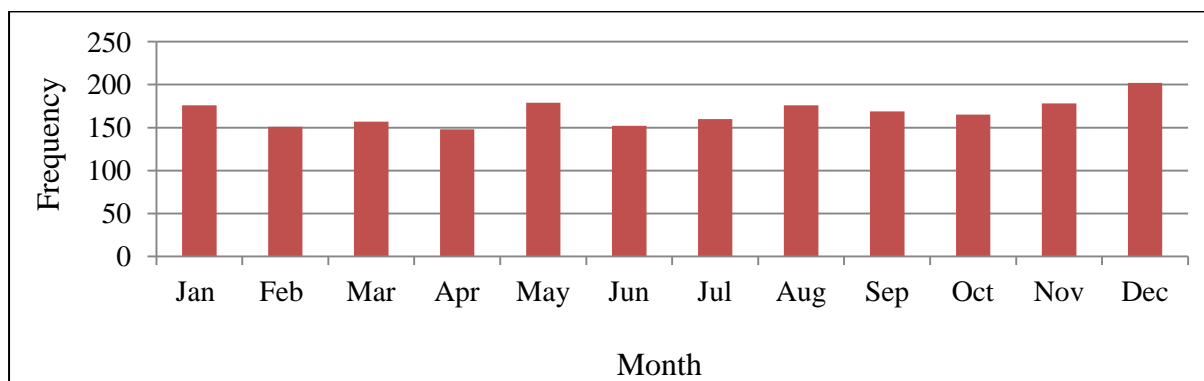
Figure 2.4ii: Count of admissions by month, 2004-2007

While it appears that there are no monthly or daily patterns to be taken into account, any increasing or decreasing patterns over time must be considered. The number of admissions per year is given in Table 2.4iii.

Table 2.4iii: Count of observations by year, 2004-2007

Year	2004	2005	2006	2007
Count of admissions	503	490	533	487

There appears to be no upward or downward trend in the number of admissions over time and therefore it can be assumed that the number of admissions per year is reasonably stable at around 500 patients, with fluctuation in the number of admissions over this period attributed to randomness. It may be expected that over a longer period of time the number of admissions would increase due to the ageing population, as discussed later in this thesis.

Day of the week, monthly and yearly analysis have all been completed in the attempt to find some pattern or trend in the number of admissions to the hospital with a trauma hip fracture with respect to time. Results have indicated that there is no time-dependency involved with these admissions.

Consider the null hypothesis of the Chi-square (χ^2) test of independence, which states that two variables do not have an association or relationship with each other. This can be utilised

here to more formally prove the lack of association between each of the three time values considered and the number of admissions. A *p*-value of less than 0.05 would suggest that the null hypothesis should be rejected in favour of their being an underlying relationship between the two variables. Results of the Chi-square test of independence on the three time variables against the number of admissions resulted in *p*-values of 0.230, 0.150 and 0.452 for day, month and year respectively. It can be concluded no evidence has been found that time-dependency, in terms of the three time variables considered here, has an impact on the admission rates of hip fracture patients to the UHW.

There have been a number of studies which have investigated seasonal changes on hip fracture risk with contradictory results. Statistical investigations using ARIMA (auto-regressive integrated moving average) modelling have shown an opposing conclusion of this, where the effects of winter were shown to significantly increase the propensity of hip fractures (Lin and Xiraxagar 2006, Modarres et al. 2012).

Conversely, seasonal variability for proximal femoral fractures by sex was shown to be not significant (Papadimitropoulos et al. 1997), while no consistent seasonal pattern in the incidence of hip fracture has also been reported (Pedrazzoni et al. 1993). One explanation of an increase in hip fracture injuries during winter months is the increased risk of slipping due to icy conditions, but it has been reported by one study that, despite this, only 4% of hip fractures were attributed to the effect of season (Bulajic-Kopjar 2000). Another study found that hip fracture risk was related to slippery winter conditions amongst women aged 45-74 years old, but not for women aged 75 years and above (Jacobsen et al. 1995).

It thus follows that, while published reports are plentiful (those quoted here are just some of the examples in the literature), conclusions are inconsistent. The conclusion of no seasonal variation for admissions to the UHW is upheld.

Of course, if admissions are looked at by the hour of the day, then some time-dependency is evident; patients are more likely to be admitted during daylight hours. However, this piece of information is not available for a significant proportion of the observations here and therefore is not considered further. Additionally, with the completion of the bed occupancy calculations (Section 2.3), it was decided that the day of admission was enough information for this study.

2.5 Comparison within specialty

Trauma hip fractures are classified as a part of the Trauma and Orthopaedics (T&O) specialty. The number of beds available and the number of beds occupied within this specialty at the University of Wales are calculated by the Welsh Assembly Government (WAG 2008b). These are included here for the years 2006-2007 as comparative figures. The average daily number of beds available is defined as the average daily number of staffed beds in which inpatients are being or could be treated without any change in facilities or staff being made, while the average daily number of occupied beds is defined as the average daily number of beds occupied by inpatients under the care of a consultant in a particular specialty. The total number of admissions under this specialty and the average length of stay are also given.

Table 2.5i: Selected figures for the specialty of T&O at the UHW, 2006-2007

Measure	Value
Average daily available beds	122.9
Average daily occupied beds	122.5
Percentage occupancy	99.67
Inpatient cases	3973
Average duration of stay (days)	11.3

It can immediately be seen that this specialty is operating at almost full capacity; the average number of beds occupied each day is almost equal to the average number of beds available. In Section 2.2 the average daily bed occupancy of hip fracture patients was calculated as 42.1, which as a percentage of all occupied T&O beds is 34.4%. The number of inpatient cases admitted in the two years of 2006 and 2007 was 3973, compared with 1020 hip fracture patients in the same time period. While it appears, therefore, that hip fracture patients account for approximately one quarter of all T&O admissions, they occupy approximately one third of the available beds. Intuitively it can be expected that this is due to differing length of stay patterns, which is supported by the reported figures. All patients within this group were reported to have a mean length of stay of 11.3 days, while the calculated mean length of stay in the hip fracture sub-group was much higher at 28.7 days.

2.6 Delayed transfers of care

The National Assembly for Wales advise that "planning for hospital discharge must begin at […] admission to hospital. It should be considered as a process and not an event." (NAFW 2005). Unfortunately in some cases discharge is not planned effectively and the patient experiences a delayed transfer of care. A delayed transfer of care occurs when a hospital inpatient is ready to move on to the next stage of care but is prevented from doing so. This may happen for several reasons. In the case of hip fracture patients, it may be because the patient needs to be discharged to a care home or rehabilitation unit and there is no current place available, for example. Thus the patient is unnecessarily occupying a bed on the ward; if these delayed transfers could be eradicated, or at least reduced, there will be benefits to the system in terms of bed availability and staff workload. It has been advised that for every day delayed while a proximal femur fracture patient is waiting for a place in a rehabilitation unit, their total hospital length of stay is increased by 0.64 days (Weatherall 2001).

The *Hospital at Home* scheme provides one method of easing pressures on the healthcare system; hospital-level care by a multi-disciplinary team to patients aged 65 or over in the comfort of their own home (Health Workforce Solutions LLC 2008), and has been suggested that approximately 40% of hip fracture patients are suitable for early discharge to a scheme such as this, leading to substantial savings in direct costs of care (Hollingworth et al. 1993).

2.6.1 Occurrence

Unfortunately there is no information available from the ward on whether or not this occurs. However, the Welsh Assembly Government disseminates statistics on this topic which can be used to make inferences about trauma hip fracture patients. These statistics are categorised by facility in terms of whether or not treatment for mental health is included, where facilities excluding mental health are classified as acute and community hospitals, rehabilitation centres and other facilities, so this group is looked at here. Unfortunately no other information on the type of treatment is provided. However, results by health board and local authority are available. The month of October 2009 is considered in greater detail here; these results are actually reported on a month-by-month basis but it is difficult to collate results. This is census data, not a retrospective analysis, meaning that results reported for October

2009 is likely to include many of the same people as months before and after October. It was found that results for this month were not too dissimilar from other months.

84 delayed transfers of care were reported this time period, a breakdown of reasons and the percentage of time this occurred is given in Table 2.6.1i (WAG 2010b). Another statistic of interest is the rate of delay by age. For October 2009 for patients aged 75 and over, the rates for Cardiff and The Vale of Glamorgan were 22.7 and 35.4 respectively, per 10,000 population in this age group (WAG 2010c). 82.1% of the 84 patients were waiting to leave the NHS while 10.7% were waiting for assessment or a move within the NHS, while this was unknown for the remainder of patients (WAG 2010d).

Table 2.6.1i: Reasons for a delayed transfer of care, Cardiff and Vale UHB, October 2009

Reason relating to	Percentage
Healthcare - Assessment or arrangements	50.0%
Community care - Assessment or arrangements	10.7%
Care home - Waiting for availability or selection	21.4%
Other	17.9%

The length of delay is also reported (WAG 2010e). Detailed information is unavailable but instead the numbers in certain ranges are reported and results are displayed in Table 2.6.1ii Note that the figures reported are from a census and therefore do not necessarily represent the final delay experienced. It can be seen that some patients had been waiting more than six months at the time of the census.

Table 2.6.1ii: Length of delay caused by a delayed transfer of care, Cardiff and Vale UHB, October 2009

Time interval (weeks)	0-3	3-6	6-12	12-26	26+
Percentage	32.1%	20.2%	26.2%	20.2%	1.2%

Taking the middle value of each range (and the lower bound of 26 for the '26 + weeks' group), some estimate of the total time attributable to delayed transfers can be made. For the

month studied, the patients waiting at that time had been waiting approximately 650 weeks in total. This is rather an alarming figure considering that with better planning and resource allocation much of this could be avoided.

2.6.2 Consideration for this study

It has been shown that around one third of delayed transfers of care occur due to reasons related to community care or care homes. Many patients presenting to the hospital with a fractured hip have an admission source of a care home or another residential care facility, but any patient admitted from these sources is assumed to not have this problem since they would be able to return to their original residency. This would therefore only be an issue for those patients who, for example, are admitted from home but are discharged to a care home. These figures are presented and discussed further in Chapter 6, where it can be seen that this group of patients is relatively small in size. While the occurrence of delayed transfers of care is rather high, it may be concluded that the proportion of these patients who are likely to be hip fracture patients could indeed be comparatively small. Many patients are discharged to a rehabilitation unit and thus could fall into the healthcare category shown in Table 2.6.1i. This is a much larger group and therefore may have an impact on patient length of stay, if a delayed transfer of care occurs. Additionally rates for the local authorities of Cardiff and The Vale of Glamorgan show the impact upon those aged 75 or above; hip fracture patients are predominantly elderly and therefore this may be an indicator of the influence of this problem on the patient group of interest.

The most important issue to consider is the extra length of stay caused by this delay in transfer. Some insight can be gained by the values presented in Figure 2.6.1ii, but the total delay incurred is unknown, as previously explained. Additionally, there is no way of knowing how many trauma hip fracture patients are in fact affected by this problem, even gaining a reasonably accurate estimate would be difficult.

Due to these difficulties, it was decided that this issue would not be investigated further for these patients. While it may be useful to consider in the future if information and data became available, currently it is simply too difficult to include this with any degree of accuracy.

2.7　Chapter summary

One of the data sources available for this project has been introduced, namely the Cardiff Hip Fracture Survey. Since this captures any patient who is admitted to the hospital of interest, and indeed the University Health Board, with a trauma hip fracture, then it can be stated that a comprehensive database of these patients has been compiled and studied. Of course, if patients have been omitted for any reason, whether purposefully or not, then the database will clearly not have all patients admitted over the relevant time period. However, the data must be taken at face-value and be assumed to be as accurate as possible. Routine validation checks were performed on the data before and during any further investigations were undertaken. Analyses were also presented to relevant members of the medical team, who confirmed that the results were as expected.

The typical patient profile of a trauma hip fracture patient has been presented. These patients are usually elderly and approximately three in every four patients is female. These figures are approximately in line with national data (see Chapter 1). Due to the high age profile of these patients, many of the other results displayed are relatively predictable. These include, but are not limited to: poor walking ability, diminished mental health and high ASA grades.

Additionally some information on the injury, treatment and patient stay has been given. A wide variety of fracture and operation types has been seen, showing the diversity of injury and treatment within this patient group. The variation in patient length of stay has been mentioned and the large fluctuations in bed occupancy have also been demonstrated.

The effect that this has on the specialty to which hip fracture belongs, Trauma and Orthopaedics, is evident. This specialty is clearly under great demand at the hospital under consideration, which is no doubt exacerbated by the lengthy time that hip fracture patients spend in the care of the hospital. Any measures to reduce length of stay, thus weakening the intense demands on this over-burdened system, would undoubtedly be welcomed.

The issue of delayed transfers of care was raised and discussed, but the final conclusion was to not consider this matter further. The proportion of affected patients and the extent of an increase upon length of stay are not accurately estimable by the current available data.

CHAPTER 3: LENGTH OF STAY ANALYSIS

3.1   Introduction

The primary aim of this chapter is to investigate which factors, if any, influence the acute length of stay for trauma hip fracture patients. Successful identification of these factors is not only interesting from a clinical and statistical viewpoint, but also may aid the hip fracture team with planning and care of these patients. There are several benefits of being able to better estimate the length of stay of a patient, arguably the most important of these being the planning of their discharge. This information would of course also be useful to the patient, their family and/or caregivers. Methods used to investigate factors influencing length of stay include linear regression and classification and regression trees (CART). These techniques and the results obtained from them are explained forthwith.

The observations given by the data capture used for these analyses are equivalent to those introduced in Chapter 2.

Results are compared with the literature, where appropriate, and an overview of other reported findings subsequently given.

3.2 Linear regression

3.2.1 Introduction

Linear regression is a statistical technique used to model the relationship between a dependent variable, y, and one or more independent variables, $X_1, X_2, ..., X_p$. The unknown parameters of the independent variables are estimated from existing data.

The resultant model takes the form $y = \beta X + \varepsilon$, where β is the p-dimensional parameter vector, X is the design matrix of regressors and ε is the vector of error terms. A constant term is usually included as one of the regressors, giving the intercept of the predictor equation.

There are two main motives for using linear regression. Firstly, this method can be used to fit a predictive model to an existing dataset. If the additional values of X_i, $i = 1, ..., p$ for a new observation are known, then the fitted model can be used to obtain a prediction for y for this observation. Secondly, linear regression can be used to analyse and quantify the strength of a relationship between the X_i and y. Another useful result of this assessment is often which of the independent variables have no relationship with y.

Linear regression is used in an attempt to find predictors of patient length of stay (LoS) and to assess which of these factors are most important, or which are not important at all. The initial factors entered into the model are the patient factors that are known on arrival of the patient or soon after the patient arrives. If patient length of stay could be predicted soon after the patient arrives, then it could help with treatment planning of the patient as well as capacity planning of the ward.

Many of the variables used here were introduced in Chapter 2 when an overview of some patient characteristics was given. They are now listed in Table 3.2.1i, refer also to Table B3.2.1a in the Appendix for more detailed information, particularly with regard to nominal variables such as *admfrom*, which needed to be recategorised as several binary variables. Other information was also recorded but data was inaccurate or mostly incomplete and was therefore excluded. *ASAnew_n* refers to the new classification of ASA grade described in Section 2.2.4, where patients with an ASA grade of I or II are combined into a new single

group ("ASA grade I&II") and grade V patients are excluded. Grades III and IV are still treated separately.

Those variables which relate to patient condition and treatment are the 15 variables that comprise the first two sections in this table, plus operative delay, while *wardlos* is the dependent variable here. Unfortunately variables in the final section were largely incomplete and thus no further investigations or analyses are completed. If more data were to become available in the future, it would certainly be an area which could expand on the work completed here. Discharge destinations are discussed in more detail in Chapters 4 and 6.

However, three variables were dropped from later analysis; *livealon* and *admdelay* were dropped from the variables collected by the hip fracture team during the study period. While some regression methods can handle missing values, the regression procedure used here requires complete information for all observations and thus the inclusion of these variables would have considerably reduced the size of the dataset available. It was decided to exclude *pathfrac* on the basis that over 95% of patients had the same value recorded for this variable (no pathological fracture), which left very few patients in the other groups, rendering analysis difficult. Patients who do not undergo an operation are also excluded, since they have no entry for operative delay. Operative delay here is a binary variable, patients are categorised according to whether they undergo surgery within two calendar days or not. The reason for this is discussed in more detail in various sections of this thesis.

The computer package SAS 9.1.3 (SAS 2002-2003) was utilised for this analysis. This is a powerful code-based statistical program which can handle vast amounts of data very quickly and accurately and incorporates a very substantial catalogue of statistical procedures.

Table 3.2.1i: Variables available from the Cardiff Hip Fracture Survey dataset

Category	Variables [*variable name*]
Patient / Admission	- Patient lives alone or in institutional care [*livealon*] - Place admitted from [*admfrom*] - Walking ability pre-fracture [*walking0*] - Walking aids used pre-fracture [*walkaid0*] - Mobility score pre-fracture [*mobility*] - Mental state on admission [*mentalst*] - WAASP (Weight, Appetite, Ability to eat, Stress factors, Pressure sores/wounds) score on admission [*WAASP*] - Age [*age*] - Sex [*sexM*] - Delay between fracture and admission (days) [*admdelay*]
Medical diagnosis	- Side of fracture [*side*] - Type of fracture [*fractype*] - Pathological fracture diagnosis [*pathfrac*] - Operation type performed [*optypenew_n*] - ASA (American Society of Anesthesiologists) grade [*ASAnew_n*]
Hospital Stay / Discharge	- Operative delay [*opdelay*] - Acute ward length of stay (days) [*wardlos*] - Death on acute ward [*survival_ac*] - Acute discharge destination [*acdisto*] - Rehabilitation placement [*rehab*] - University Health Board length of stay (days) [*finlos*] - Death in University Health Board [*survival_fin*] - Final discharge destination [*findisto*]
Follow-up	- Residency at 120 days [*resid120*] - Walking ability at 120 days [*walking120*] - Walking aids used at 120 days [*walkaid120*] - Hip pain at 4 months [*hippain4*]

3.2.2 Assumptions

A number of assumptions must be satisfied before the results of a linear regression analysis can be interpreted.

The regressors X_i are all assumed to be error-free in terms of measurement. Here it must therefore be assumed that the recorded values in the dataset used for this analysis are accurate. Where an obvious error had been entered (a negative age, for example), the value was removed.

Multicollinearity exists when there is a strong correlation between two or more predictor variables in a regression model. Perfect collinearity exists when at least one predictor variable is a perfect linear combination of the others, the simplest example of which being that two predictor variables have perfect correlation. The computer package used tests the assumption that each predictor variable is linearly independent from every other predictor variable and alerts the user to any linear combinations between the variables entered into the model. No such combinations were found and it can thus be concluded that there is no perfect collinearity in the model. Less than perfect collinearity however is virtually unavoidable, but low levels pose little threat to the models generated here since the methodology used to formulate them is generally robust enough to tackle this issue. It is important to investigate collinearity in any regression model, since while results may seemingly appear to be of good quality, as collinearity increases there are several problems which may arise, including untrustworthy parameter estimates, a limitation on the size of R^2 and difficultly in assessing the relative importance of predictors (Field 2009). Results relating to multicollinearity are discussed further in Section 3.2.3.

The final assumption relates to the error terms, those which describe the natural disturbance in the model. These residuals, simply the difference between the observed and the predicted values, must be Normally distributed with a mean of zero and constant variance. This assumption can be tested by inspection of a plot of the predicted values against the residual values. Evidence that this assumption is satisfied is also given in the next section.

3.2.3 Fitting the model

All scale and ordinal variables were standardised before performing the regression procedure for ease of interpretation. A total of 193 observations, from the complete dataset of 2182 observations, were excluded due to incomplete data.

Stepwise regression was employed within SAS here, a method in which the choice of explanatory variables is carried out by an automated procedure. Stepwise regression combines both forwards selection and backwards elimination; at each stage of the procedure testing takes place to assess the inclusion and exclusion of variables. In both cases the significance threshold was set to 5%. This is an appropriate tool since stepwise regression may be used in the exploratory phase of investigation but is not recommended for theory testing (Menard 2001), which involves the testing of a-priori hypotheses or theories relating to relationships between the variables. It is not the case here that there are any a-priori assumptions regarding the relationships between the variables but that it is the goal to discover the existence and strength of any such relationships.

Initially the requirement of random errors was not fulfilled and therefore length of stay was transformed by taking the natural logarithm of each value.

Expanding on the equation given previously, the linear regression equation then becomes

$$\ln(LoS) = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon$$

Note that an intercept, β_0, is included here.

It is important to bear in mind that the variable of interest here is length of stay, not the natural logarithm of this value. For each parameter X_i, $i = 1, \ldots, p$, the multiplicative factor is not the corresponding β_i value, but the exponent of this.

Therefore

$$LoS = e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon}$$

or

$$LoS = e^{\beta_0} e^{\beta_1 X_1} \ldots e^{\beta_p X_p} e^{\varepsilon}$$

The assumption of random errors can be seen to be satisfied in Figure 3.2.3i. A clear random scatter of points can be seen, which is also symmetrical about the line *Residual* = 0, as required.
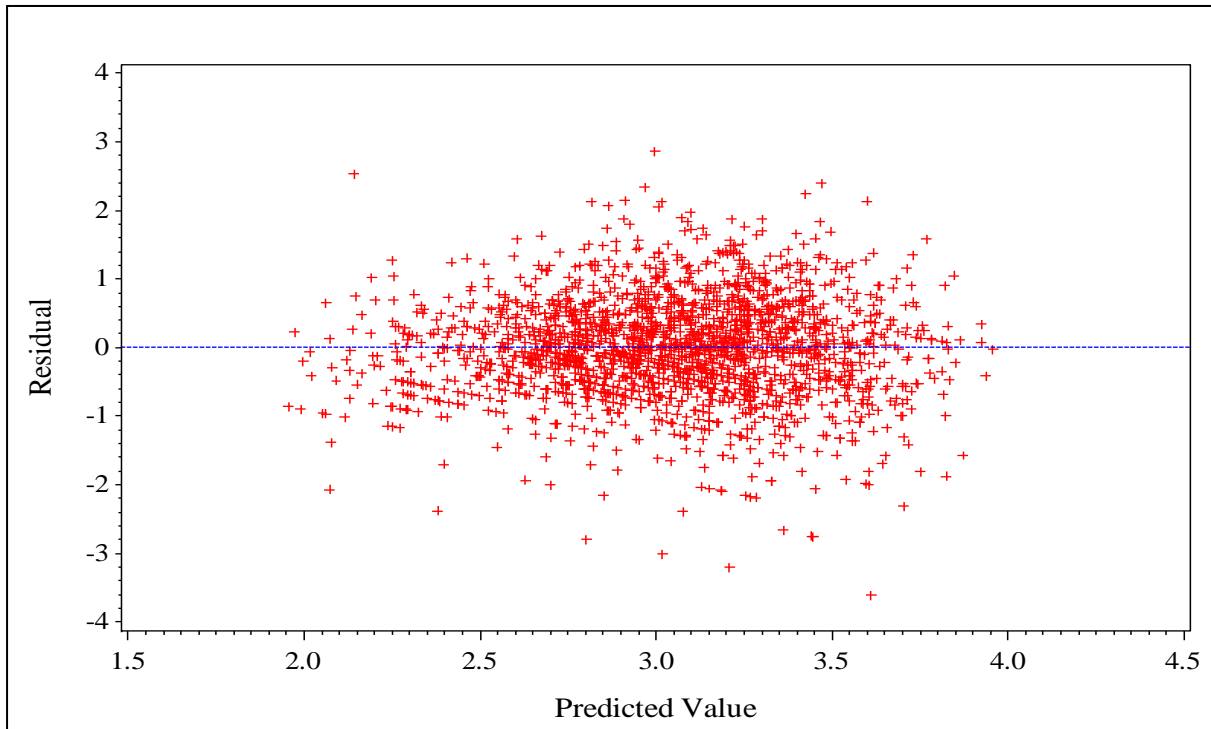
Figure 3.2.3i: Plot of *Predicted Value* against *Residual* for the linear regression model

Figure 3.2.3ii shows the observed value of log length of stay against the predicted value. There does appear to be some positive trend here which suggests that a linear regression model could be viable. A perfect straight line here would indicate a perfect fit by the model. While this is not achieved and there is arguably some considerable deviation from this, the results of a regression can still be useful.

The first important values to note in a linear regression analysis are the ANOVA results. It was indicated here by the *F*-test that the regression model fitted was significant, $p < 0.0001$; that is, the vector of parameter coefficients is significantly different from zero. This result indicates that the observed value for R^2 is reliable and not a spurious result of oddities in the data (Sweet and Grace-Martin 2008). The parameter estimates given by the linear regression procedure are displayed in Table 3.2.3iii.
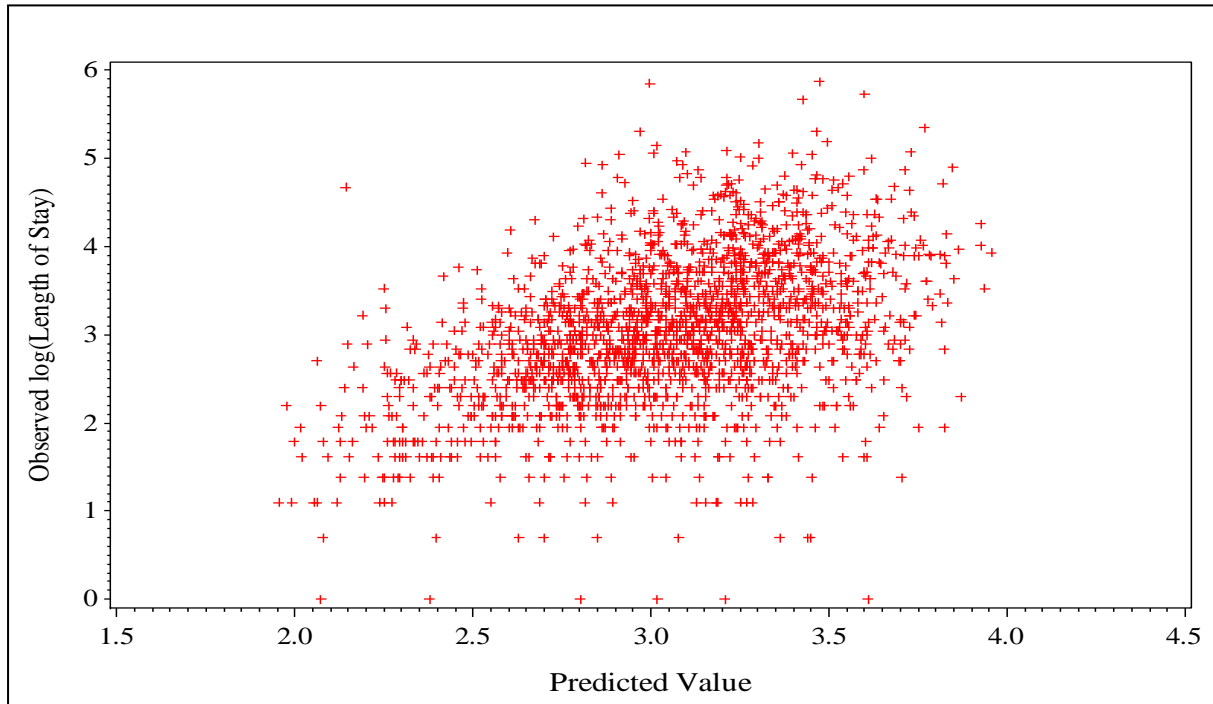
Figure 3.2.3ii: *Predicted Value* against *Observed* for the linear regression model

Table 3.2.3iii: Parameter estimates given by the multivariate linear regression model

Variable	Parameter estimate	Standard error	*F*-value	Pr $>F$
Intercept	3.0466	0.0414	5418.03	< 0.0001
Admfrom_d1	-0.1368	0.0435	9.91	0.0017
Admfrom_d4	-0.5466	0.0832	43.18	< 0.0001
Admfrom_d5	-1.0448	0.1862	31.48	< 0.0001
Mobility	0.0907	0.0211	18.57	< 0.0001
Mentalst	0.0915	0.0187	23.99	< 0.0001
WAASP	0.0555	0.0184	9.11	0.0026
Opdelay	0.2885	0.0337	73.42	< 0.0001
Age	0.1097	0.0185	35.14	< 0.0001
SexM	0.0847	0.0381	4.95	0.0262
Optypenew_d3	-0.2011	0.0702	8.21	0.0042
Optypenew_d6	-0.2408	0.0901	7.14	0.0076
Fractype_d1	-0.1618	0.0719	5.07	0.0245
Fractype_d5	0.1586	0.0444	12.74	0.0004
Fractype_d6	0.1692	0.0700	5.85	0.0157

It can be seen that all parameter estimates are highly significant at the 5% level. These results yield the following equation:

$$\ln(LoS) = 3.05 - 0.14(Admfrom\_d1) - 0.55(Admfrom\_d4) - 1.04(Admfrom\_d5)$$
$$+ 0.09(Mobility) + 0.09(Mentalst) + 0.06(WAASP) + 0.29(Opdelay)$$
$$+ 0.11(Age) + 0.08(SexM) - 0.20(Optypenew\_d3)$$
$$- 0.24(Optypenew\_d6) - 0.16(Fractype\_d1) + 0.16(Fractype\_d5)$$
$$+ 0.17(Fractype\_d6)$$

Writing this in terms of length of stay, rather than the natural logarithm of length of stay, generates the following:

$$LoS = \frac{e^{\substack{3.05 + 0.09(Mobility) + 0.09(Mentalst) + 0.06(WAASP) + 0.29(Opdelay) \\ + 0.11(Age) + 0.08(SexM) + 0.16(Fractype\_d5) + 0.17(Fractype\_d6)}}}{e^{\substack{0.14(Admfrom\_d1) + 0.55(Admfrom\_d4) + 1.04(Admfrom\_d5) \\ + 0.20(Optypenew\_d3) + 0.24(Optypenew\_d6) + 0.16(Fractype\_d1)}}}$$

The issue of multicollinearity was introduced in the previous section. One way to identify multicollinearity is to scan a correlation matrix of all the predictor variables for any high correlations, but this may miss some of the more subtle forms of multicollinearity (Field 2009) and thus this method is not used here.

Instead, some collinearity diagnostics which can be inputted into the SAS code are used, namely the variance inflation factor (VIF) and the tolerance statistic, which is the reciprocal of VIF. The tolerance is the proportion of variance attributable to a given predictor variable which is not explained by all of the other predictors, while the VIF represents the factor by which the variances of the estimated coefficient parameter is multiplied due to any multicollinearity contained within the model. It is recommended that if the largest VIF is greater than 10 then there is cause for concern and further investigation is required (Kutner et al. 2004, Myers 1990), while if the average VIF value is substantially greater than 1 then the regression may be biased (Bowerman and O'Connell 1990). It has also been suggested that tolerances below 0.2 indicate a potential problem with collinearity (Menard 2001); or equivalently, that VIF values above 5 may prove problematic.

The average VIF value was 1.30 indicating that collinearity is not a problem here by this criterion, while all VIF values were below 2. The minimum tolerance statistic was found to

be 0.51 (maximum VIF 1.97), again indicating there is not enough collinearity in the model for results to be unreliable. A full breakdown is given in Table D3.2.3a.

The R^2 value for this model is 0.1965, meaning that 19.65% of variation in the original data can be accounted for by the model. While a higher R^2 value would be desirable, the conclusions that can be drawn here are helpful nonetheless and indeed highlight the difficulty of this area.

Variables which are not found to be significant predictors of length of stay include the side of the body on which the fracture was incurred, walking aids and walking ability on admission. Operation type was found to be a significant predictor in two cases, namely for screws and total hip arthroplasty, where patients in these groups could expect a shorter length of stay. An undisplaced intracapsular fracture type was also seen to be negatively associated with length of stay (that is, be associated with a shorter length of stay), while the fracture types trochanteric (multi fragment) and subtrochanteric are shown to be associated with longer length of stay. The place a patient was admitted from was also seen to be a significant predictor in three out of eight cases; patients admitted from home, a nursing home or from continuous care as a permanent acute hospital inpatient were found to have a shorter length of stay.

3.2.4 Mortality criteria

An interesting omission from the results produced by the multivariate linear regression is that ASA grade was not found to be a significant predictor of length of stay. It is certainly expected that a patient's length of stay would be influenced by their medical fitness, as sicker patients are more likely to require more extensive and prolonged treatment. However, on further reflection this result could be explained by a number of influencing factors including the obvious observation that sicker patients are more likely to die while in hospital, thus their length of stay concludes at death and not at the end of any prescribed medical treatment. Those patients who die pre-operation or are given conservative treatment are already excluded since they do not have an entry for delay to operation, but there are many who die post-operation. Note that the modelling presented here thus prioritises length of stay as a surrogate measure for post-operative morbidity rather than for capacity/resource planning.

To account for this, the regression procedure was repeated using the subset of patients who survived their acute stay in hospital. Error assumptions were checked as previously and the natural logarithm of length of stay was used as the dependent variable. The same stepwise regression options were also used. In this case, 1770 observations were available and results are given in Table 3.2.4i.

Table 3.2.4i: Parameter estimates given by the multivariate linear regression model, surviving patients only

Variable	Parameter estimate	Standard error	F-value	Pr $> F$
Intercept	3.0399	0.0302	10104.1	< 0.0001
Admfrom_d4	-0.4428	0.0777	32.51	< 0.0001
Admfrom_d5	-1.0672	0.1759	36.80	< 0.0001
Admfrom_d7	0.2363	0.0837	7.98	0.0048
Mobility	0.1023	0.0198	26.64	< 0.0001
Mentalst	0.1219	0.0185	43.31	< 0.0001
WAASP	0.0630	0.0182	11.93	0.0006
Opdelay	0.2496	0.0332	56.59	< 0.0001
Age	0.1193	0.0177	45.72	< 0.0001
SexM	0.1158	0.0378	9.38	0.0022
ASAnew_n	0.0710	0.0181	15.40	< 0.0001
Optypenew_d2	-0.0917	0.0372	6.09	0.0137
Optypenew_d3	-0.2504	0.0674	13.82	0.0002
Optypenew_d6	-0.2865	0.0849	11.38	0.0008
Fractype_d1	-0.2061	0.0681	9.17	0.0025
Fractype_d5	0.1607	0.0448	12.89	0.0003

The average VIF value across these variables was 1.29, with all but one VIF values below 2. The minimum tolerance was 0.49, so again collinearity has not proved to be a problem for the regression model. Tolerance and VIF values are given in Table D3.2.4a of the Appendix. The R^2 value was found to be 0.2609.

Many of the original variables are kept in this second model, but some new inclusions and exclusions can be seen. An increasing ASA grade is now shown to be associated with a

longer acute hospital stay, as is having an acute hospital as their admission source. Dynamic hip screw (*optypenew_d2*) is a new inclusion here, while subtrochanteric fractures are now excluded (*fractype_d6*).

3.2.5   Analysis by ASA grade

The final investigation undertaken in this section performs the same technique but on three distinct subsets of the data created by splitting on ASA grade. Performing this split will go some way to resolving the issues discussed in the previous section. Furthermore, due to other reasoning considered previously, only those patients who survive their acute hospital stay are included. 711, 928 and 131 observations were available for ASA grades I&II, III and IV respectively. In all cases assumptions were checked and satisfied as previously described. It was also again found that there was low multicollinearity between the predictor variables and certainly not enough to cause any problems with interpretation of the results.

Table 3.2.5i: Parameter estimates given by the multivariate linear regression model, surviving patients only, ASA grade I&II

Variable	Parameter estimate	Standard error	F-value	$\Pr > F$
Intercept	2.9344	0.0416	4974.68	< 0.0001
Admfrom_d7	0.4614	0.1489	9.60	0.0020
Mentalst	0.1634	0.0301	29.48	< 0.0001
WAASP	0.0862	0.0269	10.23	0.0014
Opdelay	0.2532	0.0505	25.19	< 0.0001
Age	0.1528	0.0226	45.52	< 0.0001
Optypenew_d5	0.3212	0.0726	19.55	< 0.0001
Fractype_d1	-0.5189	0.0761	46.50	< 0.0001
Fractype_d2	-0.2657	0.0688	14.94	0.0001

For ASA grades I&II, the only operation type found to be significantly associated with length of stay was hemiarthroplasty, which is a very interesting result since this is a large proportion of all patients. Other interesting results include the use of delay to operation in this model, which was found to be associated with an increase length of stay. The R^2 value was 0.3091.

Results for ASA grade III appear to be much more influenced by admission source, with four different admission sources included, namely residential care, nursing home, permanent hospital inpatient and acute hospital; see Table 3.2.5ii. Delay to operation was again found to be related to an increased length of stay, as were increased scores for mobility, walking ability on admission score (relating to poorer walking ability) and mental state. Additionally this is the only ASA grade where sex was found to be a significant predictor, with male patients expected to incur a longer stay in hospital. The R^2 value here was 0.1692.

Table 3.2.5ii: Parameter estimates given by the multivariate linear regression model, surviving patients only, ASA grade III

Variable	Parameter estimate	Standard error	*F*-value	Pr $> F$
Intercept	3.0069	0.0409	5395.47	< 0.0001
Admfrom_d3	0.2125	0.0719	8.74	0.0032
Admfrom_d4	-0.4379	0.1003	19.08	< 0.0001
Admfrom_d5	-0.9179	0.2244	16.73	< 0.0001
Admfrom_d7	0.2453	0.1173	4.38	0.0367
Walking0	0.0953	0.0268	12.62	0.0004
Mentalst	0.1152	0.0241	22.78	< 0.0001
Opdelay	0.2325	0.0465	25.03	< 0.0001
Age	0.0821	0.0276	8.84	0.0030
SexM	0.1404	0.0536	6.86	0.0090
Optypenew_d3	-0.3045	0.0773	15.52	< 0.0001
Optypenew_d6	-0.3422	0.1458	5.51	0.0191
Fractype_d5	0.1308	0.0588	4.95	0.0263

A reduced model, in terms of the number of included parameters, was found for ASA grade IV, see Table 3.2.5iii. The only admission source which was found to be negatively associated with length of stay was permanent hospital inpatient, which may seem rather unintuitive, particularly as it is known that these patients survive their acute stay in hospital. One explanation could be that as these patients have more severe medical problems, they are moved elsewhere in order to primarily deal with these comorbidities. An undisplaced intracapsular fracture was also found to relate to a shorter length of stay, while a higher

mobility score and a delay to operation of more than two days was found to be associated with an increased length of stay in this subgroup. A moderate R^2 value of 0.1802 was given.

Table 3.2.5iii: Parameter estimates given by the multivariate linear regression model, surviving patients only, ASA grade IV

Variable	Parameter estimate	Standard error	*F*-value	Pr $> F$
Intercept	3.1429	0.1017	955.23	< 0.0001
Admfrom_d5	-1.0404	0.4068	6.54	0.0117
Mentalst	0.1349	0.0582	5.35	0.0224
Opdelay	0.4213	0.1242	11.50	0.0009
Fractype_d1	-0.4505	0.1940	5.39	0.0218

3.2.6 Conclusion

One of the most useful results gained from this analysis is that a delay to operation of more than two days is shown to be associated with a longer length of stay. Of course it must be considered that this increase in length of stay is in fact the delay itself, at least to some extent; the dependent variable (LoS) is comprised partly of delay, but this does not lead to too much cause for concern. Delay is included as a binary variable, defined as whether a patient is operated on within two days or not. Delayed patients spend, on average, 7.8 days more in hospital, compared with non-delayed patients. The difference in delay is considerably less, 4.9 days between the groups (6.2 days and 1.3 days on average, for delayed and non-delayed patients respectively). An increased mental state score, indicating diminished mental capacity, was also shown to be an important predictor in each case. Operation type also featured heavily in these analyses, with varying degrees of level and magnitude, as did admission source.

The models presented were all found to be significant by the *F*-test, $p < 0.0001$. Despite this, particularly outstanding R^2 values were not found, despite several regression methods being attempted before the final results presented here were determined. However, this will always be expected in the highly variable domain of healthcare, in particular with the rather long hospital stays experienced by this patient group.

3.3   CART analysis

3.3.1   Introduction

Classification and regression trees (CART) analysis is a widely-used data mining technique that can be utilised to help understand large volumes of data. The process begins with the entire set of data and is split into two or more subsets until an appropriate level of homogeneity is reached. A target variable is specified, on which the split decisions are made, according to the values given by numerous independent variables, until homogenous groups are established with respect to the values of the target variable. The outcome of CART can be conceptualised by a tree structure. All observations begin at the 'root', which is the initial node, and then pass through the tree based on the values they hold at each splitting decision, ultimately finishing at a terminal node, or 'leaf'.

The program TreeWorks was used to undertake CART analysis here. More information on this software can be found in the accompanying paper (Harper and Leite Jr 2008), where information on its quality and suitability for healthcare data is also discussed. Initially the data is randomly split into two subsets, a learning sample and a testing sample. The learning sample is used to create the original splitting rules and then the testing sample is passed through the model and used to validate the nodes created. 70% of the dataset is typically used as the learning sample and the remaining 30% as the testing sample. It should also be noted that TreeWorks requires a complete dataset; any observations with missing values are removed.

While TreeWorks provides the option of splitting nodes manually, the entire process was undertaken free from intervention to avoid any bias. Other options include the proportion of observations used as the learning sample (the default value of 70% was used here) and the maximum number of levels given by the final tree. The minimum number of observations in each of the final nodes is another option provided.

3.3.2   Results

Only patients who underwent an operation were included in the final dataset, since it was important to include operative delay as a variable for this analysis. This criteria and the

requirement of a complete dataset meant that there were 1980 observations remaining to use for CART; 1386 and 594 in the learning sample and the testing sample respectively.

Outputs of the program include the percentage reduction in variance (RiV) achieved and a test of the validity of each node. By altering the aforementioned input parameters, the impact on these outcomes can be assessed. While increasing the number of permitted levels and/or decreasing the minimum number of observations per node may result in a better RiV value, it can lead to a very large number of nodes and results may become uninterpretable. Many combinations were tried and the final inputs set to be a minimum of 40 observations per node, which resulted in seven levels. This gave 24 final nodes and an RiV value of 16%; the overall variation in the data has been reduced by 16% by splitting the data into groups with increased homogeneity. Three of the nodes failed validation at the 5% significance level, but this was only marginal. Results are displayed in Figure 3.3.2i and the splitting criteria used to determine the nodes of the tree are given in Table D3.3.2a. Final nodes are denoted by an 'F' after the node number.

Figures within each node relate to the node number in bold on the top line, the mean and standard deviation of length of stay in days of the observations within that node (mean | standard deviation) on the second line and the number of observations in that node, *n*, on the bottom line.
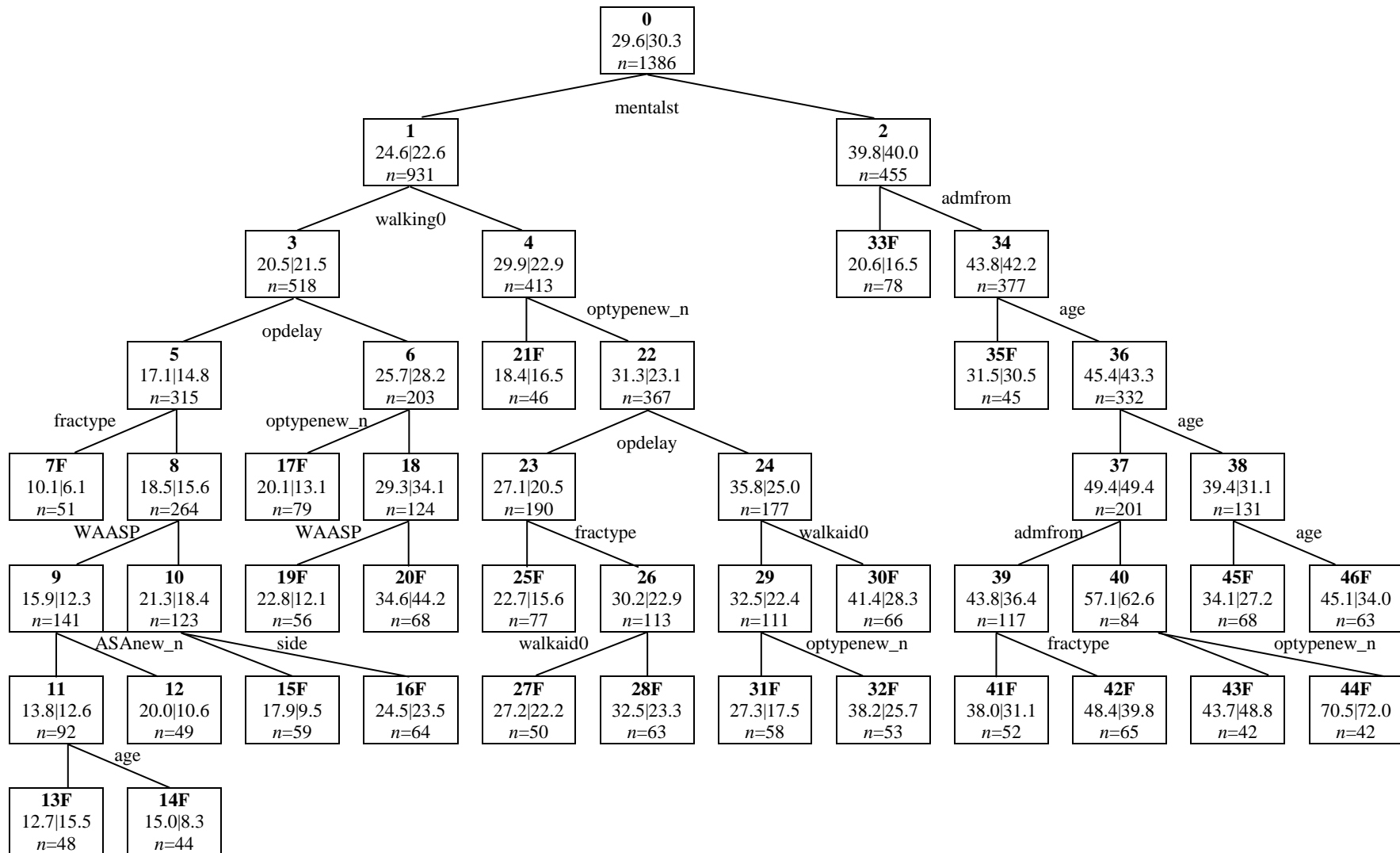
Figure 3.3.2i: CART results for length of stay

The most surprising inclusion in the results is the variable *side*, which denotes whether the injury was incurred on the left- or the right-hand side of the body. This clearly should have negligible to no effect on how long a patient should stay in hospital and therefore results are attributed to chance. Inspection of nodes 15F and 16F show that patients with a left-sided fracture have a shorter length of stay, which is also much more predictable; the standard deviation of length of stay is 9.5 days compared with 23.5 days for patients with a right-sided fracture, with coefficients of variation of 0.53 and 0.96 respectively. The coefficient of variation is a normalised measure of dispersion which is useful when comparing data with differing means since the result is dimensionless, and is calculated by dividing the standard deviation by the mean. There is no known medical explanation for this outcome and therefore results are attributed to chance; these nodes could have been removed manually from the tree, particularly as they are both terminating nodes, but are left in to draw attention to these potential problems.

An interesting result here is that mental state was used as the first splitting variable. This variable can take three levels; those patients classified as 'normal' (score of *1*) are grouped separately from those who were recorded to have known dementia or confusion (scores of *2* and *3* respectively). Levels *2* and *3* are not split later in the tree, suggesting that any level of mental illness is likely to result in a longer length of stay to the same degree.

The variables which are not used as splitting decisions are patient mobility score and sex. Despite mobility score not being used as a splitting variable, walking ability and walking aids used on admission are both present in the output, so ambulatory ability has been captured by these two variables. Looking at length of stay by a split on gender gives a difference in mean length of stay of less than one day, which may explain why sex is not included as a splitting variable. It could be argued that the inclusion of the majority of entered variables in the results does not highlight which variables are of most importance, but in essence it really highlights the complexity of this issue.

To determine which variables are of most importance, various strategies may be considered. Predictive factors higher up in the CART output is one way to assess the influence of that factor on length of stay, as well as considering the frequencies with which factors appear in the tree. Age features heavily in the CART output, where it was found that older patients consistently had a longer length of stay. This was used four times as a splitting variable,

three of which were consecutive decisions. Operation type is another variable which was used numerous times to split on, also used four times, as was the type of fracture, which was used three times. Of course, operation type and fracture type have more levels than many of the other variables (six each); it is thus even more telling if the CART procedure isolates just one or two of these levels. See node 31F for an example of this, where the operation type of hemiarthroplasty is isolated.

These results are interesting not only from a data mining and statistical viewpoint, but may also be useful for later study. Each patient will fit into one of the final nodes based upon their characteristics, so these nodes may then be used to model the variable under interest, ward length of stay. Distributional fits were found for each final node, reaffirming that this would be a suitable splitting mechanism for modelling purposes. The distributional fitting software Stat::Fit was used to find analytical fits to the data (Stat::Fit 1995-2001©).

Length of stay was found to fit a Lognormal distribution for the majority of nodes, while the remainder were found to follow a Gamma distribution. Minimum thresholds were set to zero. Results are given in Tables 3.3.2ii and 3.3.2iii, along with comparative measures for the first and second moments. These values may differ from those given in Figure 3.3.2i, since they were calculated using observations within those nodes. Recall that this tree was based a sample of the dataset, while for the following fits all observations were available.

These results all show good fits for the first and second moments, and all passed statistical goodness-of-fit tests at the 95% level of significance. Indeed, the majority of these fits would still to be found to be significant at the more stringent significance level of 99%.

Table 3.3.2ii: Lognormal fits for length of stay (days) for the final nodes resulting from the CART procedure

Node	μ	σ	Theoretical		Empirical	
			Mean	**S.D.**	**Mean**	**S.D.**
7F	2.143	0.614	10.3	7.0	10.1	6.0
13F	2.281	0.704	12.6	10.1	12.6	12.4
14F	2.647	0.561	16.5	10.0	16.5	10.4
15F	2.775	0.534	18.5	10.6	18.3	9.4
16F	2.926	0.659	23.2	17.1	23.5	20.7
17F	2.826	0.686	21.3	16.5	21.0	15.6
19F	3.012	0.511	23.2	12.6	23.0	11.9
20F	3.189	0.717	31.4	25.7	33.1	39.6
21F	2.681	0.753	19.4	17.0	19.9	21.1
25F	2.921	0.757	24.7	21.7	24.1	17.9
27F	3.037	0.801	28.7	27.2	29.2	29.0
28F	3.234	0.714	32.8	26.7	32.3	23.5
30F	3.419	0.763	40.9	36.3	40.2	31.1
31F	3.153	0.616	28.3	19.2	27.9	16.9
32F	3.313	0.666	34.3	25.6	33.8	23.0
33F	2.855	0.771	23.4	21.1	24.1	25.6
41F	3.261	0.798	35.9	33.9	35.4	32.2
44F	3.683	0.834	56.3	56.4	58.6	66.7
45F	3.309	0.857	39.5	41.1	37.8	30.2

Table 3.3.2iii: Gamma fits for length of stay (days) for the final nodes resulting from the CART procedure

Node	α	β	Theoretical		Empirical	
			Mean	**S.D.**	**Mean**	**S.D.**
12F	3.594	5.570	20.0	10.6	20.0	10.6
35F	1.457	26.461	38.6	31.9	38.6	34.4
42F	1.564	29.545	46.2	37.0	46.2	37.8
43F	1.465	29.466	43.2	35.7	43.2	47.8
46F	1.879	23.401	44.0	32.1	44.0	31.7

3.3.3 Conclusion

The CART analysis undertaken here can be deemed to be a success in terms of creating groups of observations which are homogenous in terms of the target variable, acute hospital length of stay. Despite the majority of variables being used in the splitting criteria, useful information can be gained in terms of the intricacies of the split specifications and the frequency at which various variables appear. A particular focus of this thesis is investigating the effect of delay to operation on length of stay and it is noteworthy to comment that this variable does appear more than once in the CART output.

This analysis can also be deemed a success in terms of its suitability to be used for later investigations. Length of stay for each of the final nodes was found to statistically fit either a Lognormal or Gamma distribution which could be useful for modelling purposes.

3.4   Comparison of methods and results

Two methods have been investigated and discussed in detail regarding predictive factors of acute length of stay on the trauma hip fracture ward. Five different regression models were presented which had varying conclusions, but some included and/or excluded variables were consistent throughout these analyses; these are now also compared with the CART analysis for a number of the independent variables.

3.4.1   Mental state

The variable denoting mental state (*mentalst*) was used in each of the five regression analyses and as the first splitting variable in the results produced by the CART procedure, each time signifying that poorer mental state is an indicator of increased length of stay. Mental confusion has been shown to be more likely to be present in elderly people after major surgery (Moller et al. 1998), where very old age, mobility problems prior to surgery and a history of mental health issues all increase the chances of confusion after surgery (RCOA 2010). This variable records mental state on arrival, and so the latter characteristic here is of the most importance, but the others are still important in the context of this patient cohort. These confused patients may require an extra period of recovery and this would therefore explain these results. Summary statistics of length of stay for patients who underwent surgery by mental state are presented here in order to quantify these differences.

Table 3.4.1i: Summary statistics for length of stay (days) by mental state, patients undergoing surgery only

Mental state (*score*)	Mean	S.D.	Minimum	Maximum
Normal (*1*)	24.5	23.2	1	348
Known dementia (*2*)	38.0	40.7	0	354
Confusion (*3*)	38.1	33.3	1	201

The differences between patients with a mental state score of *1* compared with those with a score of *2* or *3* are evident. Indeed, the mean length of stay for the second group is almost exactly equal to the third group. In the CART output, mental state was only used once as a

splitting rule, to separate score *1* from scores *2* and *3*, so this is verified by these summary statistics.

The presence of chronic cognitive impairment / dementia has been shown to result in a longer hospital stay after hip fracture, as was increased age (see Section 3.4.4) and a number of other factors (Clague et al. 2002). Postoperative delirium has also been found to be amongst six predictors of longer length of stay for this patient cohort (Marinella and Markert 2009). Cognitively intact and functionally independent (see Section 3.4.7) patients have been reported to not only have statistically significant shorter lengths of stay, but also achieve better changes in mobility as measured by two functional status scores, compared with cognitively impaired and dependent patients (Hershkovitz et al. 2007). Regression analysis was used to demonstrate that the presence of neurologic impairment was positively associated with length of stay in an Italian study, where it was shown that patients with an impairment experience an additional 3.8 days length of stay (Di Monaco et al. 2003). Finally, it has been shown that length of stay was increased in patients suffering from dementia, delirium and depression (Holmes and House 2000).

3.4.2   Delay to operation

Delay to operation (*opdelay*), regarded as a binary variable indicating whether or not a patient was operated on within two days of admission or not, was the only other variable to be included in all five regression models and the CART analysis, where it appeared twice. Each time a delay to operation, under this definition, signified a longer acute length of stay. Some summary statistics and a graphical representation for this are now presented; this is not examined in more detail here but is discussed in depth in Chapter 6. The relative probabilities (i.e. by delay category) displayed in Figure 3.4.2ii show the relationship between delay and length of stay clearly. Note that this graph is curtailed for display purposes. In particular consider the first group where length of stay is between zero and ten days. Almost a quarter of non-delayed patients have a length of stay which falls into this group, but just around 10% of delayed patients spend ten days or less in hospital.

Table 3.4.2i: Summary statistics for length of stay (days) by delay category

Category	Mean	S.D.	Median	Minimum	Maximum
Delayed	33.76	32.91	24	2	354
Not delayed	25.97	25.78	18	1	289

An interesting result which can be deduced from this table is that the relative spreads within each group, given by the coefficient of variation, are approximately equal. Additionally, both of these coefficients are approximately equal to one. Often this implies that a Negative Exponential distribution will adequately fit the data, but it can be seen from Figure 3.4.2ii that this is not the case here. Consider also the values of the mean and standard deviation for each group, since the median is smaller in each case then a right-sided, or positive, skew is expected, as displayed here. The differences evident between these two groups are also supported statistically by means of a Wilcoxon test, $p < 0.0001$.
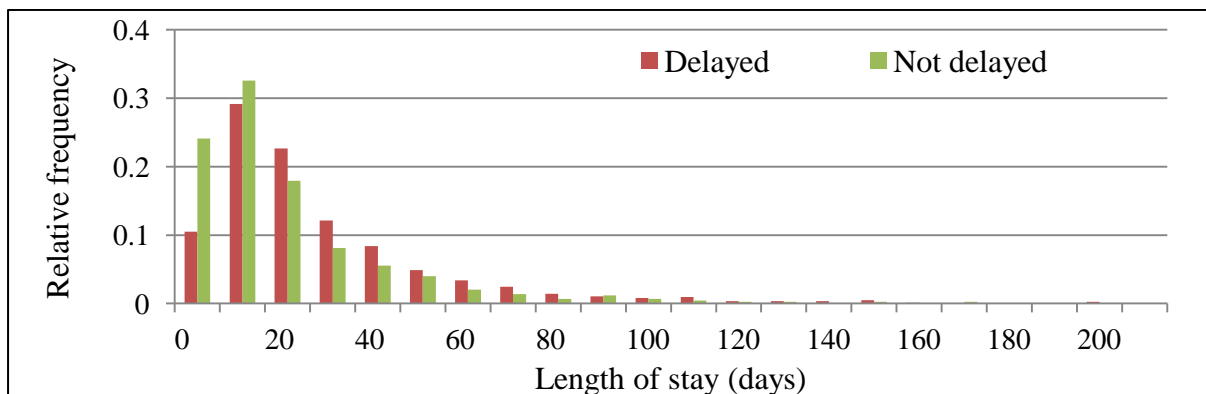
Figure 3.4.2ii: Distribution of length of stay by delay category

Results in the literature vary by delay definition and conclusion. The most commonly used definitions of a significant operative delay are 24 and 48 hours, or one or two days, though it is important to note that these are not the same. A detailed review is completed here due to the inclusion of delay as a key variable later in this thesis.

Classifying a delay as surgery after one day from admission, Verbeek et al showed that patients experienced a shorter length of stay if they were not delayed, as well having fewer postoperative complications (Verbeek et al. 2008). A significantly shorter length of stay was

also shown by another study using the same classification of delay, with means of 10.5 days compared with 12.7 days for the non-delayed and delayed groups respectively (Hommel et al. 2008b). The same conclusion was also reached by an American group, who showed a length of stay of almost two days longer for patients operated on after 24 hours. However, if post-operative length of stay only was considered, then no significant differences were found (Orosz et al. 2004). Patients delayed for longer than 24 hours due to medical reasons were compared with a matched group who were not delayed and no significant differences in length of stay post-operation were also found in an older study (Harries and Eastwood 1991).

Using a cut-off of 48 hours, it was shown that operative delay was associated with a longer post-operation stay in hospital, and that this was independent of age and comorbidities (Bergeron et al. 2006). Hoenig et al used the same cut-off and also reported a shorter length of stay for non-delayed patients (Hoenig et al. 1997), as have other studies using the same delay definition (Majumdar et al. 2006, Novack et al. 2007).

A study of 3628 patients in Peterborough also concluded that operative delay did matter, and reported a difference of mean length of stay of 21.6 days for patients operated on within 48 hours compared with 32.5 days for surgery after 48 hours, while no differences were found for lesser delays. The authors concluded that the length of hospital stay in days can be calculated as 0.1274 multiplied by the operative delay in hours. This translates to an extra day in hospital for each 7.85 hours of delay (Siegmeth et al. 2005).

Another study attempted to quantify this relationship in a similar fashion, but used post-operative length of stay instead of total length of stay. It was concluded that approximately a twofold increase in pre-operative delay increased post-operative stay by 19% (Thomas et al. 2001).

Delay to surgery was amongst four factors shown to be significant in a model which identified variables which increased time to discharge, the others being age, comorbidities and fracture type (Lefaivre et al. 2009).

A Dutch study used a cut-off of 12 hours as a significant delay, and found that there were statistical differences between the early and late groups with respect to length of stay. In this case it was shown that patients receiving surgery after 12 hours had, on average, an extra two

days in hospital compared with patients operated on within 12 hours (Rademakers et al. 2007).

However, Dolk analysed the effect of timing of surgery across several delay categories and found that there was no impact of delay upon length of stay. It was concluded that differences in hospital stay between the groups could be attributed fully to age and type of fracture (Dolk 1990). An American group also showed that the effect of delay on post-surgical length of stay is "small in magnitude", concluding that the differences in hospital stay between countries (USA and Canada) cannot be attributed to differences in operative delay (Ho et al. 2000). Parker et al also investigated differences in length of stay between hospitals, specifically between eight hospitals in East Anglia. Considerable differences in stay were found between the sites but subsequent analysis showed that operative delay was not indicative of length of stay (Parker et al. 1998).

Other studies have also reported a lack of association between time to surgery and length of stay. A study in Austria classified a wait longer than six hours as a delay. The mean difference in length of stay was just over one day between the groups (17.1 days for surgery within six hours, 18.4 days for surgery after six hours), but this difference was not significant (Dorotka et al. 2003b). Mean lengths of stay of 23.7 days and 21.5 days have been reported for delayed and non-delayed patients respectively. Here a delay was a wait longer than 24 hours and the differences were not found to be statistically dissimilar (Pathak et al. 1997).

3.4.3 Operation type

Operation type was shown to be significantly associated with length of stay in almost all cases. Consistent results include shorter lengths of stay for total hip replacements and those operations which include screws. CART separated these operations in nodes with shorter length of stay twice (17F and 21F), while all but the ASA grade IV regression results showed that these patients could expect a shorter length of stay. A total hip replacement is a more invasive procedure compared with the others and thus a patient must have an increased level of medical stability before they can be considered for this operation, which may go some way to explaining these results.

The distribution of length of stay for each of the operation types is displayed in Figure 3.4.3i. The differences in length of stay amongst the different surgical procedure groups are clearly visible, and are substantiated by a Kruskal-Wallis test, $p < 0.0001$. The whiskers displayed relate to the 1st and 99th percentiles of length of stay for each operation type.
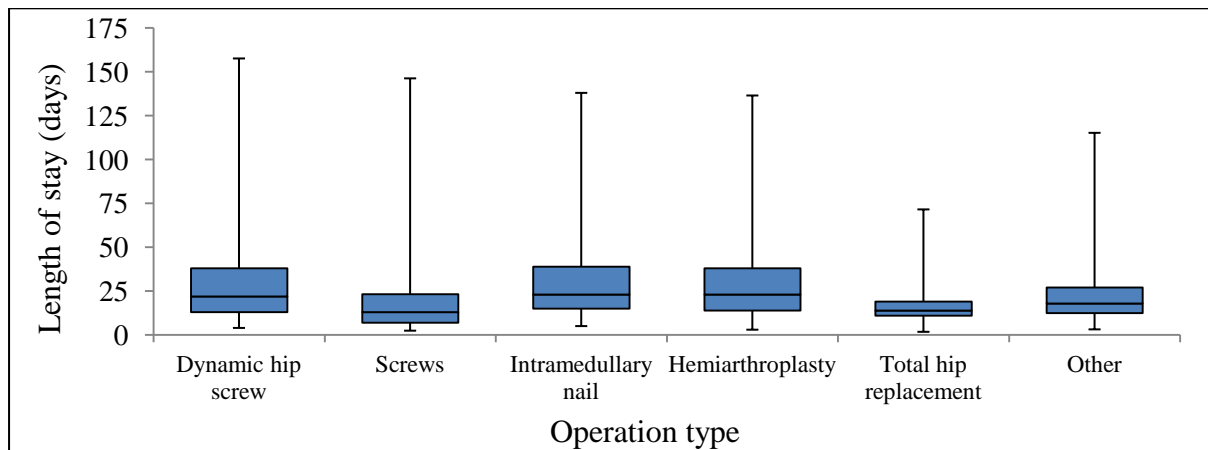
Figure 3.4.3i: Distribution of length of stay (days) by type of operation

In addition to the differences, some similarities can also be seen here. In particular, the distributions for dynamic hip screw, intramedullary nail and hemiarthroplasty appear to be alike. The shorter length of stay for the operation types of screws and total hip replacement is also evident, as found by the regression and CART analyses. The distributions appear to be positively-skewed in each case. Finally, some summary statistics are given to further compound these findings and provide additional information.

Table 3.4.3ii: Summary statistics for length of stay (days) by type of operation

Operation type	Mean	S.D.	C.V.	Median	Skew-ness	Kurtosis
Dynamic hip screw	30.41	30.47	1.00	22	4.12	28.51
Screws	20.65	25.75	1.25	13	3.75	17.54
Intramedullary nail	32.91	29.12	0.86	23	2.18	5.44
Hemiarthroplasty	31.34	29.66	0.95	23	3.97	29.26
Total hip replacement	17.61	15.19	0.86	14	3.07	12.69
Other	27.84	30.69	1.10	18	2.26	4.72

The results for coefficient of variation (C.V.) are particularly interesting here and highlight some differences where previously it appeared that certain operation types had very similar length of stay distributions. Half of the operation types have a greater standard deviation of length of stay in comparison to the mean (and therefore a coefficient of variation greater than one), while the opposite is true for the other three. This is quite a noteworthy result and operation type, with regard to length of stay, is discussed in more detail in Chapter 6.

Comparison with the literature is not straightforward as operation type classification is not uniform across all studies, though there is supporting evidence that length of stay is influenced by the type of surgical procedure performed (Clague et al. 2002) and the type of fracture, which in turn influences the operation type (Sund et al. 2009).

3.4.4   Age

An increased age was shown to be significantly associated with an increased acute length of stay in four of the five regression outputs (for all but ASA grade IV results) and was included four times as a splitting rule in the CART results. The mean and standard deviation of acute length of stay is displayed for various age groups in Figure 3.4.4i. Average length of stay is seen to increase almost consistently as age increases, where the results level for nonagenarian and centenarian patients. Standard deviation of length of stay is also shown to deviate across these age groups, with the highest value seen for age group 80-89 years. This group also has a coefficient of variation for length of stay greater than one, as does the group aged 50-59, indicating higher comparable fluctuations in these groups.
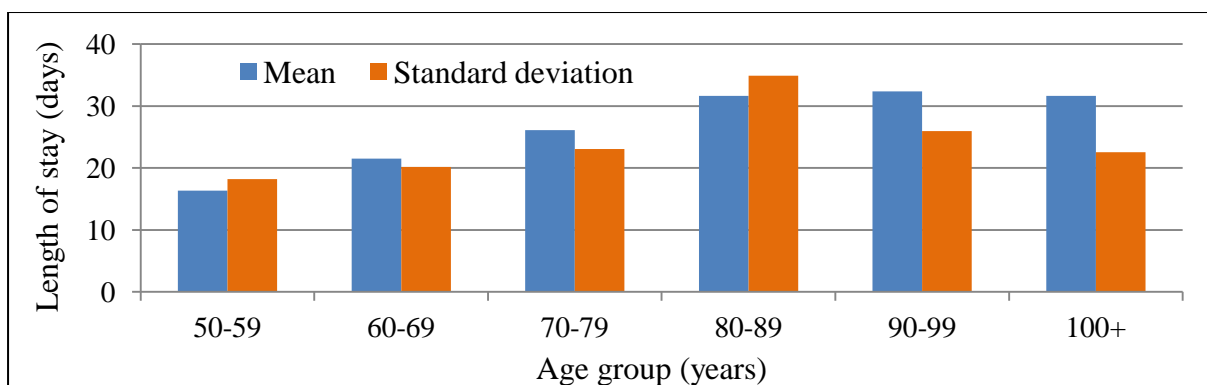
Figure 3.4.4i: Length of stay results by age group, patients undergoing surgery only

Increased age has shown to be a significant predictor of increased length of stay in other research studies, alongside a number of other factors (Lefaivre et al. 2009, Parker et al. 1998). It has been shown that patients aged 90 years or over have a hospital stay significantly greater than those aged 80-89 years, with the difference in length of stay between these groups reported to be ten days. Indeed, the notion of early surgery and early mobilisation of elderly hip fracture patients has been suggested to be extended to include all elderly patients, irrespective of fracture type (Clement et al. 2011). Another study also reported that age of 90 years or over significantly increases hospital stay after hip fracture (Shah et al. 2001).

3.4.5 ASA grade

ASA grade was not found to be a significant predictor of length of stay in the original regression model, which was not really an expected result. However, in the model for surviving patients only it was found to be a predictor variable, with a longer length of stay being associated with poorer medical fitness. The three final regression models related to each ASA grade and thus while ASA grade was therefore not an input to these models, the differences in results of these models indicate that ASA grade does matter when it comes to length of stay. ASA grade is also investigated further later in this thesis, see Chapter 6.

Additionally ASA grade has been shown to be a reliable predictor for post-operative length of stay for hip fracture patients, which in turn led to the suggestion that it is an appropriate predictor of cost (Garcia et al. 2011).

3.4.6 Sex

The final variable to be considered individually is gender. This was included in some of the regression models, but was one of the only variables to be excluded by the CART output. As discussed previously, the difference in mean length of stay between males and females is just under one day (30.05 days average for males *vs*. 29.06 for females). Where it was included by the regression analyses, males were found to have longer length of stay, so results are consistent with this. There is also a slightly greater difference in means if only surviving patients are considered, 29.92 days on average for males compared with 28.40 days on

average for females. This is consistent with the regression coefficient parameter estimates for gender found by the regression calculations, which were $\beta = 0.0847$ for the full model and $\beta = 0.1158$ for the reduced model which only included surviving patients.

There is support in the literature that male hip fracture patients experience a longer length of stay in hospital than their female counterparts (Hollingworth et al. 1995), while other studies have reported no difference in length of stay between male and female patients (Arinzon et al. 2010, Dudkiewicz et al. 2011).

3.4.7 Ambulatory measures

There are various ambulatory measures which are now considered. Mobility score was found to be significantly associated with acute length of stay in both the full regression model and the reduced model, with an increased score (indicating poorer mobility) associated with increased length of stay. These results are endorsed by calculating the mean length of stay for each score, which are 22.42, 30.42 and 35.76 days for mobility scores *1*, *2* and *3* respectively. Despite this, mobility score (*mobility*) was not used as a splitting variable in the CART results. Walking aids used (*walkaid0*) and walking ability (*walking0*) on admission, however, both feature in the CART analysis. These variables were not found to be significant predictors by the regression analyses, with the exception of *walking0* which appeared in the model for ASA grade III.

Each of the variables mentioned here are of course inextricably linked to one another, which goes some way to justifying the inconsistencies in results from the two methods investigated. Spearman rank correlation coefficients were 0.5106 between *walkaid0* and *walking0*, 0.4476 between *walkaid0* and *mobility*, and 0.8051 between *walking0* and *mobility*. A moderate to strong positive correlation is thus seen in each case; also in every instance the *p*-value was < 0.0001 against the null hypothesis of no correlation.

By including one of the variables as an indicative factor of length of stay, ambulatory level is taken into account and so another variable may not be included for this reason; it has essentially already been captured.

This conclusion is supported by a separate study which found that independent ambulation pre-fracture was a predictor of a shorter LOSE (Semel et al. 2010). LOSE is length of stay efficiency, calculated as the gain in Functional Independence Measure (FIM) divided by length of stay. It therefore measures the rate of change of FIM; a functional score which comprises of 18 parameters, each rated on a scale of one to seven according to the degree of assistance required to perform a specific activity in three domains: basic activities of daily living, mobility level and cognitive function. A cumulated ambulation score, calculated over the first three postoperative days, was also found to be a highly significant predictor of hospital length of stay (Foss et al. 2006).

3.5   Other influencing factors

There is a wealth of literature which reports on factors influencing length of stay for this group of patients and thus not all results will be detailed here. It was not possible to compare many of the published results with the results presented earlier in this chapter due to particular patient or clinical information being unavailable (patient ethnicity, for example).

Postoperative complications, reported to occur in a third of cases by one study, were shown to result in significantly longer length of stay in hospital for this patient group, while predictors for a complication are female gender and poor mobility status pre-fracture (Merchant et al. 2005). Deep wound infection is one complication that has been shown to lead to a significantly longer stay in hospital (Edwards et al. 2008), as does MRSA infection (Pollard et al. 2006).

Patients with a history of cerebrovascular accident (stroke) have been shown to have a significantly longer length of stay after hip fracture than those without, but there were no differences in in-hospital or one year mortality between these groups of patients. They also experienced the same levels of functional recovery (Youm et al. 2000).

Length of stay has also been shown to be influenced by race/ethnicity, with differences in hospital stay between non-Hispanic white and minority groups (Asian, non-Hispanic black) shown to be statistically significant. Similar conclusions were drawn relating to the probability of being discharged home (Graham et al. 2008). Patients of non-Hispanic black, Hispanic and Asian ethnicity have also been shown to experience a longer length of stay if they are discharged to a rehabilitation unit (Sterling 2011).

Albumin level, a measure of nutritional status, was found to be the only factor that significantly predicted length of stay in elderly hip fracture patients, where a negative relationship was found, with a β-coefficient of -0.23 in the linear regression model (Van Hoang et al. 1998). This is validated by another study which showed low albumin levels were a predictor of longer length of stay after hip fracture (Koval et al. 1999). Conversely, total hospital stay has been shown to be a predictor, amongst other variables, of a change in albumin levels for geriatric patients (Botella-Carretero et al. 2008). Patients with anaemia, as measured by haemoglobin level on admission, are also more likely to have a longer stay in hospital (Gruson et al. 2002).

Hospital settings were studied to identify differences in care received by hip fracture patients. It was shown that patients in an associated teaching hospital in an urban setting had approximately 14 days shorter length of stay than those in an inner-city teaching hospital, and this difference was attributed to the urban hospital having a greater supply of non-acute beds (Beech et al. 1995). Patients treated in rural community hospitals have also been shown to have longer length of stay (mean 22.5 days) than their counterparts treated in an urban teaching or urban community hospital (mean lengths of stay 17.6 days and 17.1 days respectively) (Weller et al. 2005). Focussing on location within one hospital, a Swedish group investigated the impact on length of stay for 'outliers'; patients with a hip fracture who are inappropriately admitted to another ward due to limited beds in the orthopaedic department. It was shown that patients treated elsewhere experienced an additional 3.7 days acute stay, and 13.6 extra days when rehabilitation is also considered (Hommel et al. 2008a).

Volumes relating to hospitals and surgeons have also been shown to influence hospital stay for hip fracture patients, with significantly longer lengths of stays shown for low-volume hospitals (less than 57 cases per year) and low-volume surgeons (less than seven procedures per year) (Browne et al. 2009).

A change in the on-call rota system for consultants working on an acute trauma ward was shown to statistically shorten length of stay for hip fracture patients. Significant improvements in time to theatre and promptness of discharge were also shown (Divecha et al. 2011).

Length of stay following a hip fracture has also been found to be significantly associated with marital status (living alone) and with the regular intervention of a caregiver (Pautex et al. 2005). Naglie et al also showed that additional care provision, in terms of whether patients received interdisciplinary or usual care, led to a longer length of stay (Naglie et al. 2002).

An example of a variable reported to have no impact upon length of stay is whether the hip fracture was the first fracture incurred, or whether it was a sequential fracture. It was also noted that functional recovery, as measured by the Barthel Index, was also not affected by whether it was a first or recurrent fracture (Di Monaco et al. 2002). The Barthel Index is used to measure functional ability based on ten activities of daily living, each rated on a scale of zero to ten; the scores are then summed to give a total score between zero (totally dependent) and 100 (totally independent) (Mahoney and Barthel 1965).

3.6   Chapter summary

In Chapter 2, an overview of the data available relating to trauma hip fracture patients was presented. A whole host of variables, alongside summary measures, were introduced and analysed with respect to this patient group. This chapter has expanded on those analyses and results with a particular focus on acute length of stay.

Length of stay was chosen as the target variable here for numerous reasons. Firstly, the benefits of a more predictable system are manifold; by being able to predict how long a patient will spend in hospital, decisions regarding capacity planning, manpower and resources can be made with increased confidence and intelligence. This is also more beneficial to the patient. The majority of this patient cohort spends several weeks in hospital, so a better idea of when they may be discharged would be uplifting and helpful to the patient and their family.

The results gained from the statistical methods completed have provided useful understanding into factors which influence length of stay. In order for any predictions or insight to be made in a timely fashion, only variables which are known either on arrival or soon after the arrival of a patient are considered. The techniques used here are also well-established and trusted in the healthcare field; there is little advantage to be gained by producing incomprehensible results, or results which the stakeholders will have difficulty to have confidence in.

One of the primary outcomes of this chapter has shown just how complex an issue this is. The evidence presented has shown that there are many influencing factors on length of stay and while it would be convenient if just one or two predictors were found, this was never really expected. (Indeed if this was the case, these analyses would never have been embarked upon!) Other useful results include just which of the variables are the most important and influential for this topic, as well as those which are of least importance.

Almost all of the variables considered here are determined by individual patient characteristics, which can be separated into two groups; medical and non-medical characteristics. The first group includes factors such as mental state, ASA grade and operation type, while the second is made up of factors including age, gender and admission source. Despite these differences, there is one important aspect that these two groups have in

common; the medical team have no influence or control over them. The results for these variables are thus mostly of value from a statistical point of view. Arguably more interesting, however, are those factors which the medical team do have an element of control over. It has been mentioned previously that delay to operation is a particular focus of this thesis, and indeed within the healthcare research community. Results found here have consistently shown that a delay, classed as more than two days, does have an impact on length of stay. This is a valuable result as it proves that this organisational issue, if resolved, would provide benefits to the system.

While the statistical results presented are informative in their own right and provide some useful information for the hip fracture team, here they also may be used as a decision tool. One objective of this thesis is to build a simulation model to represent the patients and bed usage on the trauma fracture ward. By compiling the results seen until now, important influencing factors can be identified with respect to what influences ward length of stay and thus the system as a whole. Despite this, the variable representing mental state, which was the only variable to appear in each statistical output, is not included in the simulation model. It was not found to be as influential when considering mortality (see Chapter 4) and since the simulation also models patient outcome, mental state was not included. Furthermore, it was decided that some measure of medical fitness should be incorporated into the model and ASA grade is a more general measure of this.

CHAPTER 4: MORTALITY ANALYSIS

4.1   Introduction

Factors influencing length of stay for trauma hip fracture patients were investigated in the previous chapter. The aim here is also to investigate which demographic and clinical variables may influence the patient's stay in hospital, but this time with respect to mortality. Due to data restrictions only acute mortality is considered. Some patients die while still under the care of the University Health Board, after their acute discharge, but this information was missing in many cases and therefore further investigation is not viable.

The same data as per Chapters 2 and 3 was used here.

Results are compared with the literature, where appropriate, and an overview of other reported findings are subsequently given. A particular focus of this chapter is the investigation of whether operative delay influences risk of mortality. This relationship is examined in detail using both multivariate and univariate analyses.

Engagement with clinicians was undertaken prior to the commencement of any statistical analyses being performed. This was to ensure that models would be fit for purpose and useful in a clinical setting. The methodologies used were also discussed to ensure that they were appropriate to the stakeholders.

4.2   Logistic regression

4.2.1   Introduction

Logistic regression is a statistical method used to predict the probability of the occurrence of a binary event, based on a number of independent predictor variables.

The logistic function takes the form

$$f(z) = \frac{1}{1 + e^{-z}}.$$

It is a useful function since the values for z can take any number in the range $(-\infty, +\infty)$, while the outcome $f(z)$ will always be between 0 and 1. The variable z represents the predictor variables, while $f(z)$ represents the probability of a particular outcome, given this set of input variables. A plot of the relationship between z and $f(z)$ is given in Figure 4.2.1i.
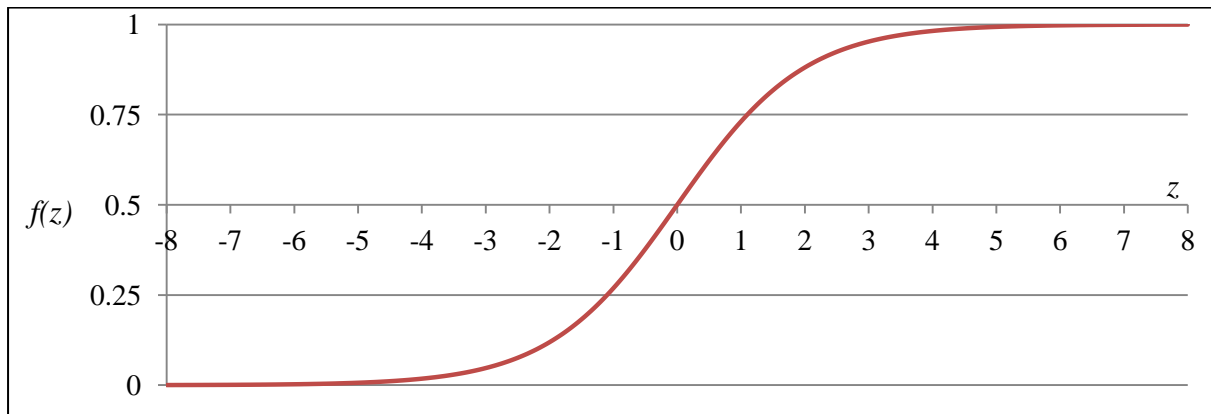
Figure 4.2.1i: A plot of z against $f(z)$

This is clearly a monotonically increasing function, so that any value which causes an increase in z will in turn cause an increase in $f(z)$, any value which causes a decrease in z will cause a decrease in $f(z)$, while $z = 0$, representing no effects, results in $f(z) = 0.5$.

The benefit of a logistic regression can be seen here by the fact that not only is the function bounded between 0 and 1, it also comes closer to being equal to these values, while linear

regression does not have this quality. Another advantage of this technique over linear regression is that it does not assume homoscedasticity of the independent variable across the explanatory variables, since this cannot occur for binary variables.

This model has an equivalent formulation which aids with interpretation. Let $\pi = f(z)$, so that now π is the probability of the event occurring.

Rearranging gives

$$\ln\left(\frac{\pi}{1-\pi}\right) = z$$

This value is also known as $\mathrm{logit}(\pi)$. While π denotes the probability of the event occurring, $\mathrm{logit}(\pi)$ represents the log odds of the dependent variable.

It was previously mentioned that z represents the predictor variables and an explicit form of this is given by

$$z = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p,$$

where β_0 represents the intercept and β_j represents the estimated coefficient for variable X_j, $j = 1,\ldots,p$.

It follows that $\dfrac{\pi}{1-\pi} = e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p}$.

4.2.2   Aim of this analysis

The aim here is to estimate the investigate mortality on the acute ward. In doing so, the important factors which influence mortality are highlighted, as well as their degree of importance. Clearly mortality is a binary variable, thus logistic regression is an appropriate tool to use here, so that now π represents the probability of dying while on the acute ward.

The statistical computer package SAS 9.1.3 was used for this analysis. As seen previously, a stepwise selection model was used with significance of entry and removal set at 5%. The variable under consideration is *survival_ac,* referring to Table 3.2.1i and Table B3.2.1a, while the aim is to find what variables predict that this takes a value of one.

4.2.3 Model validity

The validity of a logistic regression model can be assessed in many different ways. A large selection of these methods is discussed here before the model is presented.

• Likelihood ratio

The likelihood ratio tests the null hypothesis that $\boldsymbol{\beta} = 0$, where $\boldsymbol{\beta}$ is the vector of coefficients for the model parameters. Here this ratio has a Chi-square score of 219.95 where $p < 0.0001$, indicating that the vector of model parameters is significantly different from zero.

• Hosmer-Lemeshow goodness-of-fit test

Many tests for the goodness-of-fit of a model are performed by analysing residuals, but this is not possible for a binary response. This particular statistic however is only available for models where the response in binary; predicted probabilities are divided into deciles and a Pearson Chi-square statistic is computed which compares the predicted and observed frequencies across the 2 x 10 table (Hosmer and Lemeshow 1989). The result quoted by SAS is actually a lack of fit test, so a small p-value indicates that the fitted model is not an adequate model. Here $p = 0.4238$, so a significant fit has been found.

• Concordant pairs

Another output given is the percentage of concordant pairs; a pair of observations with different observed responses is said to be concordant if the observation with the lower ordered response value has a lower predicted mean score than the observation with the higher ordered response value. Here this value is 78.4%.

• Somers' *D* statistic

Similarly, Somers' *D* statistic measures the association between pairs of observations, but this time asymmetrically; that is, it also measures the direction as well as the strength of association. It is the surplus of concordant pairs as a percentage of concordant, discordant and tied pairs of observations; given the condition that a randomly selected pair of observations are not tied on the independent variable, Somers' *D* is the conditional

probability that the pair is concordant minus the conditional probability that the pair is discordant (Liebetrau 1983). It reaches a maximum of 1 for perfect association (all pairs agree) and a minimum of -1 for no association (all pairs disagree). In this case this statistic is 0.572 indicating a reasonably good level of association.

• *c*-Statistic

Another fit statistic given by SAS is *c*, a variant on the Somers' *D* statistic. It is the rank correlation when measuring on an ordinal level, where -1 indicates 100% negative association, 1 indicates 100% positive association and a value of 0 indicates the absence of an association between the two variables. Here $c = 0.786$ indicating a good level of association. Note that *c* also gives an estimate of the area under the receiver operating characteristic (ROC) curve when the response is binary (Hanley and McNeil 1982). The ROC curve obtained in this instance can be seen in Figure 4.2.3i and displays a moderate fit for the logistic regression model. A ROC curve is a plot of the true positive rate against the false positive rate and demonstrates several things. The closer the curve follows the left-hand border and then the top border of the ROC space (a "Γ" shape), the more accurate the test. Similarly, the closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test.
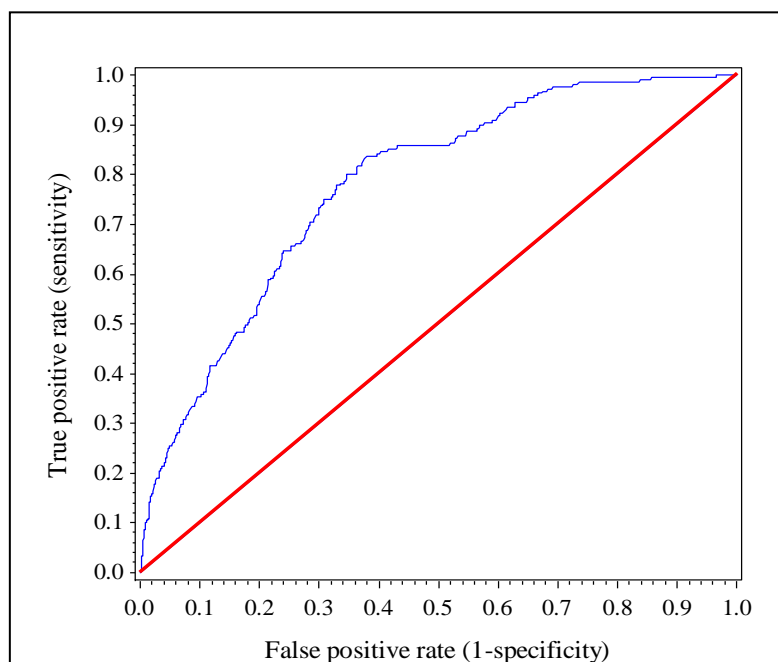
Figure 4.2.3i: ROC curve for the fitted logistic regression model

103

• Akaike's Information Criterion (AIC)

AIC measures the goodness of fit for a statistical model. SAS computes two estimates of AIC, firstly just for the intercept and then also for the intercept and the covariates. A lower value indicates a better model. Here the values were found to be 1390.360 and 1188.408 respectively, so a better model can be found when covariates are introduced.

• Maximum re-scaled R^2

A more appropriate measure to use with logistic regression is the maximum re-scaled R^2 value, denoted by $\tilde{R}^2$, rather than the usual R^2 value usually used in regression methods to measure the variation that is explained by the model. This is due to the fact that R^2 cannot achieve a maximum of 1 for discrete models. The maximum R^2 that can be achieved is equal to $R^2_{\max} = 1 - L(0)^{\frac{2}{n}}$, where $L(0)$ is likelihood of the intercept-only model and n is the sample size. To obtain the value of $\tilde{R}^2$, the following formula is used:

$$\tilde{R}^2 = \frac{R^2}{R^2_{\max}} \quad \text{(Nagelkerke 1991).}$$

Unlike R^2, this adjusted coefficient can achieve a maximum value of 1. In this case, the R^2 value was 0.1045, which is rather small. However, the value of $R^2_{\max}$ was found to be 0.5017. Using the previous formula, it is found that the maximum re-scaled R^2 value is given by $\tilde{R}^2 = \frac{0.1045}{0.5017} = 0.2083,$ which is a considerably better result.

The results discussed previously all indicate that the logistic regression model is an appropriate tool to use here. It is, however, accepted that these results are not optimal and it is always hoped that more desirable findings are found. Despite this, the conclusion can still be drawn that the method is suitable here and some insight can be gained from the analysis.

4.2.4 Results

The resultant model gave a total of nine variables as significant predictors of mortality, as displayed in Table 4.2.4i. A total of 1992 observations were available for this analysis; 221 of these patients (11.1%) did not survive their acute stay hospital. Note that multicollinearity was tested for in each of the logistic regression models presented in this chapter and levels were again low enough that they could be discounted. It has been suggested that VIF values above 2.5 may be a cause for concern in logistic regression models (Allison 1999), but VIF values produced here did not violate this recommendation in any of the models formulated.

Table 4.2.4i: Parameter estimates for the logistic regression model

Variable	Parameter estimate	Standard error	Wald Chi-Square	Pr > ChiSq
Intercept	-3.2567	0.2656	150.39	< 0.0001
Mobility	0.5169	0.0990	27.28	< 0.0001
Mentalst	0.1784	0.0730	5.98	0.0145
WAASP	0.2956	0.0930	10.10	0.0015
Age	0.3917	0.1077	13.23	0.0003
Opdelay	0.2060	0.0777	7.03	0.0080
ASAnew_n	0.4685	0.0839	31.18	< 0.0001
SexM	0.2362	0.0853	7.66	0.0057
Fractype_d1	-0.3344	0.1606	4.33	0.0374
Admfrom_d4	-0.6064	0.2051	8.74	0.0031

The regression equation is thus

$$\text{logit}(\pi) = -3.26 + 0.52(Mobility) + 0.18(Mentalst) + 0.30(WAASP) + 0.39(Age)$$
$$+ 0.21(Opdelay) + 0.47(ASAnew\_n) + 0.24(SexM)$$
$$- 0.33(Fractype\_d1) - 0.61(Admfrom\_d4),$$

while the corresponding logistic function is:

$$f(z) = \pi = \frac{1}{1 + e^{-\{-3.26 + 0.52(Mobility) + ... - 0.33(Fractype\_d1) - 0.61(Admfrom\_d4)\}}}.$$

Recall that π represents the probability of dying while on the acute ward.

A useful way to interpret the results of logistic regression is in terms of odds ratios. The odds ratio (OR) for each variable is found by taking the exponent of the parameter estimate. (Note that a more detailed definition and description of odds ratios is in given in Section 4.6.) This is with the exception of binary variables, due to the design of the regression the parameter estimate is first doubled before the exponent is calculated (Der and Everitt 2002). Note that OR results in Section 4.6 may differ from those quoted here; since logistic regression requires an entry across all variables for each observation, a slightly different subset of data will have been used there.

Table 4.2.4ii: Odds ratios for the parameter estimates of the logistic regression model

Variable	Adjusted OR	95% Confidence Interval for OR
Mobility	1.677	[1.381, 2.036]
Mentalst	1.195	[1.036, 1.379]
WAASP	1.344	[1.120, 1.613]
Age	1.479	[1.198, 1.827]
Opdelay	1.510	[1.113, 2.047]
ASAnew_n	1.598	[1.355, 1.883]
SexM	1.604	[1.148, 2.241]
Fractype_d1	0.512	[0.273, 0.962]
Admfrom_d4	0.297	[0.133, 0.665]

The most important factors seen here in terms of increasing the odds of dying while on the acute ward are mobility score, sex and ASA grade. For every unit increase in mobility score, there is 1.677 times the odds of dying, while for every unit increase in ASA grade, there is 1.598 times the odds of dying. A male patient has 1.604 times the odds of dying compared with a female patient, given that other factors remain the same. Another important result is seen for delay to operation; a delayed patient has increased odds of dying of 1.510 compared with a patient who is not delayed.

It is also interesting to look at the odds ratios that are less than one. Note that the odds for a patient dying who has incurred a fracture type coded as *1* are 0.512 compared with those with a different fracture type; that is, given that the other factors remain the same, a patient suffering from a undisplaced intracapsular fracture is less likely to die while on the acute

ward. Patients admitted from a nursing home also have lower odds of dying; 0.297 compared with those admitted from elsewhere, given that other factors remain the same.

4.2.5 Analysis by ASA grade

Logistic regression analysis was then repeated within each ASA grade in order to gain more homogenous groups in terms of medical fitness. ASA grade was found to be a significant predictor of mortality, so this has already been captured to a certain extent, but it would be interesting to see if results differ across ASA grades.

A total of 748, 1054 and 190 observations were available for ASA grades I&II, III and IV respectively. A summary of parameter estimate and odds ratio results is given in Tables 4.2.5ii-iv, but first a summary of some of the goodness-of-fit measures are displayed.

Table 4.2.5i: Goodness-of-fit measures for logistic regression performed by ASA grade

Model for ASA grade	Likelihood ratio (*p*-value)	Hosmer-Lemeshow test (*p*-value)	Concordant pairs (%)	*c*	$\tilde{R}^2$
I&II	< 0.0001	0.8864	75.4	0.801	0.3255
III	< 0.0001	0.5499	73.1	0.735	0.5173
IV	0.0029	0.7458	42.2	0.691	0.7107

Significant logistic regression models have been found in each case. In particular, consider the improvement in the value of $\tilde{R}^2$ from that seen previously, suggesting that this subsequent analysis is a viable avenue to explore. The Hosmer-Lemeshow test is not recommended when n is small, or more specifically less than 400 (Hosmer and Lemeshow 2000), so this result for ASA grade IV should be accepted with caution.

Just three variables were found to be significant predictors of mortality for ASAI&II patients, see Table 4.2.5ii; a higher probability of death for an increased mobility score and undergoing a hemiarthroplasty operation and a lower probability of death for those admitted from home. For those who underwent a hemiarthroplasty operation, the odds of dying on the acute ward were found to increase by 3.177.

Table 4.2.5ii: Results given by the logistic regression model, ASA grade I&II

Variable	Parameter estimate	Standard error	Pr > ChiSq	Adjusted OR [95% confidence interval]
Intercept	-2.7680	0.1917	< 0.0001	-
Mobility	0.6013	0.2181	0.0058	1.824 [1.190, 2.797]
Optypenew_d5	0.5779	0.1772	0.0011	3.177 [1.586, 6.363]
Admfrom_d1	-0.4751	0.2103	0.0239	0.387 [0.170, 0.882]

For ASAIII patients, a more complex model in terms of the number of parameters to retain was found, but results are still understandable and innate to what would be expected. Patients admitted from an acute hospital were found have odds of dying on the acute ward of over four compared with those admitted from elsewhere, other factors remaining the same. A delay to operation of more than two days was also shown to be a significant indicator of mortality.

Table 4.2.5iii: Results given by the logistic regression model, ASA grade III

Variable	Parameter estimate	Standard error	Pr > ChiSq	Adjusted OR [95% confidence interval]
Intercept	-1.7862	0.1940	< 0.0001	-
Mobility	0.7121	0.1352	< 0.0001	2.038 [1.564, 2.657]
Age	0.5504	0.1487	0.0002	1.734 [1.295, 2.321]
Opdelay	0.2669	0.1004	0.0078	1.705 [1.151, 2.528]
SexM	0.2690	0.1103	0.0147	1.712 [1.111, 2.638]
Admfrom_d1	0.3475	0.1217	0.0043	2.004 [1.244, 3.228]
Admfrom_d7	0.7572	0.1913	< 0.0001	4.547 [2.148, 9.623]

A higher mobility score was found to increase the probability of dying for ASA grade IV patients, with every unit increase in this variable increasing the odds of dying on the acute ward by 1.712. This model is also interesting in its simplicity, suggesting that the only

variable (under consideration) which impacts upon mortality amongst ASA grade IV patients is their mobility score.

Table 4.2.5iv: Results given by the logistic regression model, ASA grade IV

Variable	Parameter estimate	Standard error	Pr > ChiSq	Adjusted OR [95% confidence interval]
Intercept	-1.0209	0.1873	< 0.0001	-
Mobility	0.5375	0.1892	0.0045	1.712 [1.181, 2.480]

4.2.6 Conclusion

An increased mobility score, indicating poorer mobility function, was found to be a significant indicator of mortality in all cases. Admission source again featured heavily, with results for this as expected for the most part. When splitting by ASA grade, the model produced for ASA grade III resulted in the most predictor variables, while ASA grade IV had the fewest. This is influenced to some extent by the number of observations in each group. One interesting piece of information to take from this analysis relating to delay to operation is that it was included as a significant predictor for both the full model and also for ASA grade III patients after the splitting procedure. However, it must be noted that, for some patients, delay is clinically warranted (see Section 6.3.4) and the potential implications of this when interpreting statistical results should be considered. Unfortunately, delay reason was unknown for this part of the statistical analysis. An extension to this work could be to formulate sub-models, based on delay reason, but this is deemed to be beyond the scope of this thesis.

4.3 CART analysis

4.3.1 Introduction

CART analysis was utilised once more to create homogenous nodes of data, but this time with respect to ward mortality. The dependent variable under consideration this time is thus categorical and the method used by TreeWorks is different. In this case the aim is to maximise the gain in purity and thus at each stage the tree is grown by choosing which node and variable to split on in order to minimise the impurity in the dataset. Consider a dependent variable which may take up to m values, specifically the integer values $1, 2, ..., m-1, m$. Now let $f(i, j)$ represent a membership proportion for each of the m values: it is equal to the proportion of observations assigned to node i for which the value of the dependent variable is equal to j, $j = 1, ..., m$. This notation and the following equations are described in more detail in the literature from where this information was obtained (Harper and Leite Jr 2008). Consider the measure of impurity for node i, given by $I(i)$. The gain in purity achieved by splitting node i into nodes i_0 and i_1 is calculated by:

$$Gain(i; i_0, i_1) = I(i) - \left[I(i_0) p(i_0) + I(i_1) p(i_1) \right]$$

where $p(i_0)$ and $p(i_1)$ are the proportions of records assigned to node i_0 and i_1 respectively, calculated as $p(i_a) = \dfrac{s(i_a)}{s(i)}$, $a = 0, 1$, where $s(i)$ is the number of observations in node i.

The overall measure of impurity and the splitting specifications used depend upon the procedure used; there are two choices using the chosen software, the Gini Index of diversity or Information Entropy.

The Gini Index value for each node is based upon squared probabilities of membership for each of the m target categories within the node and is calculated as $I_G(i) = 1 - \sum_{j=1}^{m} f(i, j)^2$ for node i. The Information Entropy value is based upon the concept of entropy which is commonly used in information theory (that is, as a measure of uncertainty), and is calculated

as $I_E(i) = -\sum_{j=1}^{m} f(i,j) \log f(i,j)$. Note that here $m = 2$ since the dependent variable under consideration is mortality on the acute ward.

4.3.2 Results

The options were amended slightly here; as before, many permutations of the changeable parameters were tried and the most desirable used. In this case, the minimum number of observations per node was set to 30. The learning sample again comprises of 70% of all available observations, 1386 records in total, while the remaining 30% are used as the test sample for validation purposes.

The Gini Index procedure was initially used for this analysis and results are given in the form of the tree in Figure 4.3.2ii and the spitting rules in Appendix D, Table D4.3.2a. As previously, final nodes are denoted with an 'F' and the node number is given in the top line of each node. Other results given within each node are percentages relating to the number of surviving and non-surviving patients within each node, as described below.

Table 4.3.2i: Key for CART mortality results

Node number	
Surviving patients:	% of node | % of all patients
Non-surviving patients:	% of node | % of all patients

Note that due to space restrictions all figures are quoted to the nearest integer and so a value of zero may not necessarily mean that there are no patients in that node with a particular characteristic (see node 38F for an example of this).

0
88|88
12|12

walking0

1
82|46
18|10

2
97|42
3|2

admfrom

ASAnew_n

3
84|42
16|8

4
65|4
35|2

37
97|40
3|1

38F
85|2
15|0

optypenew_n

mentalst

fractype

5
90|22
10|2

6
79|20
21|5

35F
76|3
24|1

36F
49|1
51|1

39
95|20
5|1

40
99|20
1|2

ASAnew_n

admfrom

fractype

WAASP

7
96|22
4|0

8
87|15
13|2

21
73|9
27|3

22
85|11
15|2

41
94|17
6|1

42F
100|3
0|0

51
100|17
0|0

52F
98|3
2|0

walking0

fractype

WAASP

walkaid0

sexM

fractype

9
93|4
7|0

10F
100|3
0|0

13F
78|3
22|1

14
90|12
10|1

23
80|5
20|1

24
66|4
34|2

29F
97|2
3|0

30
82|9
18|2

43
96|12
4|1

44
91|5
9|1

53F
98|4
2|0

54F
100|13
0|0

optypenew_n

walking0

opdelay

walkaid0

admfrom

walkaid0

opdelay

11F
97|2
3|0

12F
90|2
10|0

15
93|8
7|1

16
84|4
16|1

25F
85|3
15|1

26F
72|2
28|1

27F
80|3
20|1

28F
50|1
50|1

31
79|6
21|2

32F
89|3
11|0

45F
93|4
7|0

46
97|8
3|0

49F
94|3
6|0

50F
88|2
12|1

mobility

WAASP

optypenew_n

ASAnew_n

17F
98|3
2|0

18F
90|5
10|1

19F
74|2
26|1

20F
94|2
6|0

33F
73|3
27|1

34F
84|3
16|1

47F
96|5
4|0

48F
100|3
0|0

Figure 4.3.2ii:CART results for mortality using the Gini Index procedure method

This case resulted in 28 final nodes. The variables which were not used as splitting criteria were side of fracture and age. The overall reduction in impurity achieved was 16%, so an increase in homogeneity has been achieved. Success or failure of these results may also be considered in terms of the predictions made using the test sample. The percentage of correct predictions, using the splitting rules, classified as dead or alive are displayed in Figure 4.3.2iii. These are given for final nodes only since these would be used should these classifications be used later; every patient would reside in one final node only.
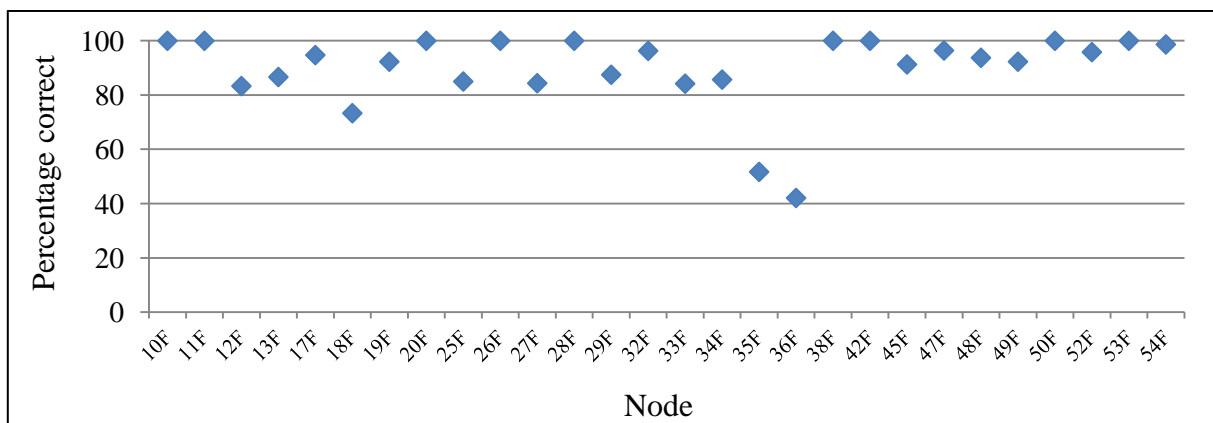
Figure 4.3.2iii: Validation of final nodes produced by the Gini Index procedure

Nine of the final nodes achieve the maximum of 100% of correct predictions, which is almost a third of all final nodes, and all but three nodes have a success rate of over 80%. Nodes 35F and 36F have a considerably lower percentage of correct values in comparison to the other final nodes, suggesting that this split on fracture type may not be appropriate. Fracture type was used on three other occasions as a splitting variable so its inclusion overall is still appropriate. The average percentage of correct values across these final nodes is 89.8% and so it is concluded that this CART analysis has provided accurate results. The standard deviation of these values was found to be 14.1 (percentage points).

ASA grade, walking ability, walking aids used and operation type are amongst those variables which appear more frequently in the output. The first split is made on walking ability on admission, suggesting that this is the most telling variable which influences mortality. However, results are unintuitive; the best and worst levels (in terms of ability) are grouped together, while the three middle levels comprise the other group. ASA grade is used on the second level of the tree, and then

two more times in lower levels, indicating this is an important predictor of mortality. Since ASA grade is a measure of medical fitness, this result is as expected. Admission source is another variable found to be used for one of the first splitting decisions and is also used twice later on.

Using the Information Entropy procedure, better results were achieved in terms of a higher reduction in impurity; this time 22% was observed, compared with 16% found previously. The tree had 60 nodes in total, 31 of which were final nodes. Results are as per quoted in Table 4.3.2i and are given in Table D4.3.2b and Figure 4.3.2iv.

Validation results are displayed in Figure 4.3.2v. In this case the percentage of correctly predicted observations is 100% in just over one third of all final nodes, a total of eleven altogether. The mean percentage of correct values is 89.7%, a very similar value to when the Gini Index procedure was used. There is less spread in this case, however, with a standard deviation of percentage correct values of 11.7 percentage points. The lowest values were found for nodes 10F and 54F, which both relate to a split on ASA grade. This may therefore suggest that such a split was not appropriate at those levels; however, it may still be concluded that using ASA grade as a splitting variable is suitable for mortality since it was used four other times in this procedure.
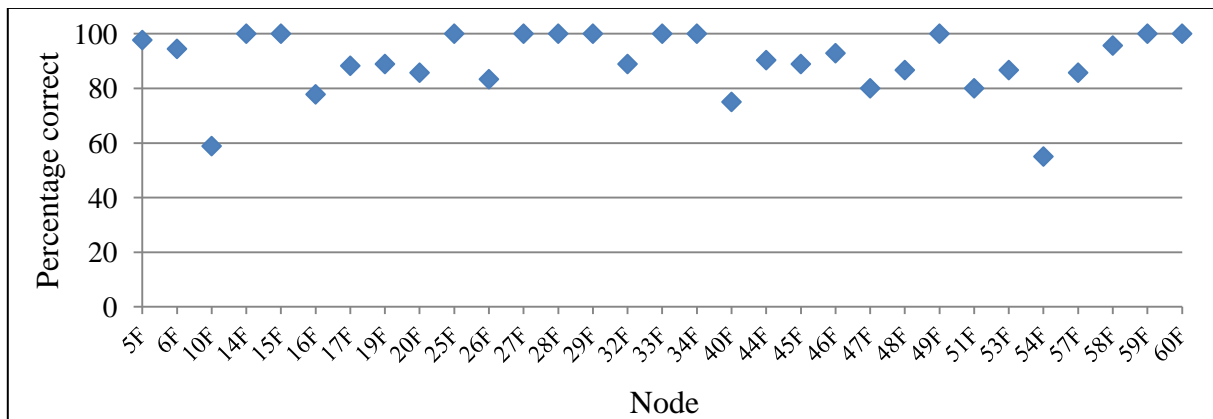
Figure 4.3.2v: Validation of final nodes produced by the Information Entropy procedure
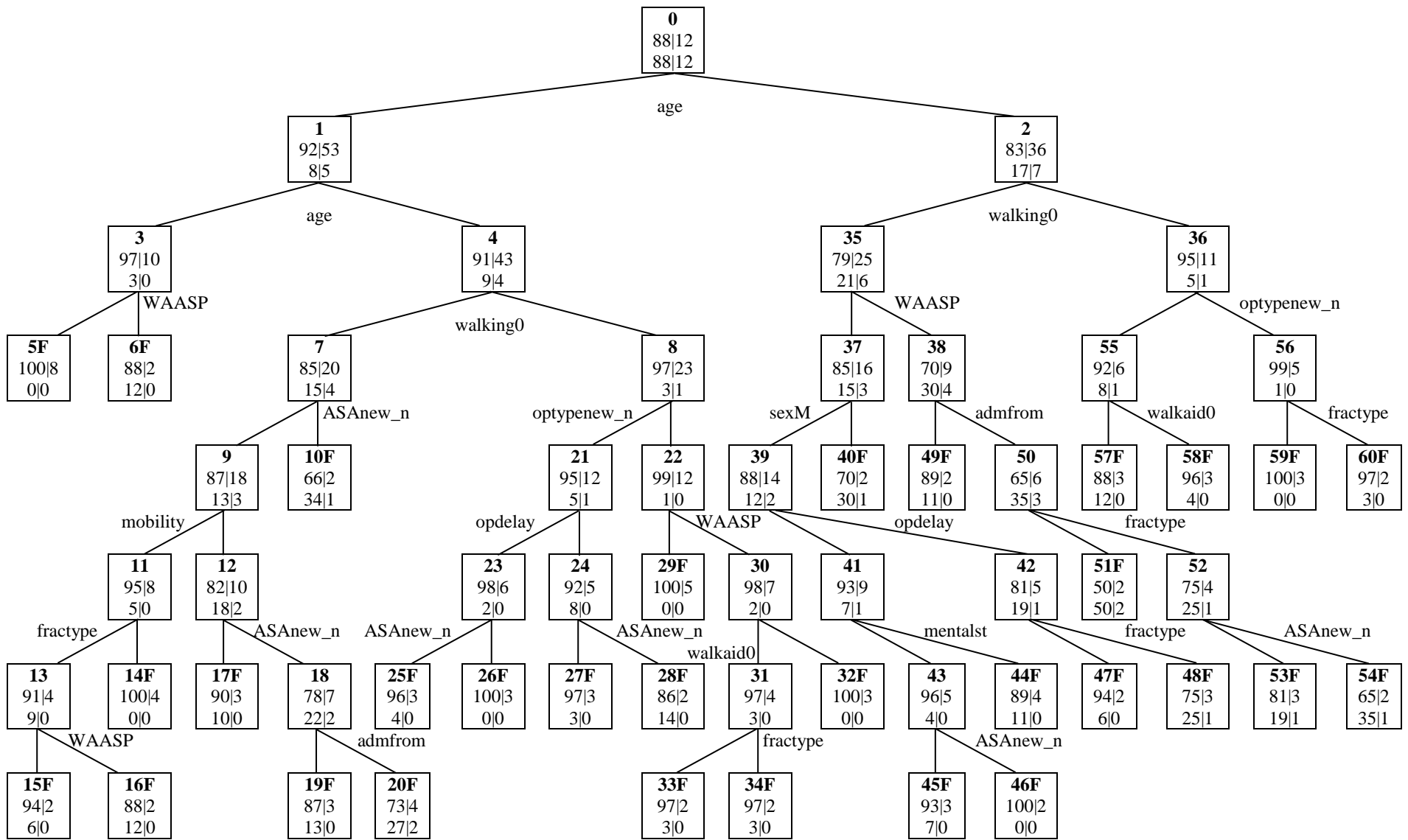
Figure 4.3.2iv: CART results for mortality using the Information Entropy procedure method

The only variable not used here is the side of the body on which the fracture was incurred. Using the Gini Index procedure, age was not used as a splitting variable, while this second method uses age as the first variable on which to split, so a conflict is observed. *Walking0* is used in the second level of the tree, indicating this is an important factor to consider with respect to ward mortality, but the same rather peculiar split across the five levels that this variable can take is seen. Type of fracture is used numerous times within the output, as is WAASP score.

4.3.3 Conclusion

One of the most useful insights to be gained from the CART analysis is that the majority of variables needed to be used in order to gain homogeneity with respect to mortality. This highlights the potential difficulties with predicting mortality as well as the complexity of this issue; the influencing factors are both numerous and varied. The only variable to be excluded by both methods was side of fracture. Other difficulties relating to this analysis is the relatively low incident (death) rate for this group. As a result, some caution should be exercised when interpreting the output; while a high percentage of nodes correctly predicted outcome, a high percentage of patients (approximately 88%) had the same outcome (of survival).

Variables which seemed to be of most importance include ASA grade and walking ability on admission, as well as fracture type, operation type and WAASP score. With regard to delay to operation, results were all indicative that delay does matter when considering ward mortality. This variable featured four times in total in the output (two per method), and after each split the node belonging to delayed patients had a higher percentage of non-surviving patients than the node for patients who were not delayed.

4.4 Comparison of methods and results

Four regression models and two CART procedures have been compiled in order to determine influencing factors for death on the acute trauma hip fracture ward. A number of variables are now investigated in more detail to ascertain the relationship between these variables and mortality.

4.4.1 Mobility score

Mobility score was the only variable to be included in all four regression equations, as well as both of the CART outputs. Indeed, it was found to be the only variable significantly associated with mortality for ASA grade IV. Recall that mobility is measured on an integer grade between one and three, with a higher score indicating a poorer level of mobility. In each case a higher score also indicated an increased probability of death on the acute ward.

Firstly, results are considered overall. 3.6% of patients undergoing surgery with a mobility score of *1* did not survive their stay in hospital. This increases to 10.3% for patients with a mobility score of *2* and increases again to 19.2% for patients with a mobility score of *3*. These results are therefore consistent with the statistical results found.

Now consider these percentages within each ASA grade, as displayed in Figure 4.4.1i. The β parameter estimates for mobility score were 0.6013, 0.7121 and 0.5375 for ASA grades I&II, III and IV respectively, indicating a greater impact for ASA grade III in comparison with the others.
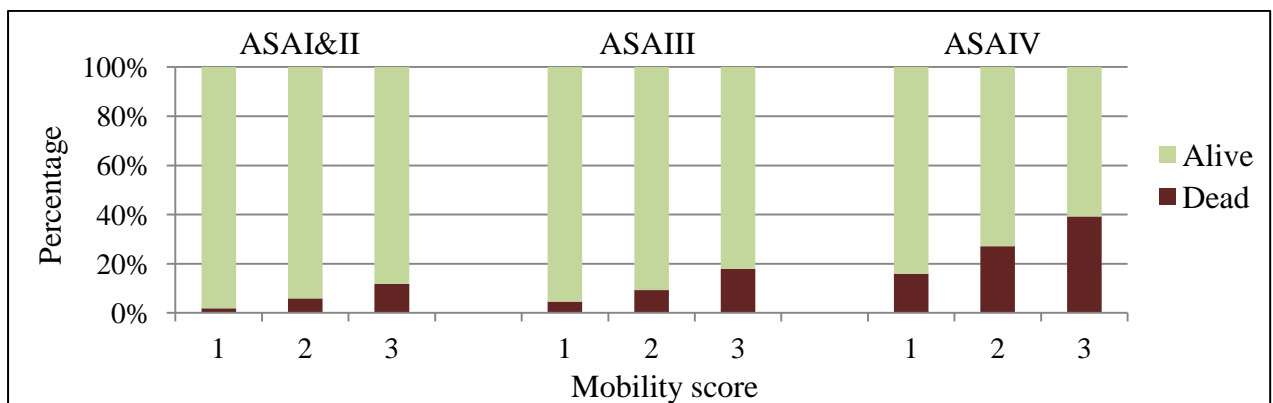
Figure 4.4.1i: Mortality results by mobility score and ASA grade

These results are also consistent with the regression results, where in each case an increased mobility score is shown to be related to an increased probability of death on the acute ward. Therefore, even when the medical fitness of patients is taken into account, there is still evidence of a relationship between mobility and survival. Each of these results is for patients undergoing surgery only.

Similar results have been found in other studies, where a significantly shorter survival time has been shown for patients with poor mobility before injury (Kopp et al. 2009). Amongst other factors, this study also showed increasing age and male gender to be related to a shorter survival time (see Sections 4.4.3 and 4.4.4 respectively). Pre-fracture functional ability has also been shown to be a predictor of six-month mortality, along with comorbidities, increasing age and surgery more than 48 hours after admission (Maggi et al. 2010).

Recovery after hip fracture has also been shown to depend largely on the pre-fracture health and functional ability of the patient, with better functional recovery recorded by men than for women (Arinzon et al. 2010). A mobility score calculated three days after surgery has also been shown to predict 30 day mortality in this group of patients (Foss et al. 2006).

4.4.2 ASA grade

While ASA grade has been considered to a certain extent in the previous section, it is now considered solely instead of in conjunction with another variable. The full regression model indicated that an increasing ASA grade was associated with an increased probability of death on the acute ward, with a parameter estimate of $\beta = 0.4685$ and an odds ratio of 1.598. Additionally, ASA grade was used as a splitting variable a total of nine times between the two CART outputs. As an example, consider the results from the Information Entropy procedure, in particular nodes 9 and 10F. This is a split made at a relatively high level in the tree. Patients with an ASA grade of I&II or III are directed to node 9, which has a survival percentage of 87%. Meanwhile, 66% of patients in node 10F, for ASA grade IV patients, survive their hospital stay. These results are therefore consistent with the regression model.

To verify these results, the percentages of surviving patients within each ASA grade are now presented. For consistency reasons, these computations were for patients who were operated on only.

Table 4.4.2i: Percentage of patients who survive their acute hospital stay, by ASA grade

ASA grade	Percentage of surviving patients
I&II	95.1%
III	88.2%
IV	69.1%

As expected, an increase in ASA grade category results in a decreased probability of surviving on the hip fracture ward. These values therefore reinforce the results found statistically.

An increasing ASA classification has been shown to predict an increase in 30-day mortality rates across all age groups. This study was not exclusive to hip fracture patients but focussed on patients undergoing surgery aged 80 years and older. A progressive increase in mortality was also shown with increasing age (Turrentine et al. 2006). Mortality at 90 days has also been shown to be dependent upon ASA grade, among other factors, exclusively for hip fracture patients (Clague et al. 2002). ASA grades III and IV have been shown to be one of the significant predictors for one year mortality (Aharonoff et al. 1997), with one study showing a nine-fold increased risk of one year mortality for these patients (Michel et al. 2002). Three-year mortality has also been shown to be significantly greater for patients with ASA grade III, IV or V (Hamlet et al. 1997), while ASA grade was one of only two predictive factors which predicted mortality at 30 days post-surgery, the other being treatment by arthroplasty (Rae et al. 2007).

4.4.3 Age

Results for age, and thus the conclusions which could be drawn from these results, were found to be contradictory between the outputs. An increased probability of dying on the acute ward for older patients was found for both the full model and the reduced model for ASA grade III patients only. These results are consistent with the CART results using the Information Entropy procedure, where age was used as a splitting rule twice, to create pairs of nodes (1, 2) and (3, 4). Despite the apparent importance of this variable here – it was used in the higher levels of the tree indicating it is a key factor influencing the formation of

homogenous nodes with respect to mortality – age was not used at all by the Gini Index procedure, suggesting that age does not matter when it comes to outcome. In both cases, the node with patients of a higher age had a smaller percentage of surviving patients.

In order to investigate this further, first consider the differences in mortality rates between age groups, see Figure 4.4.3i. There were no patients who died in hospital younger than the ages displayed. The percentage of surviving patients decreases consistently with an increase in age, thus verifying the existence of a relationship reported previously. Summary statistics of age are presented in Table 4.4.3ii, grouped by outcome.
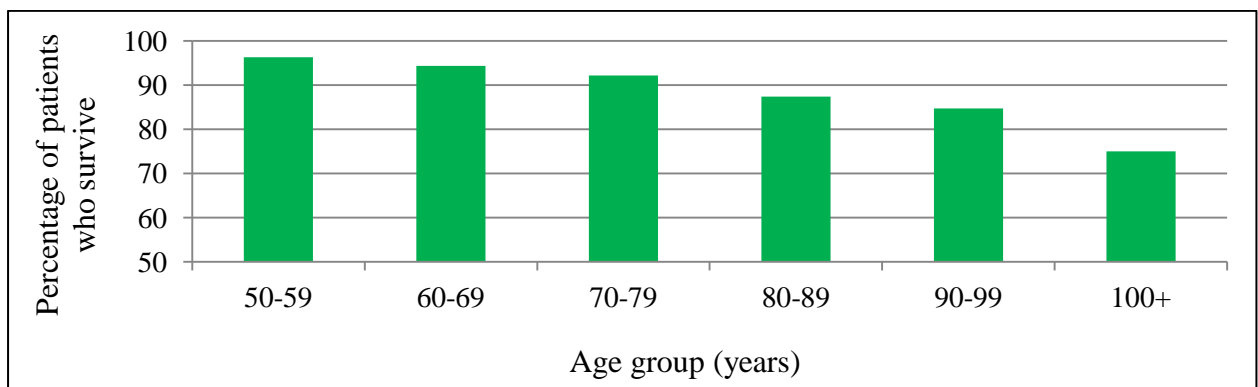
Figure 4.4.3i: Mortality results by age group

Table 4.4.3ii: Summary statistics for age (years) by outcome

Outcome	Mean	S.D.	C.V.	Minimum	Maximum
Dead	84.22	7.84	0.09	57	101
Alive	79.70	11.36	0.14	14	101

Again, death has been shown to be consistent with a higher age. The spread of ages, however, is larger in the group of surviving patients, both when the standard deviation and also the coefficient of variation are inspected.

Six month mortality has been shown to increase with older age following hip fracture, particularly when associated with dementia (Wood et al. 1992). It is interesting to note that one study demonstrated a direct relationship between increasing age and mortality for patients with intertrochanteric fractures, but no relationship for femoral neck fractures

(Kenzora et al. 1984). A similar study however found that age was the most significant predictor of mortality for these two patient groups; indeed operative delay was quoted as the only other influencing factor. Femoral neck fracture patients had a significantly shorter estimated survival, despite these patients being significantly younger than those with intertrochanteric fractures (Kesmezacar et al. 2010).

In-hospital and one year mortality rates have been shown to be greater for patients aged 90 years and over (Shah et al. 2001), while mortality at 30 and 120 days has also been shown to be higher amongst patients classified as extremely elderly, which in this case was aged 95 years or older. This was in comparison to group of patients aged 75 to 89 years old (Holt et al. 2008), while a study which concentrated specifically on centenarians showed a significantly higher mortality rate for this group during both hospital admission and at one year, concluding that there is a 20% increase in expected mortality for this age group (Forster and Calthorpe 2000).

4.4.4 Sex

Sex appeared in the full regression model, as well the reduced model for ASA grade III. It also appeared in both of the CART outputs. The conclusions that can be drawn in each case are that being of a male gender is associated with greater chance of death on the acute ward, with the exception of a splitting decision made by the Information Entropy procedure.

Calculating the percentage of deaths within each of the gender groups gives 14.4% of all patients undergoing surgery for males, compared with 9.9% for females, which explains and verifies the results found previously.

Male gender has been shown to be a significant risk factor for increased mortality after hip fracture in elderly patients, as well as being a significant predictor of sustaining a trochanteric fracture (Lin et al. 2011). Long-term survival analysis following hip fracture has also shown that men have a higher one year mortality rate after hip fracture and were also more likely to sustain a medical complication post-operation (Endo et al. 2005). Additionally, excess mortality was shown to be strongly significant for men compared with women (Kannegaard et al. 2010), a conclusion which is supported by evidence from a large hip fracture audit in Scotland (Johnston et al. 2010) and a large review article (Haentjens et al. 2010).

4.4.5   WAASP category

An increase in WAASP score implies a lower level of medical fitness, so it would be expected that this would also mean a higher death rate in these groups. This is confirmed by the logistic regression equation which considered patients as a whole, where WAASP category was included with a parameter estimate of $\beta = 0.2956$. This is not as large as the comparable β value for ASA grade (these figures are indeed directly comparable as they are both measured on a three-point scale), implying that WAASP category is related with ward mortality to a lesser extent than ASA grade. The percentage of patients within each outcome group is displayed by category in Table 4.4.5i. As expected, the proportion of non-surviving patients is smaller for lower WAASP categories.

Table 4.4.5i: Outcome by WAASP category

WAASP category	*1*	*2*	*3*
Dead	4.2%	9.2%	19.1%
Alive	95.8%	90.8%	80.9%

WAASP category was not found to be an influencing factor in any of the three regression models for each of the ASA grade categories, but this is most likely due to the similarity of these two variables. It was, however, used in both of the CART analyses, three times and four times for the Gini Index and Information Entropy procedure methods respectively. For the most part, the same conclusions would be made, but there were also some inconsistencies found. On three of the splits, categories *1* and *3* were grouped together, with a category of *2* comprising the other side of the split. The other splits, however, were all consistent with the regression output and the results displayed in Table 4.4.5i, with a higher WAASP category relating to a lower chance of survival. For each of these, categories *1* and *2* were grouped together and a category of *3* was separate. This is explained by looking at the percentage of patients who do not survive within each of the three categories.

These differences in conclusions highlight the importance of thorough investigation through data mining and statistical procedures. Uniform results are not necessarily going to be produced by the different methods but by exploring more than one technique and scrutinising the output in detail, some overall inferences can be made.

WAASP score was developed for use by and within Cardiff and Vale University Health Board and thus is not used in investigations conducted by research groups outside of the Board. However, recall that WAASP is measure of nutritional status (see Section 2.2.8 for more details) and so some general comparisons can be made.

A study conducted on the trauma hip fracture ward at the UHW compared whether employing dietetic assistants (DAs) affected clinical outcome; patients either received the conventional pattern of nurse- and dietician-led care or received additional personal attention from a dietetic assistant as well. Patients supported by a DA had significantly lower risk of mortality in-hospital and at four months (Duncan et al. 2006).

Albumin levels and total lymphocyte counts are often tested to measure nutritional status. One study found that albumin levels below three grams per decilitre was the only predictor of in-hospital mortality given by a multivariate logistic regression model (Pioli et al. 2006). Normal levels of albumin are 3.5-5.0 grams per decilitre (The National Kidney Federation 2011). Patients with low albumin levels, this time classified as less than 3.5 grams per decilitre, have been shown to experience higher levels of in-hospital mortality, while a total lymphocyte count of less than 1500 cells per millilitre was predictive of one-year mortality (Koval et al. 1999). Using the same cut-off to classify low levels of albumin, this result has been verified; it was calculated that patients in the low category have an odds ratio for mortality of 4.0 compared with those who are not, while increasing age was also found to predict mortality (O'Daly et al. 2010). The normal range of total lymphocyte counts for adults is reported to be 800-2600 cells per millilitre (Family Practice Notebook LLC 2012). Finally, haemoglobin level measured on admission showed that patients with anaemia have an increased risk of mortality at six and 12 months, but not at three months or in-hospital (Gruson et al. 2002).

Additionally, Fischer and Johnson reported that low body mass index (BMI) (and low weight / rapid weight loss), which is frequently reported the elderly and is caused by numerous physiological, psychological and social factors, stands alone as a high risk factor for mortality and morbidity in the older population (Fischer and Johnson 1990).

4.5   Other influencing factors

As with length of stay, there are a plethora of publications which have investigated factors influencing mortality and thus only a sample of these are considered here. As in Chapter 3, it is also not possible to compare some published conclusions with results found in the analysis presented in this chapter due to data restrictions.

Medical complications have been reported several times to be a significant predictor of mortality for hip fracture patients. For example, it has been shown that patients with chronic obstructive pulmonary disease (COPD) have an estimated 60-70% higher risk of death than those without COPD, after suffering from hip fracture (de Luise et al. 2008). Male veterans requiring transfusions were also shown to have a higher probability of death at 30 days than those who did not, while type of treatment was also shown to be an influencing factor (Radcliff et al. 2008). It has been reported that hip fracture patients who develop a medical complication after their operation have more than three times the probability of dying within one year post-fracture compared with those with no complications (Sexson and Lehner 1987), while Svensson et al showed that mortality at one year post-fracture could only be predicted by the number of current medical conditions (Svensson et al. 1996). Patients who have suffered a recent myocardial infarction (heart attack) are also at a much greater risk of dying (Komarasamy et al. 2007). The use of diuretics has also been shown to be the strongest independent predictor of mortality post-fracture, while the use of statins was associated with higher survival rates (Juliebø et al. 2010). Another study reported that patients with multiple comorbidities and clinically diagnosed postoperative complications, such as chest infections and congestive heart failure (CHF), were at a significantly higher risk of mortality (Roche et al. 2005).

One study analysed life expectancy at follow-up for hip fracture patients and showed that long term survival is dependent upon social dependence pre-fracture and age only. Interestingly, social dependence at six months was not found to be a significant predictor (Jensen 1984). Here social dependence was assessed by patient dependence on the social welfare system and ranked on a four-scale classification system (Thomas and Stevens 1974).

Type of anaesthetic procedure were investigated with respect to hip fracture surgery and subsequent patient outcome, where it was shown that risk of death was lower for regional

anaesthesia than for general anaesthesia. Statistical investigation showed that mortality decreased significantly for one of the regional anaesthesia types, namely the combined peripheral nerve block technique, but not the other, neuraxial block technique (Karaca et al. 2012). On the other hand, it was shown in an earlier study that the choice of anaesthesia (regional or general) made no impact upon postoperative 30 day mortality (O'Hara et al. 2000).

Hospital-wide nurse staffing levels and in-hospital hip fracture patient mortality were shown to be associated in a large retrospective American study, where it was shown that the odds of dying in hospital decreased by 0.16 for each additional full-time equivalent registered nurse per patient day (Schilling et al. 2011). Another study, which included but was not exclusive to hip fractures, also concluded that nurse staffing levels influenced in-hospital mortality rates, as did high hospital occupancy on admission, seasonal influenza and being admitted at the weekend (Schilling et al. 2010).

Overall patient volumes handled by hospitals have also been shown to influence mortality, specifically for intertrochanteric hip fracture patients. It was shown that the risk of mortality was higher for hospitals which have fewer patients; specifically that hospitals with surgeons who treated just two or three cases per year had significantly higher mortality than hospitals that employed the surgeons treating the highest volume of cases (Forte et al. 2010). A similar study showed that there was a significantly greater risk of death in-hospital for patients treated by surgeons treating less than seven cases per year, but overall hospital volume did not influence mortality. It was shown, however, to negatively impact upon postoperative infection rates, with lower volume hospitals experiencing greater rates of infection (Browne et al. 2009).

A large study of over a quarter of a million patients investigated the effect of hospital type on outcome prior to discharge. While the setting (urban/rural) and teaching status (teaching/non-teaching) were found to have very little impact on in-hospital outcome, it was concluded that age, male sex and, notably, an increased surgical delay, were risk factors for in-hospital mortality (Koval et al. 2011). However, an earlier study gave an opposing conclusion; type of hospital was shown to have a significant effect of in-hospital mortality, with teaching hospitals having a lower risk of death compared with urban community

hospitals. There was also a higher rate of mortality reported for rural hospitals (Weller et al. 2005).

The presence of cerebral dysfunction prior to operation for hip fracture was shown to increase the probability of death, while other influencing factors included older age and male gender (see Sections 4.4.3 and 4.4.4 respectively) (Miller 1978). This result is supported by the inclusion of mental state as a predictor for mortality, shown earlier in this chapter. Similarly, the relative risk of mortality for hip fracture patients who scored poorly on a mental status test was shown to be 2.3 times higher than for those who did not (Meyer et al. 2000), while the relative risk of death was reported as 6.96 higher for patients with poor mental status by another study (Alegre-López et al. 2005). Both of these latter studies also reported poor mobility pre-fracture as a predictor for mortality for hip fracture patients (see Section 4.4.1).

One study stated many of the same variables as given here as the most prominent variables influencing mortality, namely ASA grade, poor mental health, male gender and increasing age. Mental health was measured using the Short Portable Mental Status Questionnaire (SPMSQ), an easily administered ten item questionnaire designed specifically for assessment of the cognitive impairment of elderly patients (Pfeiffer 1975). The advantage of this was that the SPMSQ score could be used to provide additional information about the predicted survival time of patients (Söderqvist et al. 2009).

The level of rehabilitative care categorised by three groups, orthopaedic hospital, geriatric hospital or none (discharged home) was shown to have no impact on mortality or morbidity for proximal femoral fracture patients of normal mental status, while an improvement in Activities of Daily Living (ADL) score was seen across all three groups within six months post-fracture (Röder et al. 2003).

A large American study (324,988 patients) investigated the notion of the so-called "July effect"; that is, whether mortality rates differ by month for hip fracture patients. The relative risk of mortality was found to be 12% greater at teaching hospitals during July and August, compared with non-teaching hospitals (Anderson et al. 2009).

Finally, it has been shown that type of surgery (arthroplasty, dynamic hip screw or nails) does not influence the risk of mortality at one year, but that the probability of hip dislocation post-surgery was affected (Geiger et al. 2007). Another study measured whether type of surgery

(hip replacement or not) influenced six separate patient outcomes, including mortality, at six weeks, six months and one year post-discharge. No differences were found for any of the outcomes for any of the three follow-up periods (Burns et al. 1999). However, Bhandari et al reached the opposite conclusion after performing a review of the literature. A trend towards an increased relative risk of death within four months was shown for arthroplasty compared with internal fixation, while increases in infection rates were also reported (Bhandari et al. 2003).

A particularly interesting result has been reported by a Japanese study, which investigated the relationship between length of stay and mortality. It was shown that a shorter length of stay was associated with in increased risk of mortality, and it is suggested that this is due to patients being discharged to a rehabilitation unit before they are really ready (Kondo et al. 2010b). A similar study comparing this relationship between Japan and the United States of America concluded that for every additional ten days spent in hospital after surgery, the risk of dying was reduced by 26% (Kondo et al. 2010a).

4.6 Odds ratio, relative risk and Chi-square analysis

A particular area of interest is looking at to what extent, if any, operative delay is related to the mortality of a patient. This was not considered in the previous section as it is inspected to a much greater detail here. There are two definitions to be used here for operative delay; more than one day and more than two days. Results between the two can then be compared. Using traditional convention for these types of analyses, the 'exposed' group is defined as those patients who do not have an operation within the specified time. An introduction to these techniques is given forthwith. Observed numbers quoted throughout this section may not be consistent due to missing data; the maximum number of observations available here was 2109.

Previous analysis has classified a delay to operation as a wait of more than two calendar days. This was done to keep in line with the majority of other studies on this topic, published guidelines and the advice of clinicians. As a comparison, this section also extends to looking at a delay of more than one calendar day.

An odds ratio (OR) describes the association between two binary data values; that is, whether the probability of a certain event occurring is the same for two groups. This type of calculation thus lends itself well to the investigations into delay here, where the two binary variables under consideration are delay and mortality. Indeed, odds ratios are used widely in medical reports (Bland and Altman 2000), in part due to this suitability to analyse mortality as well as the straightforwardness of calculation and interpretation. The odds ratio is the ratio of the odds of an event (a death) occurring in one group to the odds of it occurring in another group. An odds ratio of one therefore implies that the event is equally likely in both groups. It is limited by the lower bound of zero since it cannot return a negative value, but there is no upper limitation, resulting in a skewed distribution.

Consider the joint probability distribution of the binary random variables X and Y, as displayed in Table 4.6i. The probability p_{ij} represents the joint probability that X returns a value of i and Y returns a value of j; $i, j = 0, 1$. The odds ratio is defined as $\frac{p_{00}\,p_{11}}{p_{01}\,p_{10}}$. Counts may be used instead of probabilities, as seen in the subsequent calculations, and the same result is returned.

Table 4.6i: The joint probability distribution of two binary random variables, *X* and *Y*

	Y = 0	**Y = 1**
X = 0	p_{00}	p_{01}
X = 1	p_{10}	p_{11}

Note that any odds ratios returned here will not necessarily be equivalent to those given in a logistic regression. The logistic regression procedure requires complete data across all observations and included many more parameters, while an odds ratio requires just the two variables of delay and mortality.

The relative risk (RR) involves a similar calculation to that of the odds ratio, but will always return a smaller value. It is the risk of the event occurring, relative to the exposure; that is, the ratio of the probability of the event occurring in the exposed group versus the non-exposed group. The relative risk asymptotically approaches the odds ratio for small probabilities. Consequently if the event is not rare, then the odds ratio can overestimate the relative risk (Zhang and Yu 1998), resulting in what may be a misleading approximation to the relative risk (Davies et al. 1998). For this reason, both the OR and RR values are calculated. Using the notation introduced in Table 4.6i, the relative risk for the random variable X to occur (take a value of 1) is given by $\dfrac{p_{00}/(p_{00}+p_{10})}{p_{01}/(p_{01}+p_{11})}$.

Chi-square contingency tables are used to record the relationship between two or more variables in order to assess whether or not there is an association between variables. Here a 2 x 2 contingency test is used (i.e. two rows by two columns), since there are two variables, each of which can take two levels. Expected theoretical frequencies of each event are calculated using

$$E_{i,j} = \frac{\sum_{k=1}^{2} O_{i,k} \sum_{k=1}^{2} O_{k,j}}{N}; \; i, j = 0, 1$$

where $E_{i,j}$ represents the expected frequency in row i, column j, $O_{i,j}$ represents the observed frequencies in the same location and N is the total number of observations.

The test statistic is then calculated as

$$X^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i},$$

where O_i is an observed frequency and E_i is the expected theoretical frequency for each cell i, $i = 1,...,4$. Here X^2 asymptotically approaches a Chi-square $\left(\chi^2\right)$ distribution with one degree of freedom.

4.6.1 Three definitions of delay

It appears that there is some discrepancy in the literature on how a delay is classified and clearly this will have knock-on effects when statistical results are calculated. By creating three groups of delay definition, further investigation into the impact of this can be undertaken; these are (1) operation after two days of admission, (2) operation between one and two days and (3) operation within one day of admission. To clarify, consider a patient admitted on a Monday. If their surgery was performed on Monday or Tuesday, they would be in group (3), if it was performed on Wednesday, they would be in group (2) and if it was performed on Thursday or later then they would be in group (1). Results are presented in Table 4.6.1i.

Table 4.6.1i: Timing of operation against death on the acute ward

Frequency	Acute ward outcome		Total
	Dead	**Alive**	
(1) Operation after two days	121	740	861
(2) Operation between one day and two days	36	394	430
(3) Operation within one day	65	677	742
Total	222	1811	2033

The percentage of acute deaths seen within the group experiencing one day delay and the group experiencing one to two days delay are remarkably similar, 8.76% and 8.37% respectively, whereas deaths within the group experiencing a delay of more than two days are

14.05%. Overall there is a significant relationship here between delay and acute ward outcome, Chi-square $p = 0.0005$. However, if pairwise comparisons are made then some interesting results are found, see Table 4.6.1ii.

Table 4.6.1ii: Results when comparing the three definitions of delay for acute ward outcome

Comparison	Chi-square *p*-value
(1) *versus* **(2)**	0.0032
(1) *versus* **(3)**	0.0010
(2) *versus* **(3)**	0.8195

These results show that a cut-off of two days produces the most statistically different results when investigating an association between delay and acute ward outcome. There appears to be no difference if patients undergo surgery within one day or between one and two days, but after this very significant results are seen. A summary, using this cut-off, is given in Table 4.6.1iii.

The Chi-square contingency test here suggests that the two variables are not independent, $p = 0.0001$. The odds ratio is 1.734, with a 95% confidence interval of [1.310, 2.295] and the relative risk is 1.631, with a 95% confidence interval of [1.271, 2.092].

Each of these three results are in accordance with those previously presented and indicate that there is a significant association between mortality and whether patients are operated on within or after two days of admission.

Table 4.6.1iii: Frequency of operation within or after two days against death on the acute ward

Frequency	Acute ward outcome		Total
	Dead	**Alive**	
Operation after two days	121	740	861
Operation within two days	101	1071	1172
Total	222	1811	2033

4.6.2　Operation type and ASA grade analysis

This analysis is now repeated for each operation type and ASA grade grouping, see Tables 4.6.2i and 4.6.2ii.

Of the six operation types, three could be analysed singularly to adhere to the convention of requiring at least five entries in each cell for a Chi-square analysis. In one of these three cases, namely for type B (dynamic hip screw), the same conclusions of a statistically significant result were reached, whereas results for types D (intramedullary nail) and E (hemiarthroplasty) were insignificant, although this was borderline for operation type E at the 95% level of significance.

An operation after two days of arrival is shown to significantly increase the risk of mortality for ASA grade III patients, while for ASA grades I&II and IV, the two variables of acute ward mortality and operation within two days of arrival were shown to have no significant association. These results correspond with those found previously via logistic regression.

Table 4.6.2i: Frequency of operation within or after two days against death on the acute ward, by operation type

Operation type	n	Chi-square p-value	Odds ratio [95% C.I.]	Relative risk [95% C.I.]
B	743	0.0055	1.896 [1.201, 2.992]	1.758 [1.176, 2.627]
D	205	0.1788	1.789 [0.760, 4.213]	1.668 [0.786, 3.544]
E	743	0.0543	1.511 [0.990, 2.304]	1.428 [0.991, 2.057]

Table 4.6.2iii: Frequency of operation within or after two days against death on the acute ward, by ASA grade

ASA grade	n	Chi-square p-value	Odds ratio [95% C.I.]	Relative risk [95% C.I.]
I&II	751	0.2429	1.486 [0.761, 2.900]	1.456 [0.773, 2.741]
III	1061	0.0068	1.673 [1.149, 2.436]	1.573 [1.230, 2.191]
IV	190	0.5268	0.816 [0.434, 1.533]	0.870 [0.567, 1.335]

4.6.3   Data restrictions

There is some information available regarding mortality at four months post-discharge. However, this was not collected for the full time period for which data is available. For the range of dates when this was collected, it is only complete in around half of all cases, so any accurate analysis is not possible here. Future improvement in data collection would mean that this would be an interesting avenue to explore. UHB outcome is another recorded data item which again was not complete in many cases. Since the incident rate is relatively low, small frequencies of death can have a considerable impact on results.

4.6.4   Results from the literature

Conclusions in the literature are varied and wide-ranging, a selection of which is now discussed. Not all researchers use the same definition of delay and thus results are not necessarily comparable with those quoted earlier in this chapter. The period of follow-up is also inconsistent across studies.

Adjusting for background morbidity using the Charlson Index (Charlson et al. 1987), logistic regression was used to show that operative delay does influence one-year mortality, and separate odds ratios (all greater than one) were given for varying delay categories. In-hospital and one-month mortality were also shown to be influenced by delay (Novack et al. 2007).

A delay greater than four days has been shown to increase risk of death post-operation, while specifically for delayed patients, it was shown that the risk of death within 30 days was 2.5 times greater for patients delayed for medical reasons compared with those delayed for other reasons (Moran et al. 2005).

Looking at all patients as a whole, no association between delay (one day between admission and surgery) and one year mortality was found. However, splitting the patients by ASA grade showed that for patients with an ASA classification of I or II, operation after one day was associated with a significantly higher risk of death within one year, as well as an increased risk of post-operative complications (Verbeek et al. 2008), while the same

conclusion of lower mortality for surgery within 24 hours was shown by another study to hold true regardless of ASA classification (Hamlet et al. 1997).

However, Orosz et al showed no relationship between both two- and six-month mortality and whether or not surgery took place within 24 hours (Orosz et al. 2004), while there is also evidence of the opposite conclusion for the same measures (Dorotka et al. 2003b). Indeed, Dorotka et al showed that delay influenced mortality regardless of whether a cut off point of six, 12, 18 or 24 hours was used. There was no difference for patients operated on before or after 36 hours.

Risk of in-hospital death was shown to increase by 1.13 for a one day delay or higher and by 1.60 for a delay of two days or higher. The association between delay and mortality was also shown to be strongest for patients aged younger than 70 years old and with no comorbidities, but was independent of hospital type (Weller et al. 2005). Classifying a delay as a wait of more than 48 hours between admission and surgery, it was shown that in-hospital mortality was not associated with delay. However, predictors of mortality did include some variables discussed elsewhere in this chapter, including male sex and older age (Bergeron et al. 2006). However, an earlier study showed that an operative delay of more than two days approximately doubled the risk of death within the first postoperative year. After controlling for age, sex and the severity of pre-existing medical conditions, an increase in mortality for delayed patients was again found but it was not significant (Zuckerman et al. 1995).

Among other factors, a longer time to surgery was shown to increase the risk of death within 12 months. Delay was split into five categories, ranging from less than one day to more than ten days. It was shown that in order to yield one additional survivor, 25 patients waiting between one and five days would have to have their wait reduced to less than 24 hours (Elliott et al. 2003).

However, Hommel et al showed that overall mortality was not associated with timing of surgery (within/after 24 hours). In spite of this, it was shown that specifically for medically fit patients, one year mortality was significantly higher for patients who experienced an administrative delay compared with those not delayed, with mortality rates of 33% and 21% respectively (Hommel et al. 2008b). A separate study in Peterborough showed that a patient waiting more than 24 hours for surgery had no increased risk of death at 30 days than a patient operated on within 24 hours (Pathak et al. 1997).

Instead of operative delay, which was found to have no statistically significant effect on mortality, it was shown by Swedish researchers that the delay between the trauma occurring and admission to hospital did impact upon risk of death. Those arriving within six hours of injury had a 40% reduction in risk of death within one year post-operation compared with those arriving after six hours (Vertelis et al. 2009).

A study in Spain concluded that any association between timing of surgery and morbidity/mortality can be principally explained by medical conditions which cause the delay, but after adjusting for this it was still found that a delay over five days impacted upon mortality. This was measured for death in hospital (Vidán et al. 2011).

Operative delay was also not found to be a significant predictor of in-hospital mortality by Lefaivre at el, but a relationship was found between delay and the development of medical complications and the risk of pressure sores (Lefaivre et al. 2009).

Other studies have also concluded that the timing of surgery is not associated with mortality for hip fracture patients (Dolk 1990, Holt et al. 2010, Majumdar et al. 2006, Smektala et al. 2008).

4.7   Chapter summary

This chapter has focussed on determining which factors significantly influence mortality on the hip fracture ward and in particular operative delay has been investigated in great detail. This has proven to be a complex issue with a large number of variables affecting mortality with varying degrees of magnitude, which is an interesting result in itself.

These results are useful to the hip fracture team for similar reasons as those explained for length of stay; namely for planning purposes. In addition, the families of patients with an inflated probability of dying could, if protocol allowed, be given more advanced warning of this unfortunate event.

Again the matter of control must be discussed. Operative delay is something which can be changed with better management and/or an increase in resources or operating theatre capacity. Patients who are delayed have been shown to consistently be associated with an increased probability of dying which is a clearly a valuable result. This was also true when patients were divided into more homogenous groups with respect to a certain characteristic of interest. The choice of classifying a delay as a wait of more than two days was also investigated and it was shown that this was where significant results did lie.

The simulation model of the hip fracture ward needs to include some measure of patient outcome. This is explained in more detail in Chapter 6, where in fact discharge destination is considered in greater detail than just whether or not the patient survives. By completing the analyses presented here, decisions can be made about which factors should be used within the simulation model to segregate patients into distinct groups. ASA grade has been shown to be a very important variable here, as well as for length of stay, and thus is included later for modelling purposes.

CHAPTER 5: PRINCIPAL COMPONENTS ANALYSIS

5.1   Introduction

Principal components analysis (PCA) aims to reduce an original set of variables into a smaller set of components, without sacrificing important information contained within the data. The goal is for these new components to be uncorrelated but to represent most of the information contained in the original variables. If the original dataset has a large number of variables it may be difficult for any useful interpretation or conclusions regarding relationships between variables. However, by reducing the dimensionality of the data through this structural simplification, a few components are left to interpret rather than a large number of variables.

The analysis undertaken is concerned with the variance-covariance structure of a set of variables via the method of constructing these new artificial variables, known as components or factors. With a large mass of data, it is often difficult to visualise or comprehend the associations that exist between variables within a dataset. This may then be complicated further by the redundancy that can exist between the dimensions of the dataset, which leads to high levels of multicollinearity and correlation.

PCA seeks a linear combination of all of the original variables such that the maximum variance is extracted by the data. This variance will then be removed and a second linear combination is sought for which accounts for the maximum variance explained by the remaining variables. The process is completed until all variance is accounted for, thus the maximal number of principal components which could be found is equal to the number of variables.

In general, one may perform a principal components analysis to reduce a set of p original variables to m components, that account for most of the variance of the p variables. These m underlying components are inferred from the correlations among the p variables and are estimated as a weighted sum of the p variables. The ith component is thus

$$C_i = W_{i1}X_1 + W_{i2}X_2 + \ldots + W_{ip}X_p$$

where X_j represents variable j and W_{ij} represents the weight, or component loading, of the jth variable on the ith component, $j = 1, ..., p,\ i = 1, ..., m.$

Each of the p variables may also be expressed as a linear combination of the m components:

$$X_j = A_{1j}C_1 + A_{2j}C_2 + ... + A_{mj}C_m + U_j$$

where A_{ij} represents the weight of component i for variable X_j. U_j is the variance unique to variable j, variance that cannot be explained by any of the components.

The same data as per Chapters 2 to 4 was used here.

5.1.1 Data

It would be helpful if the information contained within the Cardiff Hip Fracture Survey dataset could be reduced in a similar way here. There are 27 variables available in total (refer to Table 3.1.1i and Table B3.2.1a of the Appendix for those variables under consideration) so it is difficult to get a picture of what happens to hip fracture patients as a whole. If these variables could be collapsed then it may make the situation easier to analyse. It is intuitive that there will be some relationship between some of these variables. For example, one might expect that a patient's walking ability and mobility score to be highly related; a person is very unlikely to have a low walking ability but be highly mobile. In Chapters 3 and 4, detailed investigation was undertaken with regard to length of stay and mortality. Interestingly, tests for multicollinearity between significant predictor variables showed only minor levels of any relationships based on recommendations relating to VIF and tolerance values. However, intuition dictates that there will be some association between variables and thus this is investigated in this chapter. Note the distinct difference here however that initially the variables are not assessed with regard to any dependent variable.

5.2   Principal components analysis on the Cardiff Hip Fracture Survey data

It was decided to perform an initial principal components analysis on the variables that describe a patient's condition and situation on, or soon after, their arrival.

The reason for this is as follows; when a new patient is admitted to the ward, the medical team will only have a certain amount of information available. This includes admission and patient details, as well as some medical diagnosis information. The latter may not be readily available but should present itself in the early part of a patient's stay. All of these factors will influence decisions made by the hip fracture team, as well as help them to plan for the future. While it may be interesting to look at later, obviously discharge, hospital stay and follow-up information is unavailable at this time, as these things are yet to happen. These aspects were also looked at in detail in Chapters 3 and 4 and therefore further scrutiny is not deemed necessary. The same variables that were used as predictors in those chapters are again used here.

The active dataset to be used for the principal components analysis has thus been reduced to containing 13 variables of the original 27, and contains almost 2000 observations. Since many of the variables are categorical here, a technique called CATPCA (CATegorical Principal Components Analysis) is employed.

5.2.1   Method

The statistical package SPSS (SPSS© 2007) was used for this analysis. Since both nominal and ordinal variables are being used, a normalisation procedure must be utilised to convert the numerical codes and scores into values that can be used for this type of analysis. There are several rotation options available, the most desirable here being variable principal, which is also the most commonly used rotation method. This coordinate rotation is used to align the transformed axes with the directions of maximum variance, without changing the relative location of points on the axes to each other.

The outcome of a principal components analysis includes the factor loadings matrix, which represents correlations that exist between the original variables and the new components.

Cross loading, where variables are highly correlated with more than one component, is undesirable and the rotation helps to avoid this.

The rotation option chosen optimises the association between variables, and since this is the overall aim of this analysis then it is the preferred option here. Its goal is to minimise the complexity of the components by making the large loadings larger and the small loadings smaller within each component; components are simplified by maximising the variance of the loadings across variables and within components. Principal components can either be devised from the correlation matrix, denoted by ρ, or by the covariance matrix, Λ. The method explained here uses the correlation matrix, which has the added benefit that results are easier to interpret; the component loadings are standardised across all observations to have a mean of zero and standard deviation of one. The covariance matrix is unstandardised and can be sensitive to scale differences across the variables, so is less appropriate here.

For the following analysis, other methods were also tried but proved to differ very little from using the variable principal normalisation approach.

The scaling level of each variable is another option that can be explored. For each ordinal categorical variable, it was decided that they would be classed as spline ordinal, as opposed to merely ordinal. The result of this is that the resulting transformation is a smooth polynomial, instead of a jagged fit which merely 'joins the dots'. Each variable carried an equal weight of one.

There are several methods used to determine the number of principal components to retain, two of which are considered here:

• **Kaiser's rule**
Each principal component produced by a PCA has an eigenvalue associated with it, calculated in the usual manner. Consider the eigenvalue-eigenvector pair (λ_i, e_i), calculated from the correlation matrix ρ of the original p variables, $X_1, ..., X_p$, and the m principal components, $C_1, ..., C_m$. Kaiser recommended that only components with an eigenvalue greater than or equal to one should be considered for inclusion (Kaiser 1959). This is akin to requiring each principal component accounting for at least as much variation as one of the

original variables, thus ensuring that no component is retained which is of less value than an original variable.

• **Proportion of variance**

A pre-specified amount of variation to be accounted for by the components may be decided, so that once the proportion of variance accounted for has been cumulatively reached by the components, no further components are considered. Alternatively a proportion of variation to be contributed by each component may be specified. The proportion of variation accounted for by the ith component, of a final m components, is calculated by $\frac{\lambda_i}{m}$.

Note that there is still scope for subjectivity here and these rules should not necessarily be strictly adhered to, particularly if results become uninterpretable. One may sacrifice retaining a higher proportion of variance in return for fewer components, for example, if it makes later analysis easier.

5.2.2   Results

Using the default value of two dimensions, both principal components were found to be significant using Kaiser's rule. The eigenvalues found were 3.606 and 1.897 for Component 1 and Component 2 respectively. Varying the number of dimensions to be returned resulted in more components with eigenvalues greater than one; however on closer inspection the additional components were not interpretable in terms of the factor loadings. Keeping more components would mean that a higher proportion of variation is retained but it was decided to not do this here in order to yield interpretable results. This meant that in total 42.3% of the original variance is accounted for by the new principal components; 27.7% by Component 1 and 14.6% by Component 2.

The factor loadings for the two retained components are given in Table 5.2.2i.

Table 5.2.2i: Component loadings

Variable	Component	
	1	**2**
Admfrom	0.674	-0.050
Walking0	0.885	-0.015
Walkaid0	0.605	-0.025
Mobility	0.876	-0.016
Mentalst	0.669	-0.083
WAASP	0.582	-0.075
Fractype	0.107	0.967
Optypenew_n	-0.098	-0.963
ASAnew_n	0.436	0.003
SexM	0.113	-0.045
Side	-0.003	-0.005
Age	0.466	-0.018
Opdelay	-0.081	0.129

Component 1 thus becomes:

$$\begin{aligned} \text{Component } 1 = {} & 0.674(Admfrom) + 0.885(Walking0) + 0.605(Walkaid0) \\ & + 0.876(Mobility) + 0.669(Mentalst) + 0.582(WAASP) \\ & + 0.107(Fractype) - 0.098(Optypenew\_n) + 0.436(ASAnew\_n) \\ & + 0.113(SexM) - 0.003(Side) + 0.466(Age) - 0.081(Opdelay) \end{aligned}$$

Component 2 is formulated in the same way.

It is easier to view these results graphically; a plot of the factor loadings of Component 1 against those of Component 2 can be seen in Figure 5.2.2ii. It can be seen that there are no variables which have a high negative loading on Component 1, while there are only two, or arguably three, variables which have a loading value of note on Component 2. Most variables are positively loaded on Component 1, with varying degrees of magnitude.

Since these values represent correlations, a general rule of thumb is to regard an absolute value greater than 0.6 as a strong correlation, absolute values between 0.4 and 0.6 as

moderate correlation, and anything less than an absolute value of 0.4 as weak correlation. Using this as a guide, five of the loadings on Component 1 would be classed as strong, three as moderate and five as weak. Component 2 shows just two strongly loaded variables, the remainder having small absolute values.
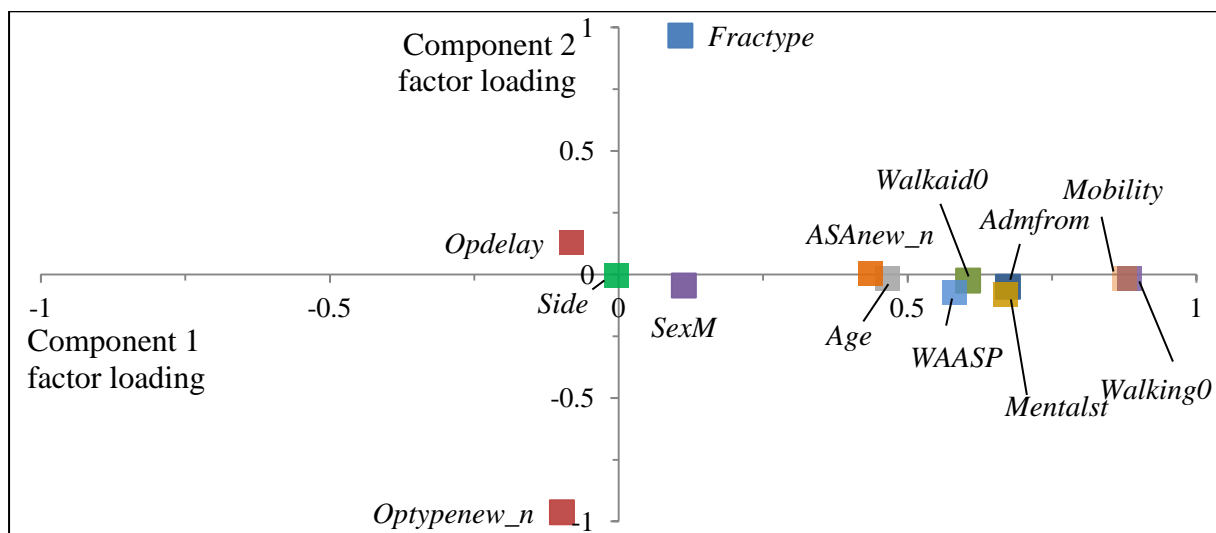
Figure 5.2.2ii: The factor loadings by variable of Component 1 against Component 2

Those variables which are highly loaded on the first component include walking ability pre-fracture, mobility score and walking aids used pre-fracture. This is intuitive; the variables that can be seen to be highly loaded on Component 1 would be expected to have some kind of positive correlation.

The nominal variables of fracture type and operation type seem to be working against each other; a high value for one would mean a low value for another. Despite the fact that these are coded as nominal variables in SPSS (thus rendering the number used to assign fracture and operation types meaningless), there is evidently still some relationship. Component 2 is therefore driven almost completely by the type of fracture incurred by the patient and the operation type performed on them. Looking at the frequencies in a cross-tabulation of these two groups confirmed this to some extent. A Chi-square test also indicated a significant relationship at the 5% level. However, the relationship is not very clear cut. Plotting each patient's component scores according to operation type shows tiers that correspond to these types, as in Figure 5.2.2iii.
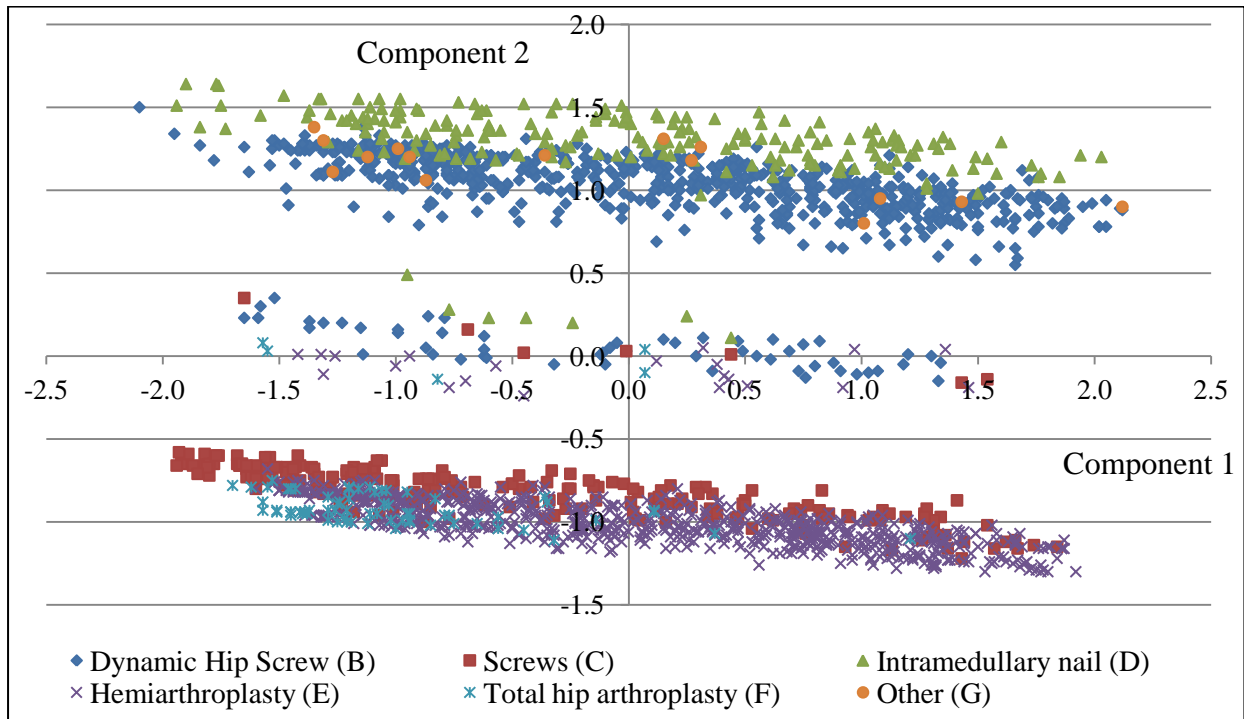
Figure 5.2.2iii: A plot of Component 1 against Component 2

It is also interesting to look at those variables which are weakly loaded on both components. It can be seen that side of fracture and sex are both very close to the origin in Figure 5.2.2ii and therefore are not guided by these components at all.

5.2.3 Conclusion

In conclusion, the data has been reduced to just two principal components, thus there has been a rather considerable decrease in dimensionality. However, in doing this, only 42.3% of the original variation in the data has been accounted for, so a trade-off between collapsing the dataset and retaining as much information as possible can be seen.

The first component retained seems to represent the medical condition of the patient, in particular with regards to aspects relating to how mobile the patient is, while the second component relates to the injury sustained and treatment given.

5.3 Principal components regression

Using the principal components found earlier, a regression analysis was performed to assess whether this reduced dataset could be used to predict length of stay and mortality. The dependent variables thus become length of stay and acute ward mortality, while the independent predictors are now the two components; each patient will have a score for both components, based on their values for the original 13 variables.

Note that multicollinearity is not an issue here since the predictor variables (the components) must be independent and uncorrelated by definition of principal components analysis methodology.

5.3.1 Length of stay

• **All data**

Initially all data was grouped together for this analysis. The natural logarithm of length of stay was taken in order to fulfil the assumptions of linear regression. A residual plot displays that the assumption of random error is satisfied, see Figure 5.3.1i.
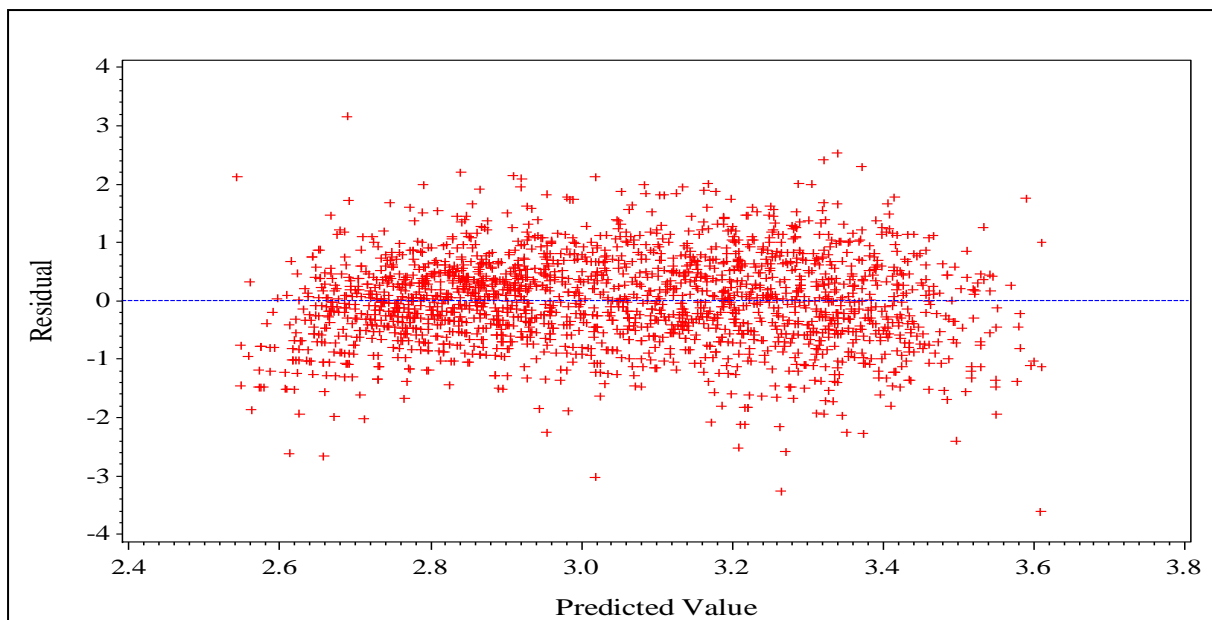
Figure 5.3.1i: Plot of *Predicted Value* against *Residual* for the principal components regression model

A plot of the predicted against the observed values produces the scatter plot seen in Figure 5.3.1ii. It can be seen that there is some positive trend, so a regression is therefore viable. However, a lot of variation can still be seen and the trend is only marginal, so the regression may not produce a particularly good fit. However, running the regression analysis produced a significant model according to the ANOVA statistic ($p < 0.0001$) and the parameter estimate results obtained are given in Table 5.3.1iii.
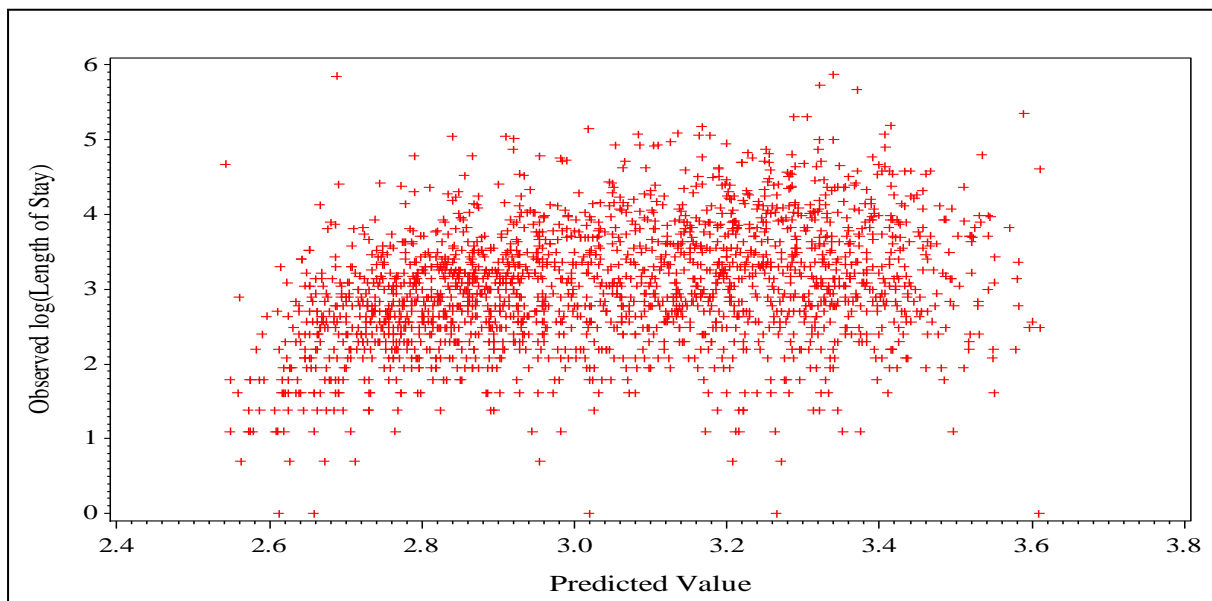
Figure 5.3.1ii: *Predicted Value* against *Observed* for the principal components regression model

Table 5.3.1iii: Parameter estimates given by the principal components regression model

Variable	Parameter estimate	Standard error	F-value	$\Pr > F$
Intercept	3.0500	0.0174	30876.4	< 0.0001
Component 1	0.2472	0.0176	197.33	< 0.0001
Component 2	0.0413	0.0170	5.94	0.0149

It can be seen that all parameters are highly significant and thus the model becomes:

$$\ln(LoS) = 3.05 + 0.25(\text{Component 1}) + 0.04(\text{Component 2})$$

$$\text{or, } LoS = e^{3.05 + 0.25(\text{Component 1}) + 0.04(\text{Component 2})}.$$

However, the R^2 value of 0.0933 shows that this is not a particularly good fit by this measure, as expected on inspection of Figure 5.3.1ii.

• Analysis by type of operation

Despite the fact that Component 2 was found to be significant, it was decided to rerun the analysis excluding this component since it essentially was comprised of the type of fracture and operation performed. The new analyses would therefore be done within each type of operation.

There are six different types of operation here so results are not discussed for each case. Instead operation type C, Screws, is focussed upon. This left 242 observations available for analysis. The following two plots show a random scatter of errors and a positive regression trend respectively, suggesting that linear regression is a valid tool to be used here.
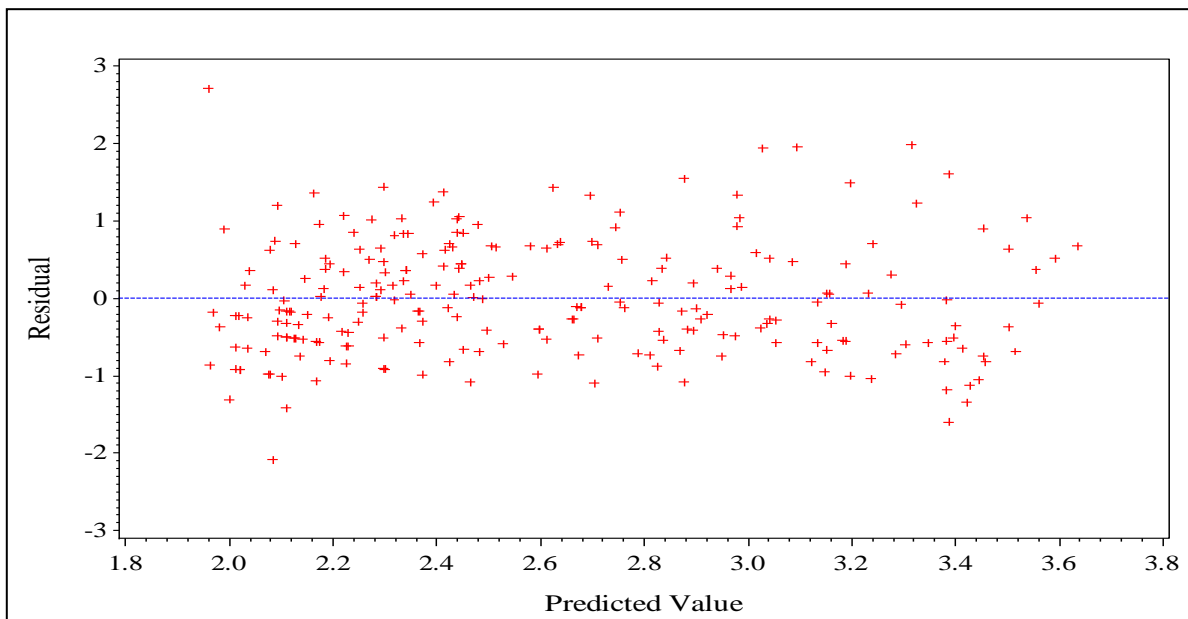
Figure 5.3.1iv: Plot of *Predicted Value* against *Residual* for the principal components regression model, operation type C only
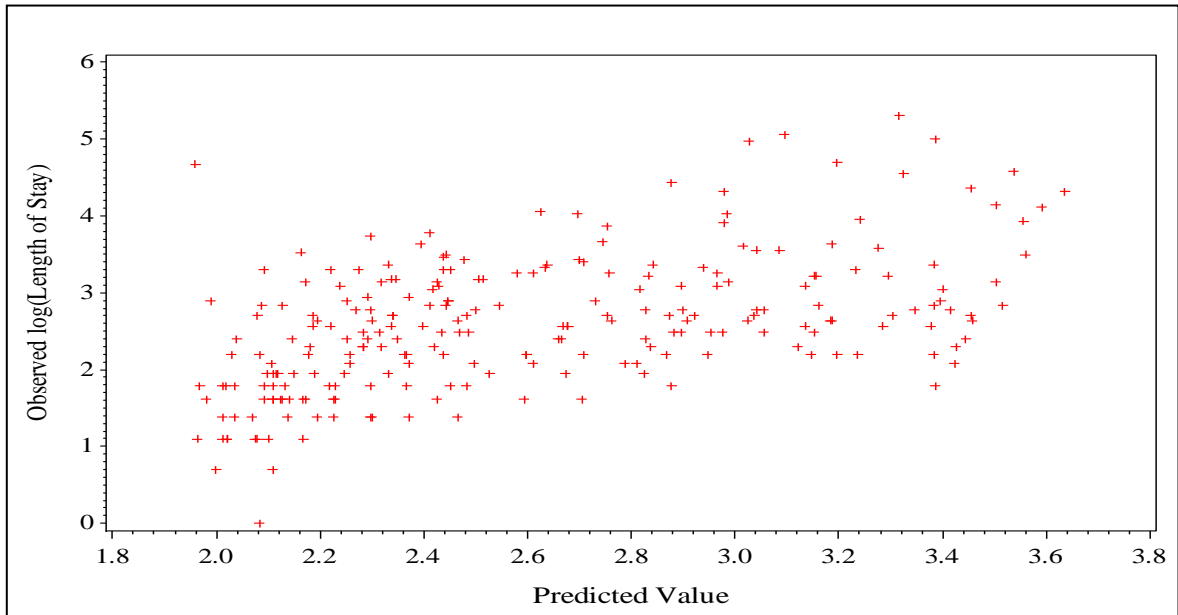
Figure 5.3.1v: *Predicted Value* against *Observed* for the principal components regression model, operation type C only

The *F*-test again showed that the regression model was significant ($p < 0.0001$); results of the regression model are given in Table 5.3.1vi.

Table 5.3.1vi: Parameter estimates given by the principal components regression model, operation type C only

Variable	Parameter estimate	Standard error	*F*-value	Pr $> F$
Intercept	2.1898	0.0526	2877.66	< 0.0001
Component 1	0.4435	0.0466	90.68	< 0.0001

The regression model is therefore

$$\ln(LoS) = 2.19 + 0.44(\text{Component 1})$$
$$\text{or, } LoS = e^{2.19+0.44(\text{Component 1})}.$$

A considerably better R^2 value of 0.2742 is now observed. It appears that removing Component 2 from the analysis and running a regression analysis within each type of operation yields better results. This was therefore repeated for each other type of operation and the results are seen in Table 5.3.1vii. Note that these are the results to predict the natural

logarithm of length of stay. No model could be found for operation type G, Other, but all other models were found to be significant.

Table 5.3.1vii: Summary of results for principal components regression analysis, by type of operation

Operation Type	n	Intercept estimate	Estimate of constant for Component 1	R^2
Dynamic hip screw (B)	728	3.0747	0.2030	0.0638
Screws (C)	242	2.1898	0.4435	0.2742
Intramedullary nail (D)	202	3.2047	0.2295	0.0791
Hemiarthroplasty (E)	720	3.1189	0.1323	0.0248
Total hip arthroplasty (F)	74	3.0912	0.4458	0.1037
Other (G)	15	-	-	-

It can be seen that the success of each model, using R^2 as a measure, varies between operation types. In particular, operation type E has a very low R^2 value and, unlike with operation type C, nothing can really be gained from this analysis.

5.3.2 Mortality

Logistic regression was employed here, as per the methodology described in Chapter 4, to assess whether the principal components can be used as predictors for acute ward mortality.

The various validation procedures described and undertaken previously were used here and, where a model is quoted later, each time indicated that the model provided a significant fit.

• **All data**

Component 2 was not found to be a significant variable when performing logistic regression on all of the data, so was excluded. The parameter estimates and their associated significance levels are seen in Table 5.3.2i, with odds ratio information in Table 5.3.2ii.

Table 5.3.2i: Parameter estimates for the principal components logistic regression model

Variable	Parameter estimate	Standard error	Wald Chi-Square	Pr > ChiSq
Intercept	-2.3135	0.0872	703.43	< 0.0001
Component 1	0.8341	0.0829	101.14	< 0.0001

The model thus becomes

$$\text{logit}(\pi) = -2.31 + 0.83(\text{Component 1})$$

$$\text{or, } \pi = \frac{1}{1 + e^{-\{2.31 + 0.83(\text{Component 1})\}}}$$

where π represents the probability of dying on the acute ward, as previously. The maximum re-scaled value of R^2 was found to be 0.1141.

Table 5.3.2ii: Odds ratios for the parameter estimates of the principal components logistic regression model

Variable	Adjusted OR	95% Confidence Interval for OR
Component 1	2.303	[1.957, 2.709]

For every unit increase in the score for Component 1, the odds of dying on the acute ward more than double. Component 1 captured various aspects of patient medical condition and treatment, with a higher score indicating poorer medical fitness.

• Analysis by type of operation

This analysis was repeated by type of operation and results are included here again initially for operation type C, see Table 5.3.2iii. The maximum re-scaled R^2 achieved in this case was 0.3528 and odds ratio results are given in Table 5.3.2iv.

The model thus becomes

$$\text{logit}(\pi) = -3.87 + 2.01(\text{Component 1})$$

$$\text{or, } \pi = \frac{1}{1 + e^{-\{3.87 + 2.01(\text{Component 1})\}}}.$$

Table 5.3.2iii: Parameter estimates for the principal components logistic regression model, operation type C only

Variable	Parameter estimate	Standard error	Wald Chi-Square	Pr > ChiSq
Intercept	-3.8719	0.6428	36.28	< 0.0001
Component 1	2.0055	0.5382	13.88	0.0002

Table 5.3.2iv: Odds ratios for the parameter estimates of the principal components logistic regression model, operation type C only

Variable	Adjusted OR	95% Confidence Interval for OR
Component 1	7.430	[2.587, 21.336]

For every unit increase in the score for Component 1, the odds of dying on the acute ward increase by 7.4.

This analysis was repeated for each other type of operation and results are summarised in Table 5.3.2v. No model could be estimated for operation type G since all patients had the same outcome of survival. Component 1 was not found to be a significant estimator for operation type F and therefore no $\tilde{R}^2$ value is given. As it was seen with length of stay, particularly good fits are not gained for each other type of operation.

Table 5.3.2v: Parameter estimates for the principal components logistic regression model, by type of operation ($\tilde{R}^2$ – maximum re-scaled value of R^2)

Operation Type	n	Intercept estimate	Estimate of constant for Component 1	$\tilde{R}^2$
Dynamic hip screw (B)	728	-2.4812	0.9461	0.1335
Screws (C)	242	-3.8719	2.0055	0.3528
Intramedullary nail (D)	202	-2.0943	0.5483	0.0501
Hemiarthroplasty (E)	721	-2.0000	0.6530	0.0718
Total hip arthroplasty (F)	74	-3.5835	-	-
Other (G)	15	-	-	-

5.4   Chapter summary

While CATPCA was useful in terms of collapsing the dataset, it has not proved to be especially useful in the further analysis seen here. However, it may be argued that sometimes finding out what cannot be achieved is as insightful as finding out what can be achieved. While a patient's overall physical and medical condition may be captured by Component 1, collapsing these into one variable does possibly oversimplify the problem. It has therefore become apparent that length of stay and mortality cannot be predicted particularly accurately by this simplification, but this highlights a problem inherent in healthcare.

When all variables were included (Chapters 3 and 4) and thus the dataset was not simplified at all, a particularly good fit could still not be found. While less useful practically, it is still interesting to note that this is an area of considerable difficulty in a retrospective statistical manner, as well as for the healthcare practitioners working in real time. Even when all the information is known, there are still problems in this area. Two patients who may be equal 'on paper' can still perform differently in terms of length of stay and mortality. The regression equations given may therefore not be completely accurate in terms of being able to substitute values in to gain an entirely correct prediction of a length of stay or mortality, but they do at least give some insight into which factors are important in these areas, while also providing the insight that these patient outcomes are difficult to predict!

CHAPTER 6: MODELLING THE TRAUMA HIP FRACTURE WARD

6.1   Introduction

Previous chapters have focussed largely on statistical data analyses. In this chapter, insight gained from these analyses, alongside additional investigations, is used to model the trauma hip fracture ward at the University Hospital of Wales. (In fact, specifically it is the patients who are modelled; a hip fracture ward, as such, does not exist physically.)

Using the information and conclusions drawn, two discrete event simulation (DES) models were built, referred to here as Model I and Model II. The process undertaken for this is discussed in detail, alongside verification and validation procedures. Though similar in some aspects, the two models also have some distinct differences, brought about due to extra data being made available throughout the course of this study. Once the models were deemed fit for purpose, they were used to analyse a number of what-if situations. This chapter focuses on these models and how changes within the system may affect its performance. Key outcomes include bed occupancy and patient discharge.

6.2 The suitability of simulation for this project

Before the DES model building process is explained in detail, first it must be considered why simulation is an appropriate tool to use here.

Simulation models are built to represent the main features of a system. While a simulation cannot fully imitate a complex system, once it is validated and believed to accurately represent the system, then it can be used to explore different strategic or tactical management options or to gain a better understanding of how it works.

The reasons and benefits of using simulation are manifold. Mathematical models are often too intractable or too simple and the system may be too complex to analyse using this method. Parameters are easy to amend in a simulation to assess their impact on results of interest, while real-life changes may not be as safe or as viable, particularly in a limited timescale.

While randomness is undesirable in any system, it is inherent in healthcare and capturing this variability is paramount to designing a useable and realistic model (Davies 1985, Harper and Shahani 2002, Lee et al. 2003, Utley et al. 2005). In this case, one example is that arrivals to the hip fracture ward cannot be predicted deterministically. While estimates of an arrival can be made based upon historical data, the number of arrivals on any given day, or indeed the inter-arrival time between patients, is not known until these events actually occur. The frequency and duration of events may only be known probabilistically, and thus a stochastic model is required. Discrete event simulation can handle this stochastic behaviour. The variation in the system is taken into account by taking samples from appropriate probability distributions. Since a single run of the simulation is therefore a sampling experiment, replications are required.

The objective here was to use DES to model trauma hip fracture patients at the UHW. Simulation is a technique that is apt to this objective due to the benefits described previously. The purpose of this exercise is to identify and investigate factors which are important to what happens on the ward.

6.2.1 Visual Basic for Applications

Both of the simulation models used here were built using Visual Basic for Applications (VBA) for Excel. There are a number of advantages of using this computer program and consequently it was chosen as an appropriate tool here.

One of the main benefits of using VBA is that the simulation model can be run by anyone familiar with Microsoft Excel, a very popular and globally-available computer package. Indeed, another user could modify the input parameters, or even the code, on which the model is based. Another advantage is the abundant flexibility provided by VBA and Microsoft Excel, both in terms of model design and results collection. Finally, no specialist software is required; users of the model are likely to be from the healthcare profession and are unlikely to have access to other, more specialist, simulation software packages.

6.3 Model I formulation

This study originally came about as clinicians within the hospital were concerned about the treatment that this patient group were receiving. Some patients were experiencing a considerable delay between admission and operation, which is not only distressing for the patient but is shown to have detrimental knock-on effects (see Chapters 3 and 4). Operative delay would therefore be incorporated into the model where appropriate. The model was developed for two primary reasons. Firstly, model development inherently means a critique of the current system and its data. Secondly, parameters within the model can be adjusted in order to safely investigate the impact of system changes in a timely manner.

6.3.1 The National Hip Fracture Database

In March 2007, the data collection procedure at the UHW for hip fracture patients was adjusted. Many of the same variables were collected, but the database was brought in line with the National Hip Fracture Database (NHFD). NHFD is a joint initiative between the British Orthopaedic Association (BOA) and the British Geriatrics Society (BGS) which, amongst other objectives, aims to collect continuous comparative data to create a drive for sustained improvements in clinical standards and cost effectiveness. Data collected includes information regarding case-mix, process and outcome.

This new process of data collection meant that the number of hours a patient was delayed getting to theatre could now be identified where previously only the number of days could be calculated. This extra data became available throughout the duration of this study and therefore was not considered in previous chapters. A total of 1223 patients were available in this dataset. Note that the value of n may not be consistent throughout this analysis as each data item may not be available for every observation.

The British Orthopaedic Association state that "All patients with hip fracture who are medically fit should have surgery within 48 hours of admission, and during normal working hours" (BOA 2007). For this reason, along with guidance given by medical experts involved in the project, a cut-off value of 48 hours is used. Recall that a significant operative delay has been regarded as one greater than two days throughout much of this thesis, so choosing 48 hours is also sensible for consistency reasons.

6.3.2   Patient classification

Looking at all patients as a whole makes any analysis too general. An agreed method of classifying patients was to split by ASA grade and operative delay. ASA grade is the best way to categorise patients according to medical fitness, while operative delay is a key focus of this study. The decision to make these splits is endorsed by results from Chapters 3 and 4; in particular see Sections 3.4.2, 3.4.5, 4.4.2 and 4.6. Chapter 5 informed model development to a lesser extent, but results do confirm that medical fitness (Model I) and operation type (Model II) are appropriate splitting variables.

A particular feature of interest is patient length of stay, which is now further investigated and shown graphically using survival curves. Consider the survival distribution function, $S(t)$, usually used to describe the lifetimes of a population of interest. Let T be the lifetime of a randomly selected experimental unit within the population, then the survival distribution function evaluated at t represents the probability that the experimental unit will have a lifetime which exceeds t; that is, $S(t) = \text{Probability}(T > t)$.

The probability that their lifetime does not exceed t is given by the cumulative distribution function, denoted $F(t)$, where $F(t) = 1 - S(t)$. The probability distribution function, $f(t)$, takes its usual form; $f(t) = \dfrac{\mathrm{d}F(t)}{\mathrm{d}t}$. $h(t)$ is commonly used to represent the hazard function, given by $h(t) = \dfrac{f(t)}{S(t)}$.

As suggested by the name, this topic is usually used in conjunction with estimating mortality amongst a population but here it is utilised in a different and novel way. Instead of the 'event' under consideration being death, consider this instead to be the departure from the ward, thus the survival distribution function estimates the probability that a patient will have a length of stay exceeding t, using the same notation.

The results of the regression analysis undertaken in Chapter 3 indicated that delay to operation was a significant predictor of length of stay. ASA grade was initially not found to be a predictor in this instance but it was decided in agreement with medical experts that this should still be included; see Figure 6.3.2i for the survival curve.
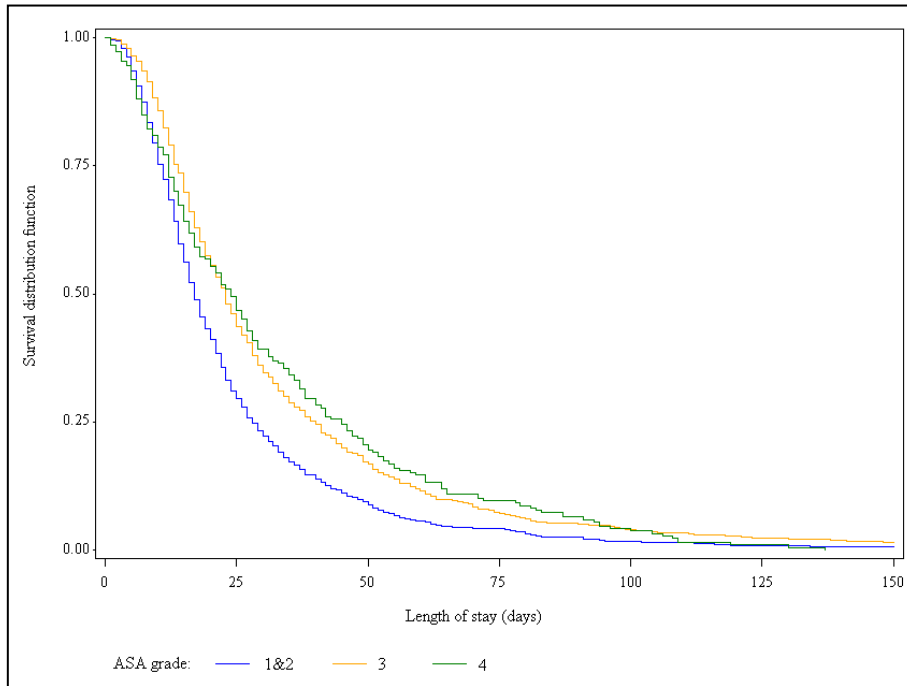
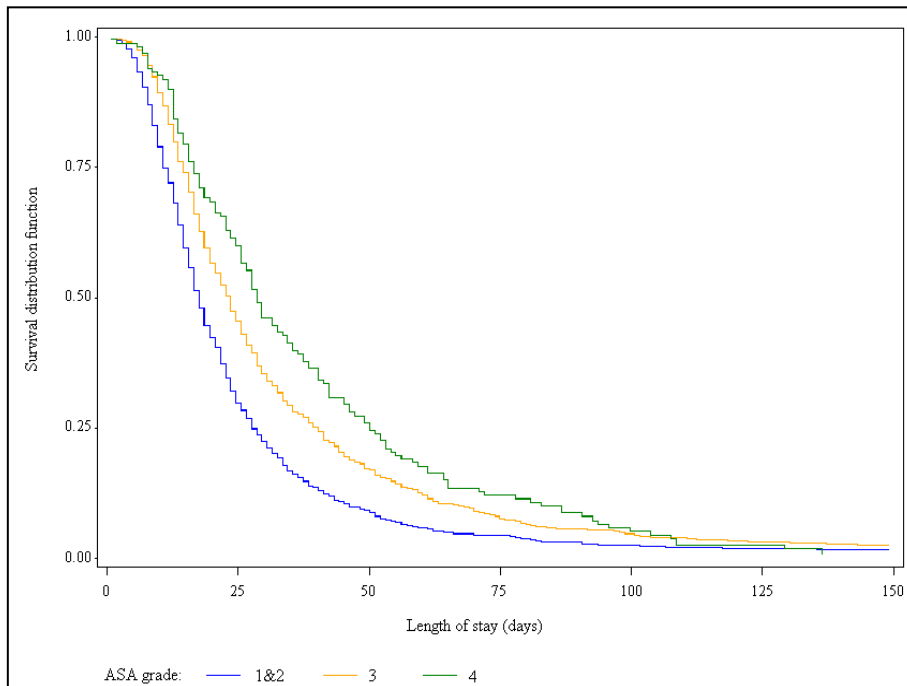**Figure 6.3.2i:** Survival curve (LoS) for all patients by ASA grade

Figure 6.3.2ii: Survival curve (LoS) for surviving patients by ASA grade

These results may be influenced by mortality rates due to the differences in medical fitness between these patient groups so analysis was repeated using a subset of the original data

158

which only included patients who survived their acute stay in hospital. Results are displayed in Figure 6.3.2ii.

For both of these analyses, the survival analysis returned significant differences between the groups with respect to length of stay, $p < 0.0001$. A clearer distinction between groups is seen when just surviving patients are considered, in comparison with all patients. The difference in length of stay between ASA grades is therefore evident and will be included in the model where appropriate.

6.3.3 Operative delay in hours

Summary statistics on the number of hours delayed before surgery, categorised by ASA grade, are given in Table 6.3.3i ($\leq$ 48h – operation within 48 hours; > 48h – operation after 48 hours). The percentage of patients entering theatre within 48 hours was 58.1% for ASAI&II patients, 48.1% for ASAIII patients and 28.7% for ASAIV patients. It is clear that fitter patients enter theatre quicker.

Table 6.3.3i: Summary statistics on operative delay (hours) for all patients

ASA Grade	Delay	n	Mean	S.D.	Minimum	Maximum	Skewness	Kurtosis
I&II	All	353	56.64	65.71	3	791	5.61	50.29
	$\leq$ 48h	205	25.43	12.41	3	48	0.25	-0.99
	> 48h	148	99.88	82.98	49	791	5.11	35.58
III	All	559	63.53	49.43	1	368	2.09	6.67
	$\leq$ 48h	269	29.01	11.84	1	48	-0.01	-1.10
	> 48h	290	95.56	49.49	49	368	2.23	6.78
IV	All	87	102.70	111.05	5	835	4.11	22.59
	$\leq$ 48h	25	31.96	11.16	5	47	-0.20	-0.35
	> 48h	62	131.23	120.24	50	835	3.95	19.53

6.3.4 Investigating the causes of delay

Another aspect of the NHFD is that the reason for operative delay, both at 24 and 48 hours, is recorded. There are nine categories for each of these fields, reduced to three here for

simplicity; in general the reasons for operative delay can be segregated into one of two classifications; system-related or medical-related (Shiga et al. 2008). One category of 'Other' is ignored; this had a very small number of entries and does not aid the analysis. The three reduced categories are:

(i)   No delay; the patient has their operation within the relevant time limit;

(ii)   Clinical delay; the patient is not currently medically fit enough to undergo an operation within the relevant time limit;

(iii)   Administrative delay; no theatre space, for example.

The percentage of patients not experiencing a delay at 48 hours was given in Section 6.3.3. Of the remaining patients, 31% of delays were due to clinical reasons and 69% to administrative reasons.

Results relating to delay reason in the literature are inconsistent. For example, in one study 56% of delays were directly attributable to medical problems and 20% were due to patients awaiting medical investigations, compared with administrative reasons such as lack of theatre space causing 24% of delays (Charalambous et al. 2003), while another reported that a much larger proportion of delays (69%) were due to lack of theatre space (Petermann et al. 2003). Figures comparable to those reported for the UHW were found by a group in Spain, with 61% of delays over 48 hours due to non-availability of an operating room and 33% due to medical problems (Vidán et al. 2011). Other results include 59% of patients being delayed due to waiting for medical review, compared with 29% waiting for organisational reasons (Youde et al. 2009), while respective percentages of 52% and 29% have also been reported (Orosz et al. 2002).

Reducing administrative delay is not easy. Dy et al reported on two treatment strategies, compared in order to investigate factors influencing timing of surgery. It was found that a systems-based solution can be cost-effective in minimising delay, via the use of a dedicated on-call support team (Dy et al. 2011). Taking an x-ray during triage is another way of reducing delay (Chia et al. 2008).

Logistic regression was utilised to determine and quantify the predictors of a surgical delay, with the target variable of surgery after two days. Age, ASA grade and admission day were among the significant predictors found, but there were no predictors over which the hospital

has any control (Fantini et al. 2011). The presence of comorbidities has also been shown to explain delays over 48 hours (Bergeron et al. 2006). Patients using the anticoagulant drug warfarin have also been shown to experience a longer delay, which is reported to most likely to be due to comorbidities. Patients taking this drug experienced almost double the surgical delay of non-users (23 compared with 12 hours), and it was also a predictor of longer hospital stay (Ranhoff et al. 2011). Time to surgery has also been shown to be dependent upon the method used for the cessation of warfarin prior to surgery, which is required to reverse anticoagulation in order that the operation can be performed. Simply stopping taking the drug led to a two day increase in delay compared with cessation and additional pharmacological treatment (4.4 days compared with 2.4 days respectively) (Ashouri et al. 2011).

An investigation into whether the place of fall (outside home, at home, residential/nursing home, hospital inpatient) had any relationship to the time to the commencement of specialist hip fracture treatment found that, rather unexpectedly, patients already under maximal healthcare treatment had to wait the longest time for referral (Khan et al. 2011).

Inspection of the data showed that there were some patients in the group ASAI&II who did not undergo surgery within 48 hours due to clinical reasons. However, this is a rather small group and therefore for the purposes of the model they are ignored. All delayed patients in ASA group I&II are assumed to experience administrative delays only. This decision was also undertaken on the advice of a senior clinician involved in the project. Figure 6.3.4i shows that almost two-thirds of all delayed ASA grade III patients are still waiting at 48 hours due to administrative reasons, reducing to just under a third in ASA grade IV patients.

Clinical delays are justifiable and are not easily reducible by tightening routines and for modelling purposes it is assumed that these delays are confined to patients with an ASA grade of III or IV only. These unavoidable delays can be used to gain improvement in the clinical condition of the patient (Buck et al. 1987), but on the other hand, chasing unrealistic medical goals should not lead to delay (SHFA 2008).

Administrative delays are not justifiable and are reducible and, on the advice of clinicians, is something that crucially could be modelled as being identical for the ASA I&II group and the

ASA III and ASA IV groups. Thus clinical delay becomes the total delay for the relevant ASA group (III or IV), minus the total delay for ASA group I&II. This is important since once a patient is delayed for longer than 48 hours, the reason for their delay is not recorded but, by making this assumption, it can at least be surmised.

Using empirical data, it is decided whether a patient still waiting at 48 hours is delayed for a clinical or non-clinical reason. This is because those still delayed for clinical reasons were shown to have a significantly longer delay than those delayed for non-clinical reasons ($p <$ 0.0001).

There are therefore two categories for ASAI&II patients, operation within 48 hours and operation after 48 hours, while there three categories for each of ASAIII and ASAIV patients, delayed patients are further split by delay reason; see Figure 6.3.4i for a breakdown of how patient numbers within these categories are distributed.
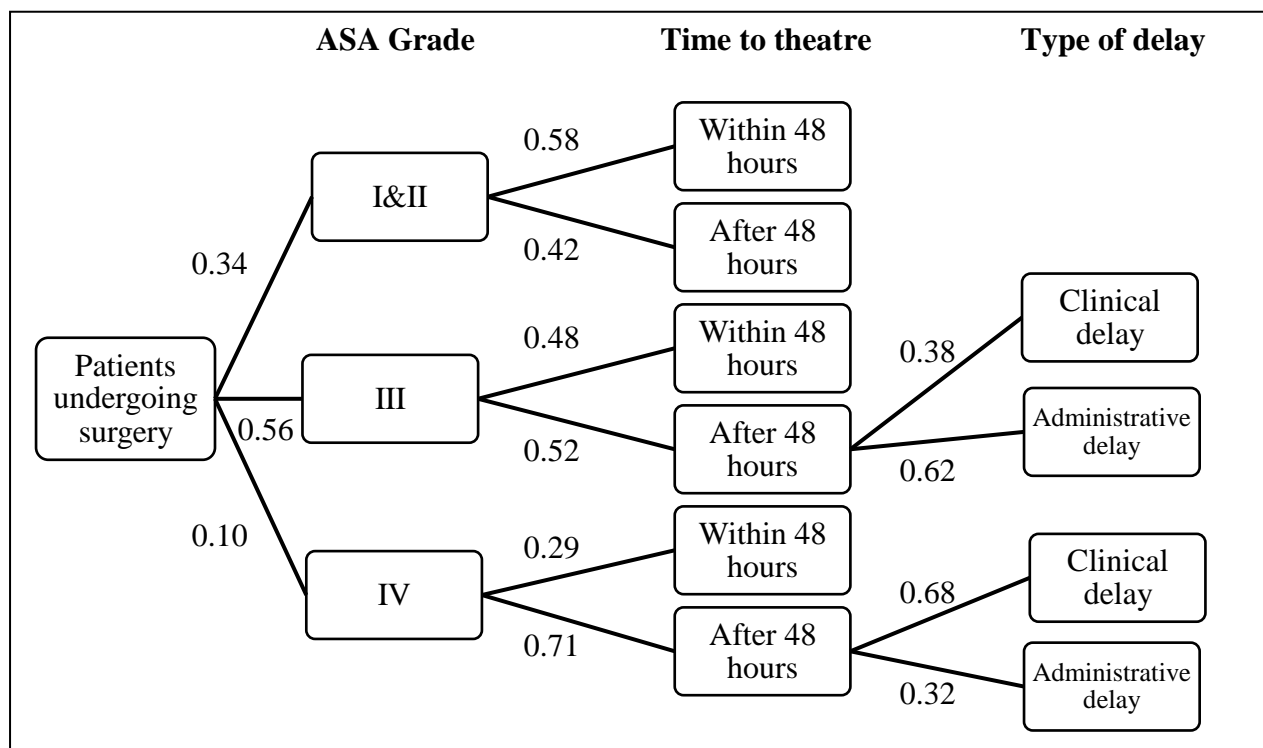
Figure 6.3.4i: Probability pathway for ASA grade and operative delay, for patients undergoing surgery

162

Each patient in ASA groups III or IV, who is clinically delayed at 48 hours, has their total delay estimated and then an administrative delay removed; the remainder being the clinical delay and must exceed 48 hours. To avoid the potential problem of the sampled administrative delay being greater than the total delay, the random number used to determine the time spent administratively delayed is scaled to be in the range $[0, r]$, where $r = F_x(t) = P(X \leq t)$. The earlier sampled value of the total delay is represented by t and F_x represents the cumulative distribution function (CDF) of the time spent administratively delayed.

The distribution of the number of hours delayed for ASAI&II patients was found to follow a Lognormal distribution; the fit can be seen graphically in Figure 6.3.4ii while maximum likelihood estimates and some comparative empirical figures are given in Table 6.3.4iii. Recall that the value removed from the total delay for ASA grade III and IV patients comes from a curtailed version of this distribution.
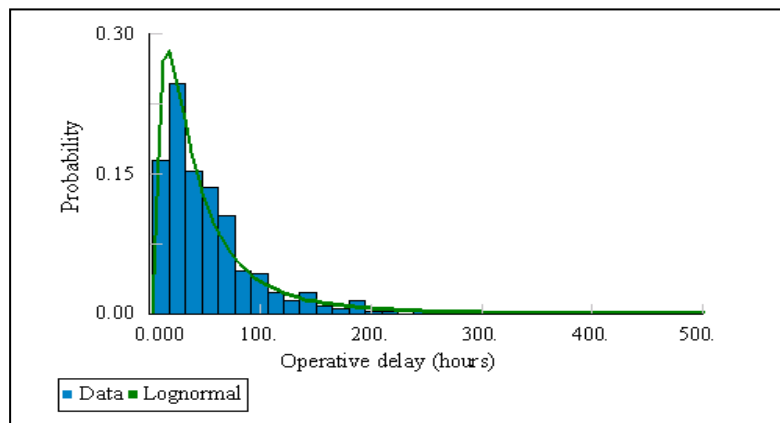
Figure 6.3.4ii: Distribution of operative delay in hours for all ASAI&II patients against the Lognormal distribution with parameters min = 3, $\mu = 3.523$ and $\sigma = 0.954$

Table 6.3.4iii: Lognormal fit for operative delay (hours) for all ASAI&II patients

Category	μ	σ	Min	Theoretical		Empirical	
				Mean	S.D.	Mean	S.D.
ASAI&II	3.523	0.954	3	56.1	64.2	56.6	65.7

For delayed patients, statistical fits could be found for each combination of ASA grade and delay reason, for the total time spent delayed. These are given in Table 6.3.4iv and all are for the Negative Exponential distribution. Each time there is a shift of 49, as this is the minimum value that can be achieved for these patients. (In one case the shift was forced to be this value since the minimum value from the data was not 49, and extreme outliers were removed where necessary.) Goodness-of-fit tests were passed in all cases.

Table 6.3.4iv: Negative Exponential fits for operative delay (hours) for patients operated on after 48 hours (Admin. – administrative)

ASA Grade	Type of delay	μ	Theoretical		Empirical	
			Mean	S.D.	Mean	S.D.
III	Admin.	38.01	87.01	38.01	87.01	38.33
	Clinical	66.60	115.60	66.60	115.60	62.74
IV	Admin.	40.06	89.06	40.06	89.06	25.73
	Clinical	93.81	142.81	93.81	142.81	145.20

The fits for ASAIII patients delayed at 48 hours for administrative and clinical reasons are given in Appendix D, Figure D6.3.4a. Note that graphs may have been curtailed for display reasons.

The operative delay for each patient is generated in this case via the inversion method, described as follows. (Note that the values for μ quoted in Table 6.3.4iv must be inverted so that they are in the same form as seen here.) The Negative Exponential distribution has a probability density function of $f(t) = \mu e^{-\mu t}$, where μ^{-1} is the mean of the distribution. Integrating with respect to t (time) over the interval $(0, \infty)$, the CDF can be found; $F(t) = 1 - e^{-\mu t}$. Setting $u = 1 - e^{-\mu t}$, where u is a random number between 0 and 1, yields

$$u - 1 = -e^{-\mu t}$$

$$1 - u = e^{-\mu t}$$

$$\ln(1 - u) = -\mu t.$$

Rearranging for t gives $t = -\dfrac{\ln(1-u)}{\mu}$. Since u is a random number between 0 and 1, $(1-u)$

is also a random number between 0 and 1, and so simplifying gives $t = -\dfrac{\ln(u)}{\mu}$.

The appropriate value of μ is then substituted into this expression and a value of 49 is added in order to gain the operative delay in the simulation.

Looking at the non-delayed patients, their distributions are required to be upper- and lower-bounded, which is where problems arise. However, by combining all patients by ASA grade led to more positive results. For both ASAIII and ASAIV, operative delay was found to fit the Gamma distribution, see Table 6.3.4v (results for ASAI&II were given in Table 6.3.4iii); maximum likelihood estimates were found using Stat::Fit and the fits are displayed graphically in Figures D6.3.4b and D6.3.4c. Clearly, the Gamma distribution provides a very good fit to the operative delay data, for both ASAIII and ASAIV patients. The slight discrepancies can be attributed to the 'jaggedness' of the actual data, which is to be expected especially with the ASAIV patients since this group only contained 87 data points.

Table 6.3.4v: Gamma fits for operative delay (hours) for all patients, ASAIII and ASAIV

Category	α	β	Min	Theoretical		Empirical	
				Mean	S.D.	Mean	S.D.
ASAIII	1.971	31.745	1	63.6	44.6	63.5	49.4
ASAIV	1.629	59.978	5	102.7	76.6	102.7	111.1

It is however not desirable to just use these distributions to estimate operative delay due to loss of generality. As described previously, the type of delay is also important. Therefore, if the sampling procedure produces a value less than or equal to 48 (hours), then the operative delay for that patient is taken to be that sampled value. However, if the value produced is greater than 48, then the patient is classed as 'delayed' but their operative delay value is discarded. Instead, they enter either the clinical or administrative delay group; the appropriate Negative Exponential distribution for that ASA grade and delay type is then sampled from. This excludes the group of ASAI&II patients, whose delay type is not

classified further. The original sampled value can therefore be used. To summarise for ASA grade III and IV patients, delay is calculated as per the steps given in Figure D6.3.4d.

In order for this method to be suitable, the distributions used must estimate the number of non-delayed patients accurately. By using the appropriate cumulative density function, the percentage of non-delayed patients predicted by the distribution can be calculated. 58% of ASAI&II patients were operated on within 48 hours, while the result from the appropriate Lognormal distribution is 62%. For the Gamma distributions of ASAIII and ASAIV, the results are 45% and 26% respectively. These compare correspondingly with 48% and 29% from the data. It can be seen that reasonable approximations are gained using this approach and therefore the method is deemed fit for purpose.

6.3.5   Inter-arrival times

After implementation of the NHFD, the distribution of inter-arrival times could be calculated in terms of hours, as displayed by hour in Figure 6.3.5i. As commonly seen with inter-arrival times, particularly relating to unplanned hospital admissions, the shape is approximately Negatively Exponential in shape (Moore 2003), but no statistical fit could be found to represent this data. A two-period moving average trendline is overlaid to further indicate the awkwardness in the general shape of this distribution.

For the purpose of this model, however, modelling arrivals to this level of detail is not essential and daily admissions would suffice. Pre-NHFD data was merged with NHFD data in order to collate as much information as possible, meaning that almost six years' worth of data was available. One outlying value of eight days was removed. The values used for modelling purposes are given in Table 6.3.5ii. The mean inter-arrival time was 0.727 days with a standard deviation of 0.834 days.
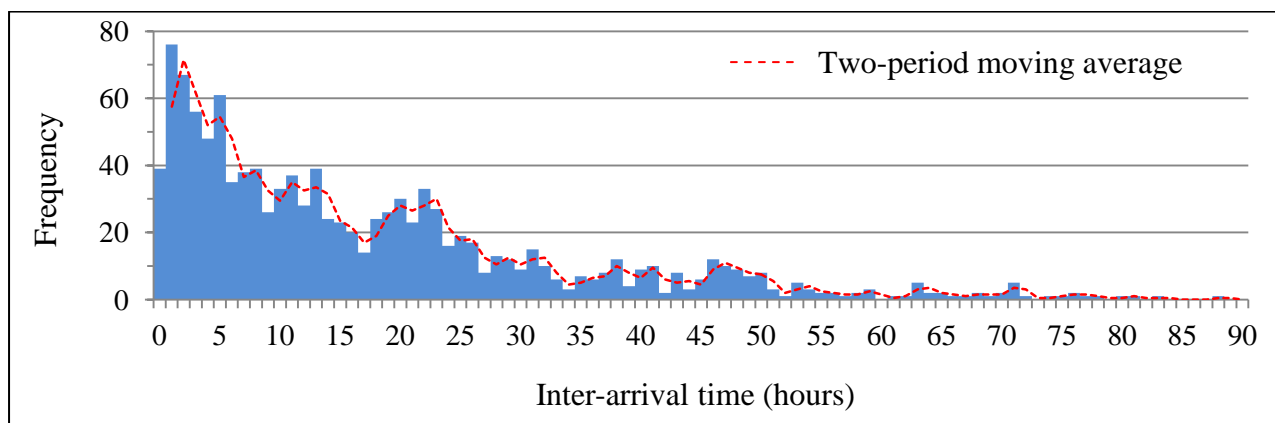
Figure 6.3.5i: Histogram of inter-arrival times in hours

Table 6.3.5ii: Distribution of inter-arrival times to the hip fracture ward

Inter-arrival time (days)	0	1	2	3	4	5	6
Percentage	46.2%	39.5%	10.9%	2.6%	0.6%	0.1%	<0.1%

6.3.6   Length of stay

Post-operative length of stay only is considered here, since length of stay pre-operation is already accounted for. This then also eliminates the operative delay itself when looking into whether or not delay affects length of stay. This analysis is done by ASA grade in order to gain more homogenous groups of patients and summary statistics are given in Table 6.3.6i.

Examination of these statistics would give rise to the possible conclusion that delay does matter. Statistically, however, there are no differences between delayed and non-delayed patients for ASA grades III and IV, while for grade I&II the difference is statistically significant at the 5% level.

Subsequent analysis showed no difference between grades III and IV, when looking at delayed and non-delayed patients in turn. This meant that these patients could be grouped and results in the following summary statistics. Additionally no differences were found between delayed and non-delayed patients, thus all ASA grade III and IV patients were grouped together, as per the last line in Table 6.3.6ii.

Table 6.3.6i: Summary statistics for post-operation length of stay (days), split by ASA grade and operative delay

ASA Grade	Delay	*n*	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurtosis
I&II	≤ 48h	204	18.78	19.29	1	117	2.71	8.93
	> 48h	145	23.14	19.90	0	115	2.11	5.67
III	≤ 48h	270	27.34	25.22	0	182	2.75	10.83
	> 48h	290	31.18	29.80	1	266	3.30	16.85
IV	≤ 48h	25	23.68	16.40	0	74	1.12	2.48
	> 48h	61	32.56	30.33	0	126	1.28	0.95

Table 6.3.6ii: Summary statistics for post-operation length of stay (days) for ASA grades III and IV, split by delay category

ASA Grade	Delay	*n*	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurtosis
III & IV	≤ 48h	295	27.03	24.59	0	182	2.76	11.21
	> 48h	351	31.42	29.85	0	266	2.93	13.86
III & IV	All	646	29.41	27.64	0	266	2.92	13.69

There are thus three groups of patients to consider with regard to post-operation length of stay. Each of these groups were found to follow a Lognormal distribution; the Lognormal distribution is commonly fitted to length of stay (Marazzi et al. 1998, McClean and Millard 1993), in part due to its long tails. Maximum likelihood estimates of the first two moments are given in Table 6.3.6iii, while the fits are displayed graphically in Figure 6.3.6iv. It is clear that reasonable fits were found in all cases, as was supported formally.

Table 6.3.6iii: Lognormal fits for post-operation length of stay (days)

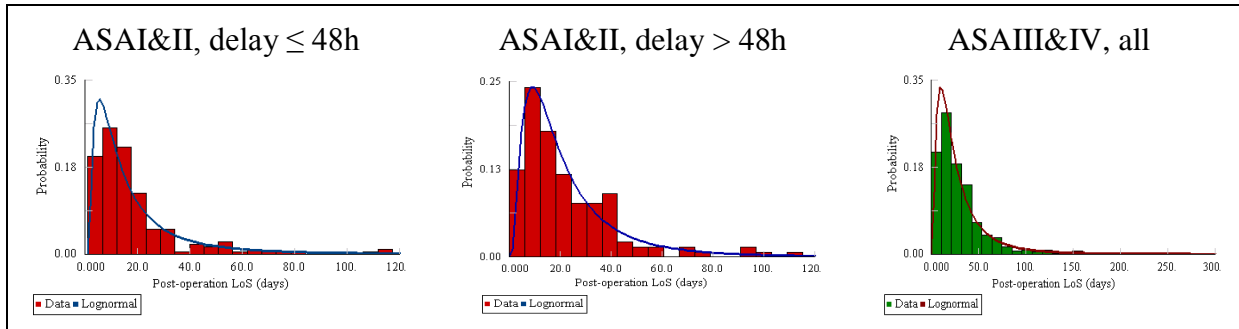| ASA Grade	Delay	μ	σ	Min	Theoretical		Empirical	
					Mean	S.D.	Mean	S.D.
I&II	≤ 48h	2.454	0.980	1	19.8	23.9	18.8	19.3
	> 48h	2.842	0.798	0	23.6	22.2	23.1	19.9
III & IV	All	3.052	0.857	0	30.5	31.8	29.4	27.6

Figure 6.3.6iv: Distribution of post-operation length of stay against the Lognormal distribution

6.3.7 Admission source and discharge destination

Pre-NHFD and NHFD data was again combined here in order to collate enough data for any meaningful analysis. This meant that operative delay had to be classified in days and not hours. While this loss of accuracy is unfortunate, it is also unavoidable. (So, as earlier in this thesis, a delayed patient is one who waits longer than two days.) There are eight separate sources from which a patient may arrive (see Section 2.2.2). Patients admitted either from home or from sheltered housing are grouped together into a new group called 'home', patients admitted from a residential or nursing home are grouped together into a new group called 'care home' and the remaining four groups are merged into a group called 'healthcare institution'. A significant relationship was found between admission source and ASA grade ($p < 0.0001$) and inspection of the data showed that fitter patients are more likely to be admitted from home.

Two discharge destinations are recorded for each patient: acute destination and final destination. Discharge destination is another area of interest and thus the effect of delay on this were investigated in order for it to be incorporated into the model. There are ten acute discharge destinations (see Section 2.2.13). These are grouped as per for admission source, with the addition of another group representing those who die in hospital. Final discharge destination is not considered further due to a lack of data. When considering discharge destination, admission source must also be considered for consistency.

The relationship between ASA grade, operative delay and acute discharge destination is shown for each admission source in Appendix D, Figures D6.3.7a-c. For admissions from

home, delay category was shown to be significantly associated with discharge destination for ASAI&II and ASAIII patients. No significant differences were found for ASAIV patients, or for any ASA category for care home and healthcare institution admissions. However, a problem was data constraints; after splitting by ASA grade, delay and discharge destination some groups were left fairly small and it would be inadvisable to make any resolute conclusions on this basis. The percentages are thus taken as indicated by the data and a value is sampled from the continuous Uniform distribution (range [0, 100]) and then compared with empirical values. However, an element of randomness is brought in with respect to mortality in order to emulate real-life unpredictability and a summary of mortality figures is given in Table 6.3.7i (≤ 2 days – operation within two days; > 2 days – operation after two days).

Table 6.3.7i: Percentage of patients who do not survive their stay in hospital, grouped with respect to admission source, ASA grade and delay

Admission source	ASA grade	Delay	Mortality percentage
Home	I&II	≤ 2 days	2.49%
		> 2 days	3.81%
	III	≤ 2 days	8.23%
		> 2 days	13.91%
	IV	≤ 2 days	37.50%
		> 2 days	25.56%
Care home	I&II	≤ 2 days	12.12%
		> 2 days	16.67%
	III	≤ 2 days	10.14%
		> 2 days	16.88%
	IV	≤ 2 days	30.00%
		> 2 days	35.71%
Healthcare institution	I&II	≤ 2 days	12.50%
		> 2 days	17.65%
	III	≤ 2 days	16.36%
		> 2 days	20.00%
	IV	≤ 2 days	27.27%
		> 2 days	40.00%

Introducing randomness here is achieved via use of the Poisson distribution. Consider, as an example, the sub-group of patients who are admitted from home with an ASA grade of I or II and who do not undergo surgery within two days. 3.81% of these patients do not survive their stay in hospital. The model uses the Poisson distribution about the value of 3.81 to create a new mortality percentage and the remainder of the discharge destination percentages are then adjusted proportionally.

The probability mass function (PMF) of the Poisson distribution with a mean value of 3.81 is shown in Figure 6.3.7ii. The probabilities at each value of x are calculated using $P(X=x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$, where λ represents the mean of the distribution. It can be seen that the most likely value to be sampled is 3, with the probability of sampling 4 just slightly smaller. While using this method results in extra running time for the model due to the additional calculations required, the advantage of bringing in the extra variability outweighs this issue.
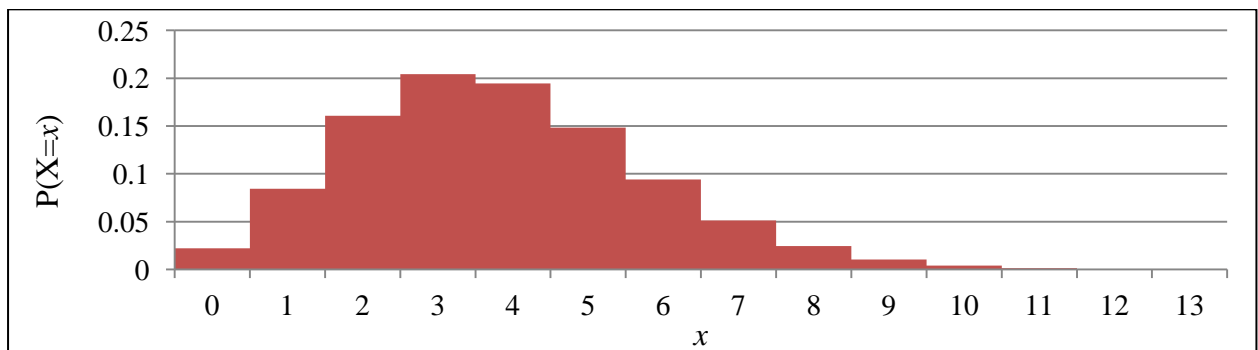
Figure 6.3.7ii: The PMF of the Poisson distribution about the mean value of $\lambda = 3.81$

6.3.8 Patients not undergoing surgery

Much of the model formulation so far has concentrated on patients undergoing surgery, but the small group of those treated conservatively or who die pre-operation must also be considered. 12.5%, 20.8% and 66.7% of patients were ASA grade I&II, III and IV respectively. Length of stay was found to fit the Lognormal distribution, see Table 6.3.8i.

52% of patients were admitted from home, 25% from a care home and 23% from a healthcare institution, but data restrictions meant that analysis of discharge destination by admission

source (/ASA grade) was not possible. On discharge, 67% of patients had died, 10% went home, 8% went to a care home and 15% went to another healthcare institution. The acute discharge destination was decided at random based on these discharge percentages, with the only exception being that a patient admitted from a care home could not be discharged home.

Table 6.3.8i: Lognormal fit for length of stay (days) for patients not undergoing surgery

μ	σ	Min	Theoretical		Empirical	
			Mean	S.D.	Mean	S.D.
2.189	1.247	0	19.4	37.5	18.7	30.4

6.3.9 Initialisation bias

Initialisation bias must be considered before results can be recorded. The model starts as an empty system, which is clearly unrealistic. For a non-terminating simulation such as this, the removal of this bias must be addressed to avoid misleading results.

There is an abundance of accessible research which details studies into the initial transient period, including methods on how to distinguish this from the steady-state period and also how to remove or deal with it. There are two general methods of dealing with this problem: the inclusion of a warm-up period or intelligent initialisation (Nelson 1992).

The first approach lets the model run for a specified warm-up period then, once it has reached steady-state conditions, data collection may begin. Alternatively, data is collected from the beginning of the simulation and then deleted from the results once the model has reached steady-state. It is vital to not underestimate the length of this transient period in order to avoid biased results. It is preferable to gain as accurate a figure as possible in order to also not overestimate the transient period, thus wasting useful results. For the second approach, the modeller must choose initial conditions which the simulation model begins from and data is collected from the commencement of the model. The main challenge here is deciding what these conditions should be.

The former approach was used for this study. There are over forty methods in the literature which deal with the issue of how long to set a warm-up period. Three of the most commonly

used methods are employed here and explained forthwith; time-series inspection, Welch's method and the MSER-5 method.

The interested reader may wish to refer to other statistical (Kelton and Law 1983, Robinson 2007, Yücesan 1993) or heuristic-based (Conway 1963, Fishman 1971, Gafarian et al. 1978, Pawlikowski 1990) methods available. There are also methods which do not determine the length of a warm-up period, but detect whether initialisation bias is present in a series of data (Goldsman et al. 1994, Schruben et al. 1983, Vassilacopoulos 1989). Note that this is just a small sample of the plethora of literature available on this topic.

(a) Time-series inspection method

Inspecting a time-series of the simulation output is the most straightforward method to identify how long a warm-up period should be. This time-series should display the key response of the simulation and in this case bed occupancy used. The simulation starts in an empty condition; that is, the bed occupancy is zero. The bed occupancy needs to reach steady-state before results are recorded. However, inspecting a time-series of a single run can be misleading or difficult to analyse as data can be very noisy; it is better to take several replications and then take the averages across the replications for each point on the time-series. The graph is then inspected to see where it becomes smooth.

The model was run 20 times and the bed occupancy at each day of each replication was recorded. The results of the mean bed occupancy can be seen in Figure 6.3.9i. The graph continued similarly past the 700 time periods shown on the graph but this is not included for presentation reasons. More replications were also tried but results were not found to differ significantly.

It appears that the graph is beginning to smooth at around 400 periods. There is still some variation beyond here but this is expected and does not lead to any cause for concern; a stochastic model will always result in fluctuations in outcome. This is clearly a subjective method and therefore other methods are also considered.
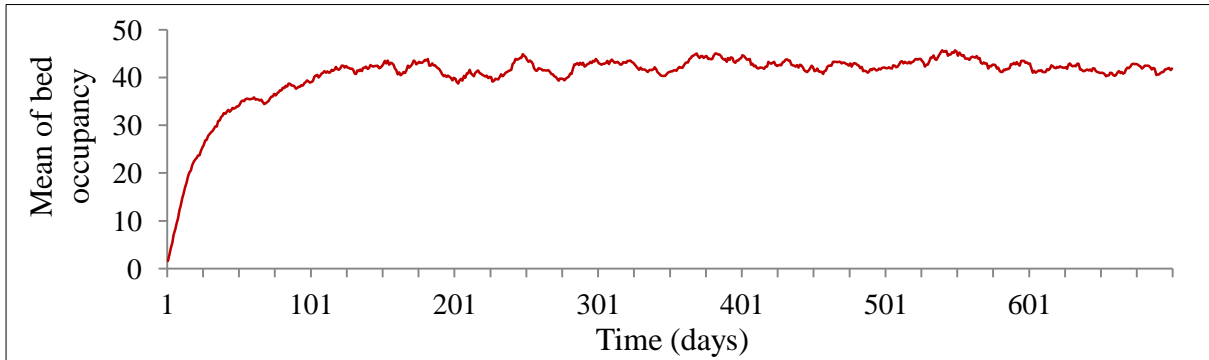
Figure 6.4.1i: Results of the time-series inspection method

(b) Welch's method

Welch suggests a method to determine the warm-up period of a simulation based upon moving averages (Welch 1983), a technique popularised in later years by Law and Kelton (Law and Kelton 2000). These moving averages, calculated using a window of size w for a maximum of m results in the series, are plotted on a time-series graph. If the data is smooth, then the result is acceptable; if not, then w is increased and the process is repeated. The warm-up period is then identified as the time that this time-series becomes flat. It is aimed to minimise w while still obtaining a reasonably smooth line; short-term fluctuations are thus removed but the intervals are not so long that they may distort the long-term trend.

Let $\bar{X}_t(w)$ be the moving average of window size w. The moving averages are calculated as follows:

$$\bar{X}_i(w) = \begin{cases} \dfrac{\displaystyle\sum_{j=-(i-1)}^{i-1} \bar{X}_{i+j}}{2i-1} & \text{for } i = 1, \ldots, w \\[3em] \dfrac{\displaystyle\sum_{j=-w}^{w} \bar{X}_{i+j}}{2w+1} & \text{for } i = w+1, \ldots, m-w \end{cases}$$

The results from this process are seen in Figure 6.3.9ii, where $w = 5$. Other window sizes were tried and results were not found to be notably different. Again the decision on what to use as a warm-up period is subjective but the graph provides an aid to this decision. As seen with the time-series inspection method, the graph appears to smooth at around 400 days.
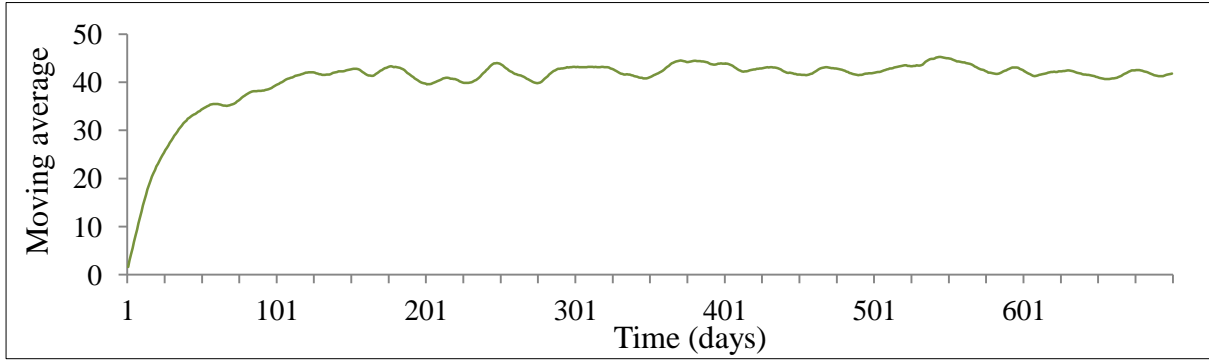
174

Figure 6.3.9ii: Results of Welch's method

(c) MSER-5 method

The Marginal Standard Error Rules (MSER) method of determining the length of a warm-up period was first introduced by White Jnr (White Jnr 1997) and later extended to the MSER-5 method (Spratt 1998). Full details of this method can be gained in the first instance from these sources but a brief overview is given here. These rules are based on heuristics and determine the truncation point as the value of $t*$ which returns the best trade-off between improved accuracy, measured in terms of bias elimination, and decreased precision, measured in terms of sample size reduction. Evidence of the efficiency and effectiveness of this method, in particular with regard to its superiority over other techniques, can also be found in the literature (Franklin and White Jnr 2008, Hoad et al. 2008b, White Jnr et al. 2000).

Consider the finite output series $\{X_i : i = 1, 2, ..., n\}$, representing the result of interest for the first n time periods of the simulation. For the MSER heuristic, the optimal truncation point for this sequence of results is given by

$$t* = \arg\min_{n \gg t \geq 0} \left[\frac{1}{(n-t)^2} \sum_{i=t+1}^{n} \left(X_i - \bar{X}_{n,t} \right)^2 \right]$$

where $\bar{X}_{n,t} = \frac{1}{n-t} \sum_{i=t+1}^{n} X_i$ is the truncated mean after the first t observations are removed from the series $\{X_i\}$.

The adaptation made by the MSER-m heuristic is that instead of using the raw output series $\{X_i\}$, a series of batched averages $\{B_j : j = 1, 2, ..., b\}$ is used instead. Here $b = \left\lfloor \frac{n}{m} \right\rfloor$,

175

where $\lfloor . \rfloor$ represents the standard floor function and the batched averages are calculated

using $B_j = \dfrac{1}{m} \sum\limits_{k=1}^{m} X_{m(j-1)+k}$. The graphical results of this method can be seen in Figure

6.3.9iii. The value of MSER-5 is minimised at period 24, with a value of 0.0039. Since the batch size is five, the warm-up period given by this method is 120 days.
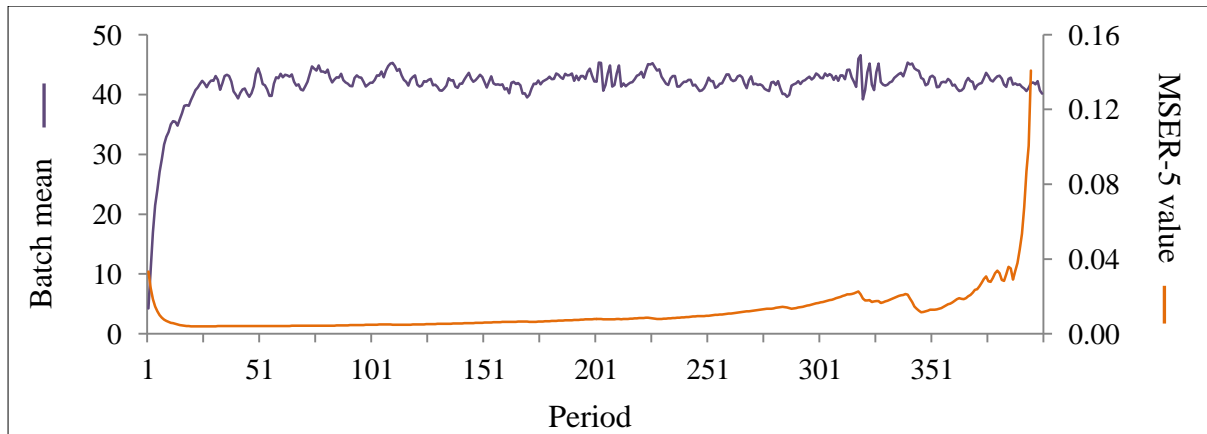
Figure 6.3.9iii: Results of the MSER-5 method

The first two methods considered appear to give similar answers for estimating the warm-up period and while they are subjective, unlike the more formal MSER-5 method, they should still be considered. MSER-5 suggests that 120 days would be adequate; however, on inspecting Figures 6.3.9i and 6.3.9ii it would appear that this would not be long enough. It is very important to avoid initialisation bias and therefore the higher estimate is taken. This figure was then inflated to 500 days, in order to be absolutely sure that ample time has been taken to initialise the system. This makes negligible difference to the run time of the model.

6.3.10 Run length

The model was initially run for an additional 2000 days, after the warm-up period had been completed. This was then increased up to a maximum of 20,000 days and various results were inspected. It was found that running the model for just 2000 days did not produce particularly stable results, despite the inclusion of a warm-up period, and the model needs longer to 'settle down'. It was then decided that 10,000 days would be a sufficient length of

time; the model has reached steady-state by this point, while not leading to any considerable extra running time. Despite the fact that the model is now representing a longer time period than the data on which it was based is of no real cause for concern; running a model for a longer time can be thought of in the crudest sense of simply several replications after each other – naturally, the decisions to be made concerning run length and the number of replications for a simulation model are inter-linked, so the number of replications is considered next. Additionally, mean figures will be used for later analysis, as well as yearly rates, and so any bias caused by run length is limited.

6.3.11 Number of replications

The approach used here to determine the number of replications required is based on confidence intervals. It is akin to the sequential procedure described by Chow and Robbins (Chow and Robbins 1965). A brief overview of this method is introduced here but a wealth of literature exists should the interested reader require a more detailed description (Hoad et al. 2008a, Law 2007, Robinson 2004).

For this method, the user chooses a pre-determined significance level and an output variable of interest; in this case the significance level is set at 95% and the output variable is again bed occupancy. Confidence intervals around the mean of this variable are then calculated at the significance level specified, using sequential cumulative means.

Results can be seen in Figure 6.3.11i, the blue line represents the cumulative mean while the red dashed lines represent the upper and lower limits of the 95% confidence interval. Steady-state values are achieved rather rapidly, which is no doubt in part due to the long run length. This graph was also constructed for up to 1000 replications but results are not included for display purposes; the lines continued in an almost perfect horizontal fashion.
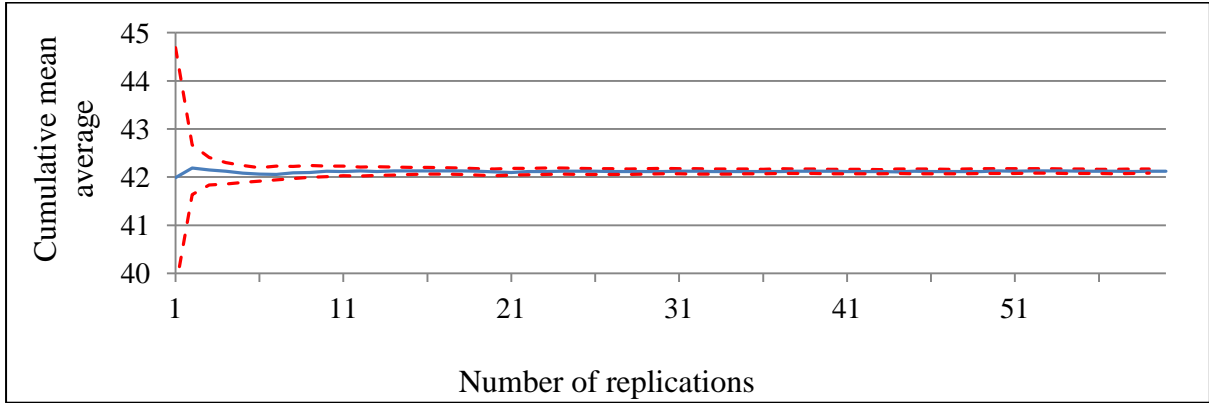
Figure 6.3.11i: Results of average bed occupancy with respect to the number of replications to perform

Precision criteria may also be set, or inspected later, by the user. This precision is defined as half of the width of the confidence interval, expressed as a percentage of the cumulative mean. Let r be the number of replications currently carried out while the finite output series $\{R_i : i = 1, 2, ..., r\}$ represents the results of the r replications for the output of interest. Also let $\bar{R}_r$ be the cumulative mean and s_r be the estimated standard deviation, where both of these values are computed using results $\{R_i : i = 1, 2, ..., r\}$. The precision at r replications, p_r, is consequently defined as

$$p_r = \frac{\left(s_r / \sqrt{r}\right) t_{r-1,\alpha/2}}{\bar{R}_r} \text{ x } 100\%,$$

where $t_{r-1,\alpha/2}$ is the value from the Student's t-distribution with $(r-1)$ degrees of freedom and a significance level of $(1-\alpha)$. Note it is not recommended to stop once a desired level of precision has been reached as this result may be reached by chance; premature convergence may occur. Inspection of plots and of the precision values obtained by the subsequent replications is therefore useful in this case.

The precision of mean bed occupancy was calculated to be 0.22% at 50 replications, while inspection of Figure 6.3.11i and the values following this both indicate that this was not an anomalous value. This was repeated for other measures of bed occupancy, namely the standard deviation, minimum and maximum, since variation and extremities in the system are

178

also important. In order to ensure anomalous values were not reached at these values, plots of p_r and the cumulative mean values were inspected in each case are presented in Appendix D. The cumulative mean plots for standard deviation, minimum and maximum of bed occupancy can be found in Figures D6.3.11a-c, while p_r is plotted for each measure in Figure D6.3.11d, beginning at $r = 3$ for display purposes.

Table 6.3.11ii: Precision values obtained for various bed occupancy measures at different values of r

Measure	Precision value, p_r (%)			
	$r = 50$	$r = 100$	$r = 250$	$r = 500$
Mean	0.22	0.08	0.07	0.07
Standard deviation	0.32	0.23	0.15	0.14
Minimum	1.69	1.27	1.00	0.61
Maximum	0.32	0.22	0.19	0.10

Relatively high precision is gained by using just 50 replications. However, if this is increased to 500 replications then all values are within a precision of 1%. For this reason, r is set to 500. Run time of the model is not of particular concern here and therefore shorter runtime is sacrificed for higher precision.

6.3.12 Model I summary

To summarise, patients arrive according to an inter-arrival distribution based on empirical data. Each patient is probabilistically assigned surgical or conservative treatment, an ASA grade, delay category and delay reason (if appropriate). Delay is modelled in hours according to a theoretical distribution based on ASA grade, delay category and delay reason. Post-operation length of stay (or total length of stay for patients treated non-surgically) is modelled in days, again according to theoretical distributions. Discharge destination is determined probabilistically using empirical data, with the Poisson distribution attached to mortality.

Key outputs recorded by the model include arrivals, daily bed occupancy and discharge destinations, split by relevant variables.

6.4   Validation and verification

Validation is the process of ensuring that the model is sufficiently accurate for the purpose at hand (Carson 1986) and is a binary decision; a model is either adequately accurate for its purpose or it is not, there is no grey area in-between. It is, however, not possible to prove that a model is valid – instead it is better to think in terms of the confidence that can be placed in the model (Robinson 2004). Of course it must be ensured that the conceptual model accurately represents the real world problem. Through thorough investigation of the variables and detailed discussion with the clinicians, it was decided that the most appropriate variables had been selected and that the conceptual model was fit for purpose. Data validation is a very important issue too; if the data on which the simulation was based is inaccurate then it is likely that the model becomes invalid. In this case, it is assured by the relevant staff that the data is recorded to as high a degree of accuracy as possible. The dataset was carefully inspected and any erroneous values were excluded before analysis.

Comparing the model to the real world system is a useful form of validation. If the inputs to the model are the same as the inputs to the real world system, then the outputs should be approximately equal for this type of validation, known as 'black-box validation' (Robinson 2004).

Table 6.4i gives some comparative bed occupancy outputs. The percentage of time that total bed occupancy exceeds 38 is of particular interest since currently hip fracture patients are not currently all on one ward, but spread over many different wards. The hip fracture team would like to centralise their patients and the possible wards available for this have 38 beds. Utilisation of this centralised ward is therefore vital. It can be seen that these results have a high level of accuracy.

Another measure of validation used relates to the acute discharge destination, see Table 6.4ii. As described in Section 6.3.7, discharge destination was split by admission source and delay. It was also split by whether the patient had an operation or not, which in turn influenced admission source. A check is therefore made with the purpose of ensuring that making these splits has not distorted the overall numbers discharged to each destination, which may have been a possibility due to the groupings made and the stochasticity built into the model. It is concluded that the model accurately represents acute discharge destination routes.

The slight discrepancies in both sets of results do not result in any cause for concern. Indeed, no model is ever 100% accurate, but is a simplified means for understanding and exploring reality (Pidd 2003), according to this definition the model is unquestionably valid.

Table 6.4i: Comparison of output and real world system, bed occupancy

Measure	Model	Data
Mean	42.4	42.1
Standard deviation	7.1	8.1
Minimum	20.7	22
Maximum	69.8	71
Percentage of time > 38	70%	66%

Table 6.4ii: Comparison of output and real world system, acute discharge destination

Acute discharge destination	Percentage to each destination	
	Model	**Data**
Home	34.6	34.5
Care home	14.0	13.8
Healthcare institution	39.2	38.8
Died	12.2	12.9

The task of verification is rather narrower; it is the process of ensuring that the model design has been transformed into a computer model with sufficient accuracy (Davis 1992); that is, assessing whether or not the computer model accurately represents the conceptual model. Verification, if completed properly, therefore ensures that the computer programming and implementation of the conceptual model are correct (Sargent 2000). Static testing is one basic approach to test simulation software, where the computer program is scrutinised to assess its correctness (Fairley 1976). Techniques utilised here include a structured walk-through of the model; the model was continually checked during its development and any mistakes were rectified and the appropriate section of code was retested.

It can be concluded that the simulation model has been tested thoroughly and has been successfully validated and verified, and thus may be used to investigate a number of changes.

6.5 What-if A: Varying the percentage of delayed patients

The percentage of delayed patients used as an input for the model was decided by historical data and split by ASA grade. By changing this, the effect of any increase or decrease of the percentage of delayed patients can be seen and the scale of the improvement or the deterioration on the system can then be judged in relation to the adjustment of the number of delayed patients. Each category of ASA grade used here is considered in turn.

6.5.1 ASA grade I&II

This patient group is deemed to be the simplest to treat due to the lesser severity of their medical circumstance and so with better resources and/or better management of these patients, improved conditions may well be achievable. All patients with an ASA grade of I or II are assumed to be administratively delayed only, therefore with optimal management and infinite resources, no patient in this group would ever have to wait for an operation. The observed percentage of patients who were not delayed was 58.1%; this value is now varied at every integer between 0 (worst case scenario) and 100 (best case scenario) percent.

Since ASA grade III and IV patients are not considered here (they can still be administratively delayed), there will still be some inefficiency in the system. However, a large change in the number of days spent delayed can still be seen, just by improving the treatment of these patients over whom there is more control.

If all ASA grade I&II patients were to be delayed, results show that approximately 1350 days in total would be spent administratively delayed by all patients each year. However, if no ASA grade I&II patient is delayed, then this reduces to approximately 780 days per year, a saving of 570 days each year. Currently the total number of days spent administratively delayed stands at approximately 1000 days per year, so a saving of 220 bed days is gained annually by improving the treatment of around 70 patients over the same time period.

It was shown previously that delayed patients experience a longer post-operation length of stay, in comparison with those operated on within 48 hours of admission, so by altering the delayed/non-delayed ratio the effect on length of stay is realised.

182

Table 6.5.1i: Results on post-operation bed usage when the percentage of non-delayed patients is altered for ASA grades I and II

Percentage of patients not delayed	Extra bed days		Percentage change from current situation
	per year	per day	
0	373.1	1.02	2.81
10	341.0	0.93	2.57
20	276.2	0.76	2.08
30	203.6	0.56	1.53
40	154.8	0.42	1.17
50	71.0	0.19	0.53
60	-20.2	-0.06	-0.15
70	-47.0	-0.13	-0.35
80	-94.0	-0.26	-0.71
90	-170.1	-0.47	-1.28
100	-199.8	-0.55	-1.51

If all ASA grade I&II patients were delayed (0% not delayed), then on average the hip fracture ward would have just over one more bed occupied each day. However, by ensuring that no patient in this patient group has to wait more than 48 hours for an operation, around 200 post-operation bed days per year can be saved. Again this is a fairly substantial saving considering it would mean a change in treatment for just 70 patients. Once this is combined with the gains made pre-operation, around 440 bed days are saved annually.

Changing the percentage of delayed patients also impacts upon discharge destination are results are given in Figure D6.5.1a of Appendix D. By improving efficiency and delaying less ASA grade I&II patients, almost six percent more patients are found to return home after hip fracture, while the percentages of patients discharged to a care home, another healthcare institution or who die are all found to decrease. For example, if one hundred percent of all ASA grade I&II patients are operated on within 48 hours, results from the simulation show that the average number of yearly deaths will fall from 57.1 to 55.9, a decrease of 2.1%. This would mean that an improvement for approximately 70 patients will save, on average, around 1.2 lives per year.

6.5.2 ASA grades III and IV

The case of patients with an ASA grade of III or IV is not quite as simple; some of these patients are clinically delayed and thus even under perfect circumstances, not everybody could be operated on immediately. The percentage of patients who are delayed at 48 hours for medical reasons is therefore kept fixed as it is assumed that this cannot be altered. These values are 19.7% and 48.5% for ASA grades III and IV respectively. The remainder of patients are either administratively delayed or not delayed.

A total of 81 scenarios are investigated for ASA grade III patients, where the percentage of non-delayed patients is tested at every integer in the range [0%, 80%]. 52 scenarios are investigated for ASAIV patients, across the range [0%, 51%]. These scenarios were run separately and results are presented in Figure 6.5.2i. It is immediately obvious that a greater effect is seen when the numbers of ASAIII patients are varied, in comparison with ASAIV patients. However, large variation in bed occupancy is not seen and so in terms of this measure it could be concluded that there are no huge advantages to be gained from making the extra effort or changes required for these scenarios to become reality.
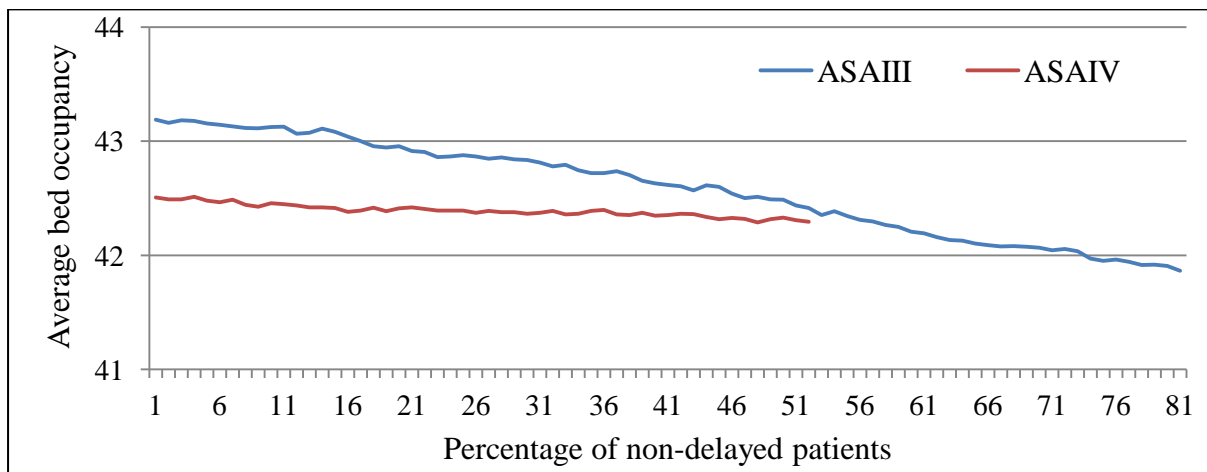

Figure 6.5.2i: The result on the average bed occupancy for different scenarios

6.6   What-if B: Altering the distribution of delayed hours

The number of hours patients spend delayed is currently modelled by statistical distributions based upon historical data. Consider now the effect of changing these values; the data has shown that length of stay and acute discharge destination may be affected by the time spent waiting for surgery, which will in turn result in subsidiary consequences for bed occupancy and patient outcome figures. These values are now changed in a variety of ways in order to quantify how altering the distribution of the number of delayed hours will affect the system.

6.6.1   Setting to a pre-determined value

Initially, the operative delay distributions are discarded and instead the number of delayed hours is set to a fixed value, so this aspect of the model is now deterministic. This is varied between 0 hours and 96 hours (four days) across all ASA grades.

Bed occupancy results are given in Table 6.6.1i. One of the main marker variables in this simulation model is whether or not a patient has their operation before or after 48 hours of arrival at the hospital and thus the biggest differences between the results are seen when the 48 hour mark is exceeded. Despite this, the bed occupancy figures still vary within the 0-48 hours and 49-96 hours intervals, due to the change in pre-operation length of stay. If all patients were operated on within the first hour of arrival, the mean bed occupancy would fall to 38 beds, which is the limit of the proposed dedicated hip fracture ward. Increasing this to 96 hours would result in an increase from this of around 8.1 beds on average. Inspection of the bed occupancy results shows that while a more volatile system is expected in terms of an increasing standard deviation, the coefficient of variation actually decreases as the set number of delayed hours increases, so the relative variation is lower.

Acute discharge destination is affected only by whether the patient undergoes surgery within 48 hours or not, so results at 0, 24 and 48 hours will be similar, but different to those at 72 and 96 hours. There are therefore just two groups to consider in terms of results here, see Figure 6.6.1ii. The percentage of patients discharged home decreases from approximately 32% to 25% if all patients are operated on after 48 hours instead of within 48 hours, while the percentage of deaths increases from 10% to 14%. The numbers discharged to a care home

remain relatively unchanged; while a relative percentage increase of around 4% more patients are discharged to another healthcare institution in the 'after 48 hours' group.

Table 6.6.1i: Results of setting the number of operative delay hours to a pre-determined value on bed occupancy

Fixed number of hours delayed	Bed occupancy				Percentage of time over 38 beds
	Mean	C.V.	Min	Max	
0	38.0	0.175	17.8	64.3	45.2%
24	39.3	0.172	18.8	66.2	53.1%
48	40.6	0.170	19.6	67.8	60.7%
72	44.7	0.163	22.7	73.2	80.2%
96	46.1	0.160	23.6	75.3	84.8%
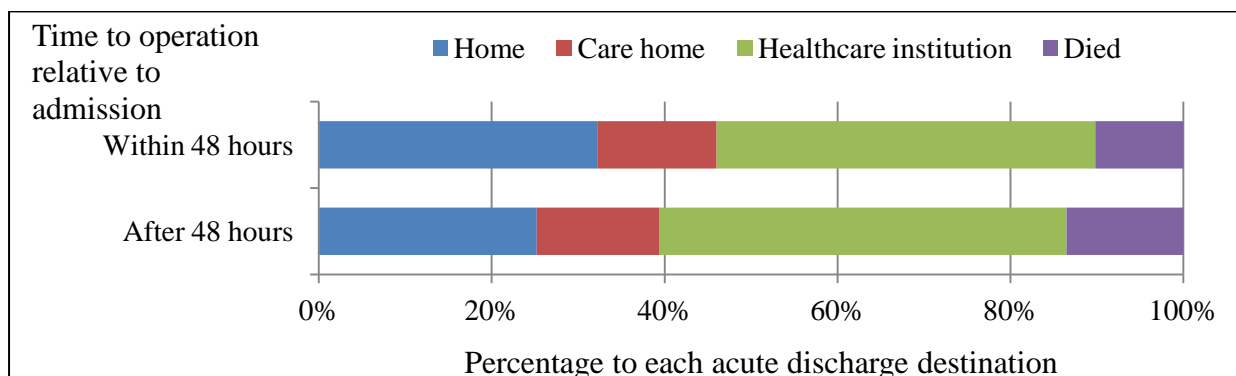
Figure 6.6.1ii: Results of setting the number of operative delay hours to a pre-determined value on acute discharge destination

6.6.2 Altering the distribution of delayed hours for ASA grades I&II

Operative delay for ASA grade I&II patients was modelled using a Lognormal distribution, which was also used to model the number of hours ASAIII and ASAIV patients spend administratively delayed. The Lognormal distribution has three parameters: the mean μ and standard deviation σ of the included Normal distribution, plus a minimum should one be required. The Lognormal distribution is a continuous distribution which is bounded on the lower side. It always returns a value of zero at minimum x and then rises to a peak that

depends on μ, σ and the minimum value. It will then decrease monotonically as x increases. Both the mean and standard deviation of the distribution also depend on both μ and σ, where the mean is also shifted by the minimum if one exists.

The fitted Lognormal distribution in this case had values of $\mu = 3.523$, $\sigma = 0.954$ and a minimum of 3, giving a mean of 56.4 hours and standard deviation of 65.1 hours. Altering the input parameters and rerunning the model will give an insight into the sensitivity of the simulated results on this distribution. The mean and standard deviation of the distribution are fixed in turn, while the other will vary; thus the results will reflect firstly a change in location and secondly a change in spread.

The value of σ is changed systematically across the interval [0.1, 2.6] and the value of μ then altered in order to keep the value of the standard deviation constant, but a change in the mean value will result from this. Note that it is not a requirement that $\mu \geq 0$. The resultant changes in the mean and μ are presented in Figure D6.6.2a.

Similarly, one can fix the mean but change the standard deviation of the distribution. The value of σ was again changed systematically but the value of μ then altered in order to keep the mean value constant, see Figure D6.6.2b. The average time spent delayed is thus not changed, but the effect of more or less variation in the system can be investigated.

These changes were made in turn as described and results are given in Figure 6.6.2i. On fixing the standard deviation but altering the mean, a rather modest effect on bed occupancy is observed, both in terms of the mean and the coefficient of variation, but much more unpredictable results are seen if the mean is fixed and the standard deviation is fluctuated.

The effect on discharge destination is also investigated. The percentage of patients discharged to each of the four acute discharge destinations is given in Figure 6.6.2ii. Similar but not identical results are seen. The lines appear to fluctuate less when the standard deviation is fixed, while fixing the mean results in starker changes across these values. Therefore despite keeping the overall average of the delay distribution the same, introducing more or less variation again shows a bigger change in results compared with varying the average of the same distribution.
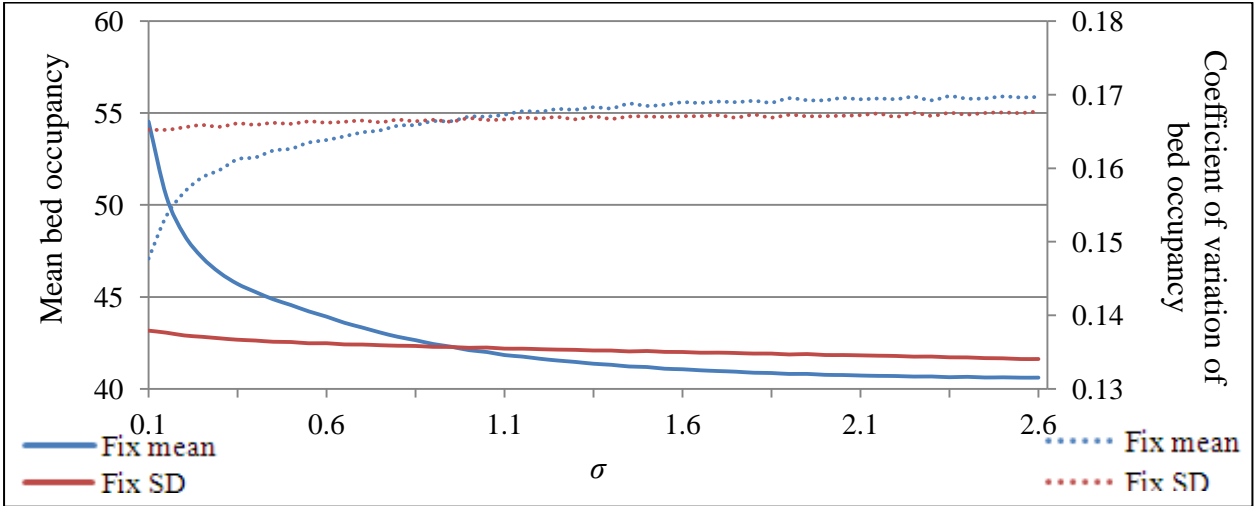
Figure 6.6.2i: Results of changing the distribution of delay of ASA grade I&II patients on bed occupancy measures
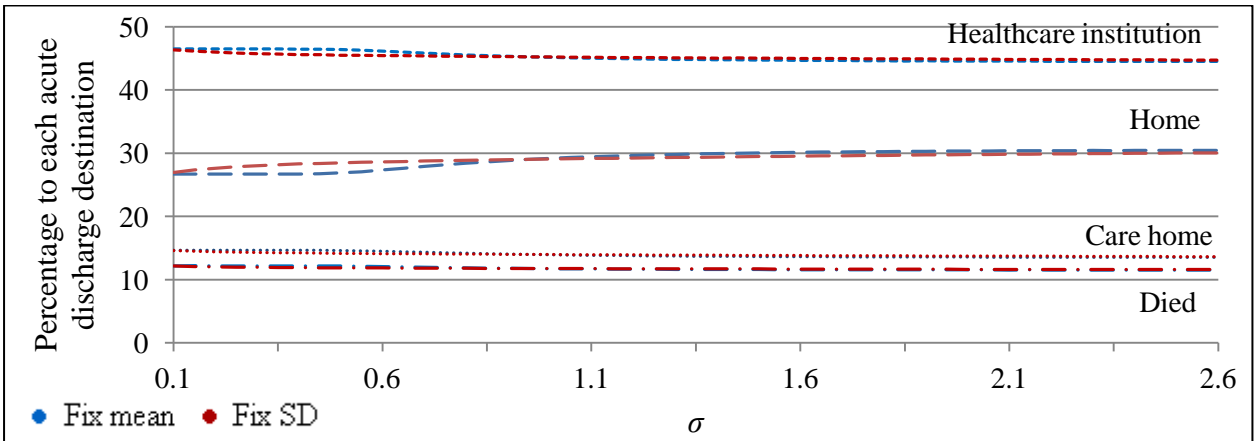
Figure 6.6.2ii: Results of changing the distribution of delay of ASA grade I&II patients on acute discharge destination

6.6.3 Altering the distribution of delayed hours for ASA grades III and IV

Operative delay for ASA grades III and IV was modelled using a grade-dependent Gamma distribution. The Gamma distribution is governed by three parameters: a shape parameter α, a scale parameter β and a minimum value if necessary. Note that if this distribution returned a value greater than 48, indicating that the patient is delayed, then their actual time spent delayed was decided by a different Negative Exponential distribution. The Gamma distribution is a continuous distribution bounded at the lower side with three distinct regions.

If $\alpha = 1$, then the Gamma distribution reduces to the Negative Exponential distribution. For $\alpha < 1$, the Gamma distribution tends to infinity at minimum x and decreases monotonically for increasing x. For $\alpha > 1$, the Gamma distribution returns a value of zero at minimum x, then rises to a peak which depends on both α and β, decreasing monotonically thereafter.

In this case, the value of α is varied and then the value of β is calculated based on α and either the resultant mean or standard deviation, whichever is required to be fixed. α is set to each 0.05 increment in the range [0.1, 4], producing 79 scenarios in each case in total.

The Gamma distribution used for the number of hours spent delayed for ASA grade III patients has a mean of 63.6 and a standard deviation of 44.6, with respective values for ASA grade IV of 102.7 and 76.6.

Firstly, values of α and β are changed in turn to keep the standard deviation fixed while increasing and decreasing the mean. The resulting change in the average (inclusive of the minimum value) in this case varies between 15.1 and 90.1 for ASA grade III and 29.2 and 158.2 for ASA grade IV. The relationships are displayed graphically in Figure D6.6.3a. It can be seen that as α is increased, β is forced to decrease in order to keep the standard deviation constant. These variations in α and β are shown to increase the mean of this distribution.

Firstly, the impact upon acute discharge destination is examined. The percentage change from the current situation is used as measure for each destination. Results are displayed in Figure 6.6.3i. The percentage of deaths is seen to vary rather dramatically when these distributions are amended; as expected, as mean delay increases, the number of deaths increases, while at the same time the percentage of patients discharged home falls.

The impact on various bed usage measures is also considered and results are seen in Figure 6.6.3ii, where the dashed lines indicate the current situation. By changing these distributions to the minimum mean levels as described, over 750 bed days can be saved each year. Additionally, the percentage of time the ward would operate at over 38 beds would reduce to less than 60%.
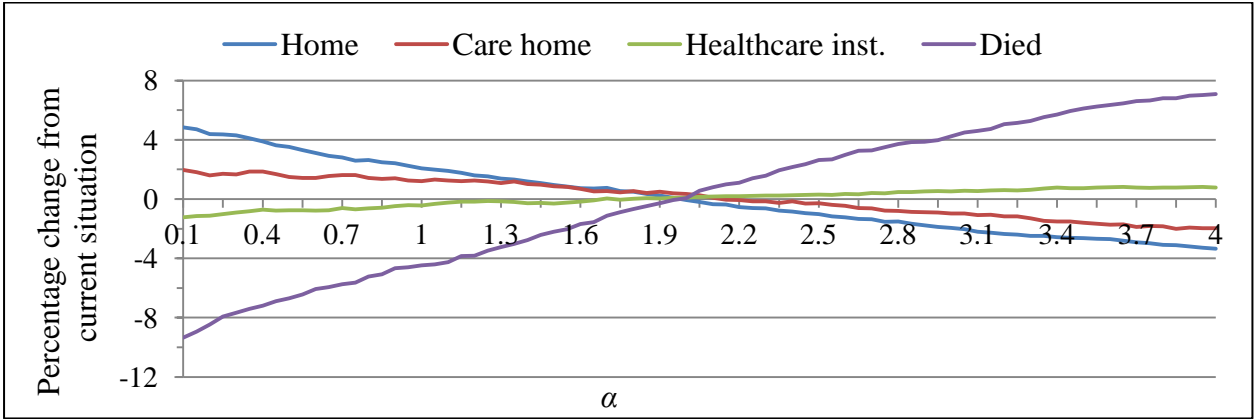
Figure 6.6.3i: Results of changing the mean of the distribution of delay of ASA grade III and IV patients on acute discharge destination
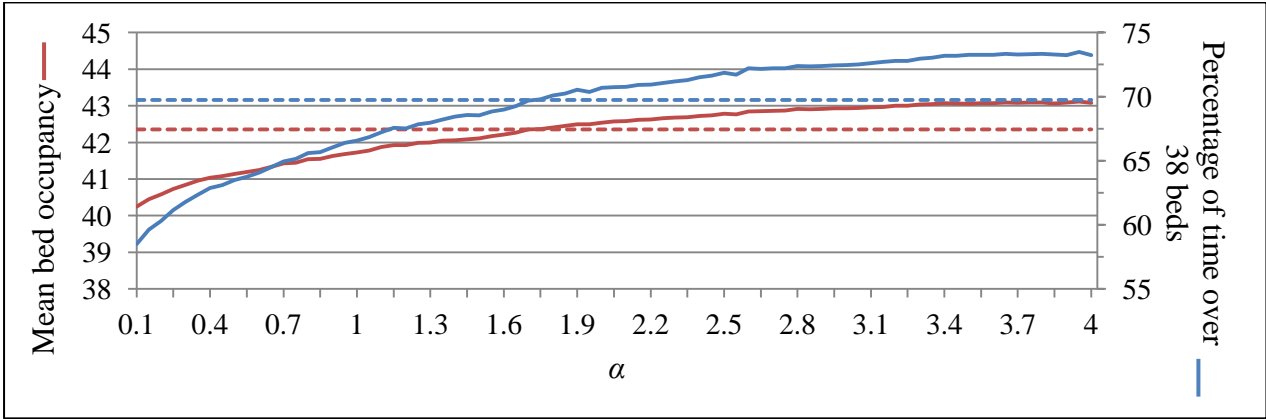
Figure 6.6.3ii: Results of changing the mean of the distribution of delay of ASA grade III and IV patients on bed usage

The mean values are now fixed in turn, while the variation in the system is altered. This results in a minimum standard deviation of 31.3 and a maximum of 197.8 for ASA grade III, with the values for ASA grade IV being 48.9 and 309.0. The standard deviation therefore increases at most by more than a factor of twelve in each case, in comparison to the original value, resulting in a rather more unpredictable system. Again, β is forced to decrease, this time to keep the mean constant, while the standard deviation decreases also. See Figure D6.6.3b for a graphical display of this relationship. Here the number of hours spent delayed per year is considered. Since delay was split by administrative and clinical delay, the same is done here with particular regard to the more controllable administrative delay. Comparative figures with the current situation are also given, see Figure 6.6.3iii.

190
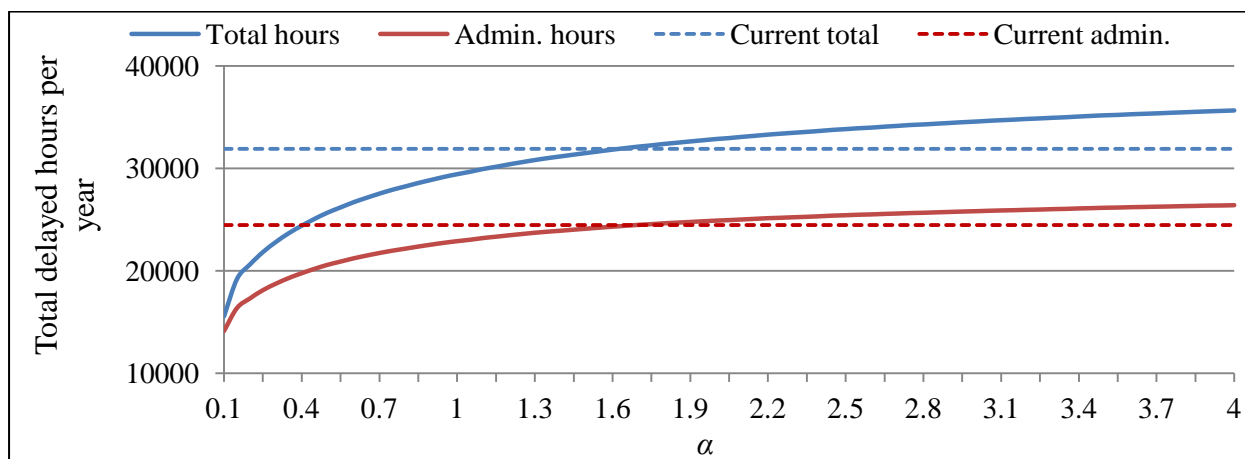
Figure 6.6.3iii: The effect of a change in variation of the distribution of delay of ASA grade III and IV patients on the total number of delayed hours per year

Finally, as a comparison between the consequences of alternately changing the mean or the standard deviation, the result on the total number of bed days per year is considered. This is calculated as the percentage change from the current situation and results are given in Table 6.6.3iv. A more extreme change is seen when the standard deviation is fixed and the mean is varied, rather than vice-versa. Thus not only does more stable system leads to more predictable results, as expected, better results may also be gained by controlling this variation.

Table 6.6.3iv: The effect of a change in mean or variation of the distribution of delay of ASA grade III and IV patients on the percentage of total bed days

α	**Fix mean**	**Fix standard deviation**
0.5	-1.63	-2.96
1	-0.79	-1.55
1.5	-0.07	-0.59
2	0.40	0.43
2.5	0.67	1.02
3	0.82	1.39
3.5	0.96	1.70
4	1.07	1.79

6.7 What-if C: Altering the number of arrivals

6.7.1 An ageing population

The issue of an ageing population in the United Kingdom is something that has been highly publicised over recent years. Population projections for those aged 50 and over have been computed by the Welsh Assembly Government for each local authority in Wales up until the year 2031 (WAG 2009a). These calculations are based on assumptions about births, deaths and migration. Changes in lifestyle, living conditions and health and social care provision are all reported to have led to an improvement in life expectancy (Ezzati et al. 2003, Khaw et al. 2008, WHO 2002).

The recorded numbers in this age group are reported to have increased by 15.6% in the years 1991-2007, standing at approximately 1.11 million in 2007. This is projected to increase to 1.42 million by 2031, a further 20.8%. This equates to 37% of the Welsh population being aged 50 and over in 2007, rising to an estimated 43% in 2031. This rise is attributed mainly due to an increase in life expectancy; over the thirty years from 1976 to 2006, life expectancy increased by 8.4 years and 6.3 years for males and females respectively. Indeed, the population aged 85 and over in Wales is projected to more than double in size between 2007 and 2031 (WAG 2009a).

Since the majority of hip fracture patients are elderly, it is important to plan for future provision requirements on the basis of an ageing population. Population pyramids are a useful way to view the distribution of ages within a population by gender; actual results for 2006 and projected results for 2031 are presented in Figure 6.7.1i for the local authorities of Cardiff and The Vale of Glamorgan (VoG). Note that the 2006 are overlaid onto the 2031 figures (the 2031 values were larger in every case).

The shape of the pyramid appears to be more symmetrical for 2031 compared with 2006, suggesting that the trend of a higher increase in life expectancy for males could continue. A striking increase can be seen in the ninety and over age group, which is estimated to show a huge increase in numbers over the coming years, more than quadrupling for males and more than doubling for females.

The population projections provided were used in combination with admission data to gain an expected arrival rate for the year 2031. Each age was weighted according to the proportion of admissions within that group in order to gain a more accurate estimation of a projected increase in arrivals; these were taken singularly between 60 and 89, with two other groups which captured patients aged under 60 and patients aged 90 and over. The rate of arrivals for under 60s was kept constant. See Figure 6.7.1ii for the results of this exercise.

In total, the average number of admissions per year is expected to be in the region of 860 patients by 2031, translating to an increase of 72.4% from the current observed numbers. At the same rate of increase per group, the number of annual admissions is estimated to double by the year 2041.
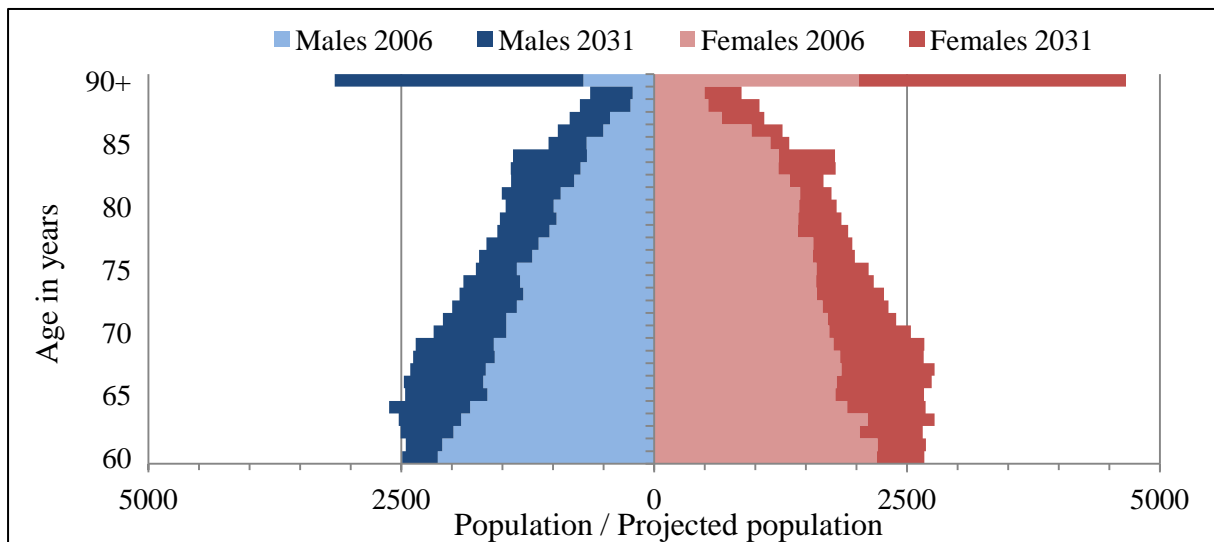

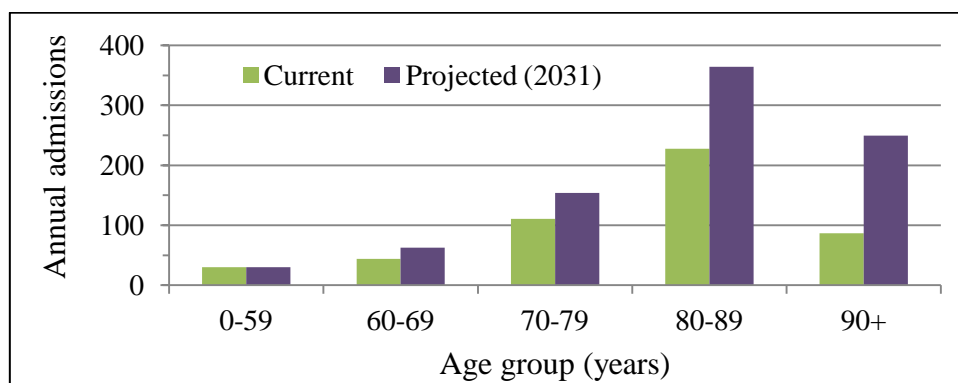**Figure 6.7.1i:** Population / Projected population pyramid for Cardiff and The VoG

Figure 6.7.1ii: Current and projected annual admissions of trauma hip fracture patients

6.7.2   Issues to consider

This issue of an increase in the ageing population is raised here as it will undoubtedly have an impact on the demand for the provisions provided by the trauma hip fracture service. However, there are various issues and assumptions to consider when interpreting these results, with regard to an increase in admissions. The estimated increase in demand is valid on the assumption that everything else influencing patients and patient care remains the same and relate solely to an increase in the number of patients being admitted under the care of the University Hospital of Wales with a fractured hip.

There is no way of knowing any organisational changes which may be introduced in the meantime; such as deferring some patients elsewhere or extra resources becoming available. These may include, but are not limited to; staff, theatre space and theatre time. An increase in any of these could mean that patients are treated more promptly.

There could also be clinical advances in this field which may influence the treatment of this patient cohort, or an increased awareness of this injury and its causes may also lead to fewer incidences. Alternatively an increase in the use of hip protectors may be seen, which may reduce the proportion of people suffering this injury, although the efficacy of these protectors is still under debate (Birks et al. 2004, Gillespie et al. 2010, Kiel et al. 2007) and it is reported that patient compliance also remains an issue (Sawka et al. 2005).

This is by no means an exhaustive list of issues which may have an effect on not only the number of future admissions, but also the way in which these patients are cared for and managed.

6.7.3   Changing the inter-arrival pattern

The mechanism used in the simulation model to generate arrivals is based on the inter-arrival times (in days) from historical data. By altering the proportion of arrivals which occur of the same day (i.e. an inter-arrival time of zero days), the overall number of entries to the system can be altered. The proportion of inter-arrivals which then fall at one, two, three, four, five or six days is then calculated based on the original proportions.

The data showed that the inter-arrival time between patients was zero days in 46.2% of cases. A decrease in this value is also considered for academic purposes. While it is unlikely that this will happen, due to the reasons explained and evidence presented previously, it is still interesting to see the effect on the system.

Firstly consider how the mean number of arrivals will vary by changing the inter-arrival pattern. Each integer value between 35% and 80% is considered for the percentage of cases that the inter-arrival time between patients is zero days, with the remaining values then adjusted accordingly. The effect on the mean number of arrivals per year can be seen in Figure 6.7.3i. The mean number of patients admitted is seen to approximately double when the inter-arrival time is zero days is in the region of 73-74%.
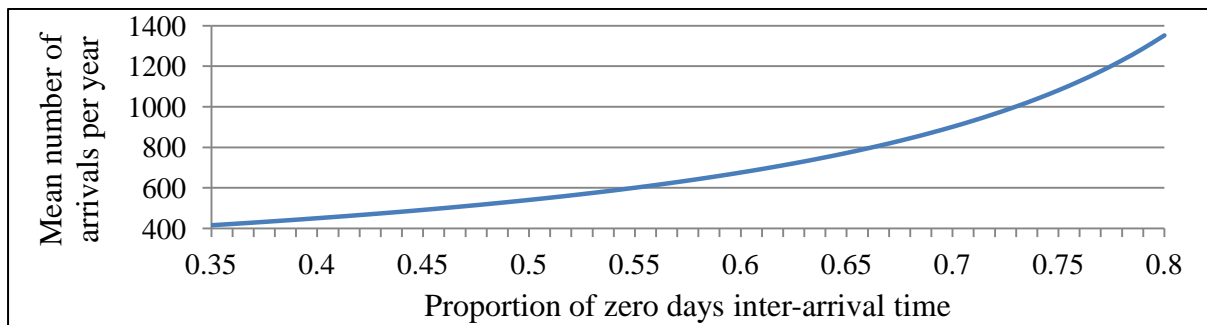
Figure 6.7.3i: Impact of a change to the inter-arrival pattern on annual admissions

Now consider how this increase in the number of arrivals will affect bed occupancy, see Figure 6.7.3ii. Intuitively an increase in the mean number of beds occupied is expected but quantifying this is worthwhile. Results may help healthcare workers with planning for the future in terms of resources and manpower. Note that the horizontal axis does not increase proportionally but relates to the increase in the number of arrivals seen in Figure 6.7.3i.

If the projected situation of around 860 admissions per year by 2031 is realised, then the hip fracture team can expect an average of approximately 72 beds occupied at any one time, with approximate minimum and maximum values of 36 and 122 respectively. Under these circumstances, the percentage of time that hip fracture patients would exceed the proposed dedicated ward capacity of 38 beds would stand at 99.92%.

The result seen for the current situation of average occupancy now becomes the roughly the best case scenario of the minimum number of beds occupied (42 beds occupied) if admission rates are doubled. The mean and maximum values are now 85 and 142 respectively, with the threshold of 38 beds exceeded 99.99% of the time.
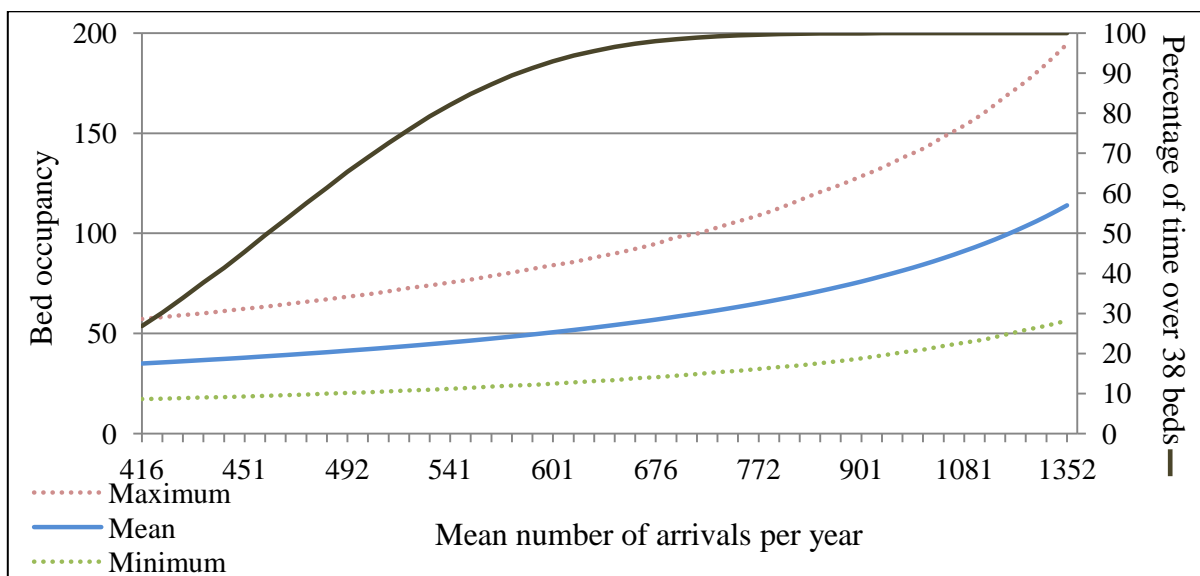
Figure 6.7.3ii: Bed occupancy results for a varying number of arrivals

It is also interesting to look at the variation in bed occupancy. As seen in Figure 6.7.3iii, more arrivals bring about more variation in the system. Of course, as bed occupancy increases, the standard deviation is expected to increase, since larger numbers will return a larger standard deviation, even if the relative changes are similar. For this reason, the coefficient of variation is also included. The relative variation is therefore also seen to increase, thus the hip fracture team will not only have to deal with more patients at any given time, but also with a more volatile system. This result may seem unintuitive and it is conjectured to be due to the process of changing the arrival pattern, as previously explained.

While these results may be expected by extrapolating the results seen for the current situation, any scale or precision of the change is unknown (or can be no more than supposed). By running these scenarios the precise figures can be gained, thus enabling more accurate strategic planning for the hip fracture ward. Indeed, since an accurate forecast of the future

situation has been garnered, then results and conclusions from some of the other what-if scenarios may prove useful in planning for the inevitable increase in admissions.
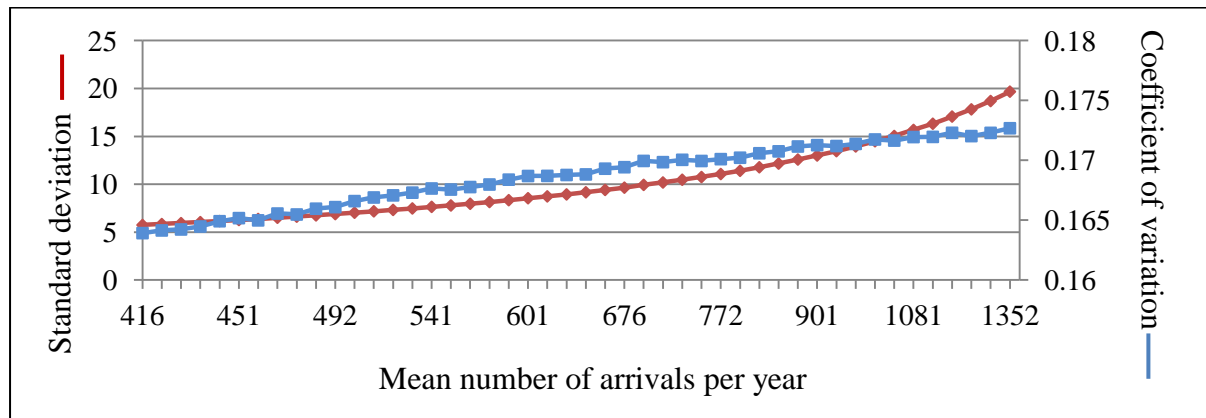
Figure 6.7.3iii: Results of the change in the standard deviation and coefficient of variation of the number of beds occupied as the mean number of arrivals per year is varied

6.7.4   Preparing for an increase in arrivals

The effect that an increase in the number of arrivals is likely to have on this system has been shown and, according to the evidence, it seems that this eventuality is inevitable. While it may be useful to be aware of this for planning reasons, the model can also be used to assess the impact that any changes which could be made may have on the system. Various strategies have been proposed and investigated thus far, and amending the percentage of delayed patients with an ASA grade of I or II is considered in further detail here. The reason for this is that, as explained in detail previously, this patient group is one which is deemed to be the easiest to impose changes upon. The percentage of patients delayed is varied at intervals of 25 between 0 and 100 percent inclusive.

Bed occupancy figures at each of these levels are displayed graphically in Figure 6.7.4i. Note that the sets of values used to create this graph are statistically different, both for overall group and individual pairwise comparisons, $p < 0.0001$. This is true for each of the three summary measures plotted. As an example, the biggest difference in mean bed occupancy is almost ten days; this is a rather remarkable figure when the overall number of bed days saved per year is considered.

Similarly, the effect on the yearly total number of bed days is investigated. The percentage changes from the current values are given in Table 6.7.4ii. Note that the current average number of arrivals per year is approximately 500. Results were interpolated so that these results could be quoted at tidy numbers of arrivals per year.
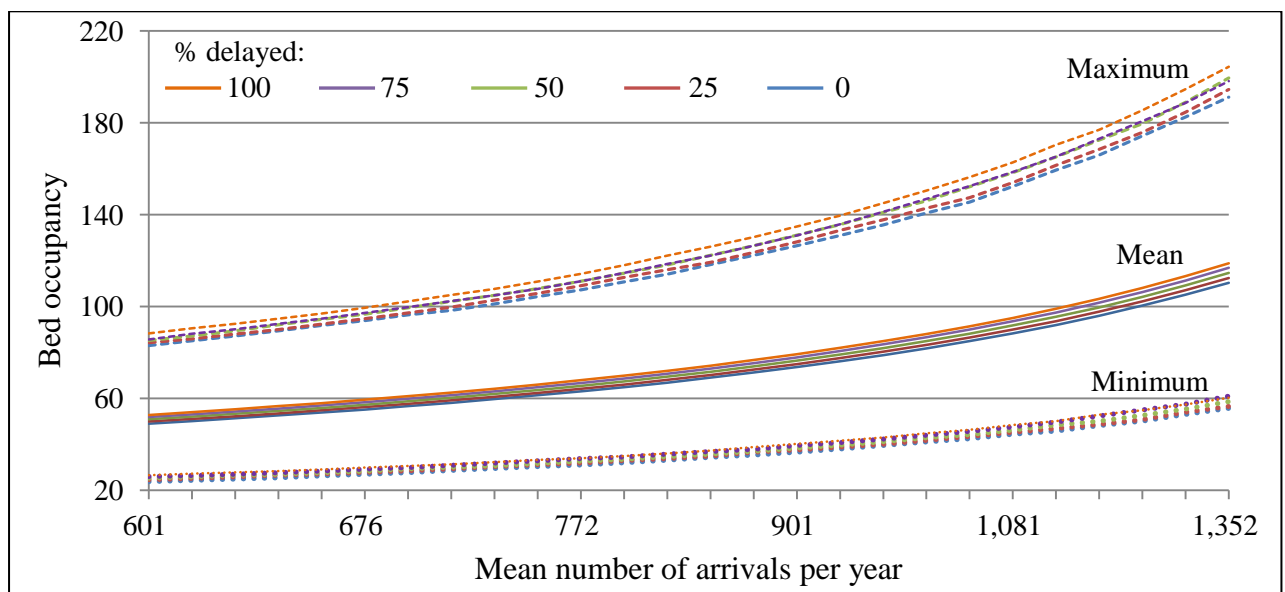
Figure 6.7.4i: The effect on bed occupancy due to changes in the percentage of delayed patients (ASA grade I&II) and an increase in arrivals

Table 6.7.4ii: Percentage increases in total number of bed days by altering the number of arrivals and the percentage of delayed patients, ASA grade I&II

Average number of annual arrivals	Percentage of delayed patients (ASA grade I and II)				
	0	25	50	75	100
600	15.7	17.9	20.4	22.7	25.0
800	54.1	56.9	60.2	63.4	66.3
1000	92.9	96.6	100.7	104.7	108.1
1200	131.2	135.5	140.3	145.2	149.5

Finally, the effect on mortality of this change in the percentage of delayed ASA grade I&II patients is investigated; results are given in Figure 6.7.4iii. On inspection of the left-hand *y*-axis, results may not initially appear relatively astonishing, but consider the number of

patients that these values are referring to. With an approximate average of 500 arrivals per year, a change affecting 0.2% of the patient group translates as one person.

Consider now the anticipated increase in the long-term future to 1000 admissions annually, the same percentage change now affects one more additional person. The numbers of average deaths for this admission level are therefore also given. It is also important to bear in mind that just the patient group of ASA grade I and II patients are under consideration here.
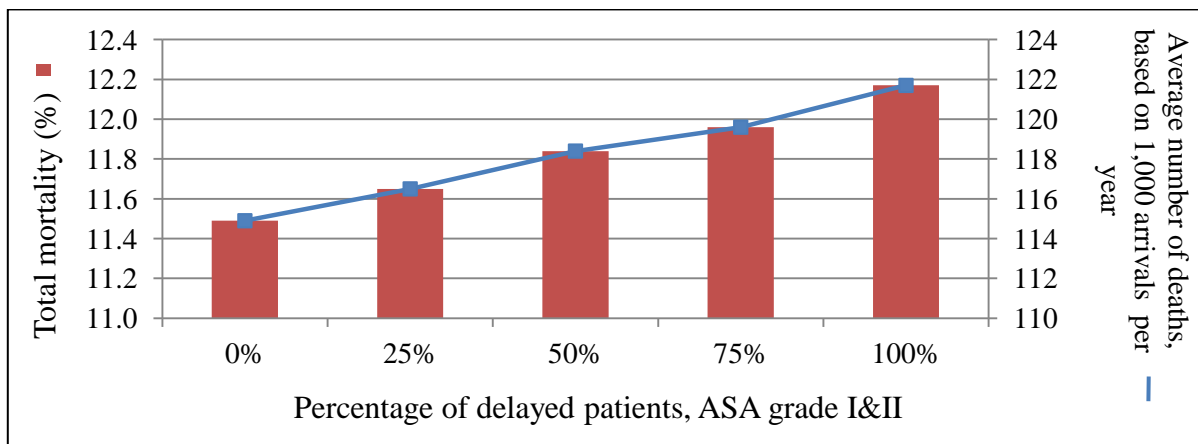
Figure 6.7.4iii: Percentage changes in mortality by altering the percentage of delayed patients, ASA grade I&II

6.8   Model II

Model II was actually formulated prior to Model I but was use of it was discontinued once NHFD data became available. While it is the same patient group that is modelled both times (but not entirely the same data used for the model formulation), they focus on different perspectives, with the main difference between the models being that operation type was the main focus of Model II instead of ASA grade. The operation type parameter was dropped in favour of concentrating on operative delay and ASA grade for Model I. This decision was taken on the advice and support given by a senior clinician involved in this work.

Rigorous set-up, validation and verification procedures were undertaken as previously described before the model was concluded to be fit for purpose and used for later analysis.

6.8.1   Model formulation

Recall that surgical procedure type is divided into seven categories; one for no operation (A, 3.4% of all patients) and six operation types (B-G). Type A patients were modelled as per Model I.

Significant differences in length of stay between types B-G were suggested by both a Kruskal-Wallis test and survival analysis, which indicated significant differences between the strata (as explained in Section 6.3.2), $p < 0.0001$. Using operation type as an indicator of length of stay is additionally verified by results from Chapter 3. Post-hoc analyses were then undertaken to assess differences in length of stay between delayed and non-delayed patients for each operation type and these were incorporated into the model where appropriate, using a cut-off of two days. (Only information on the number of *days* between admission and operation was available at this time.) A split by ASA grade was also used where appropriate.

Length of stay for the majority of groups could be modelled using a Lognormal distribution, see Table 6.8.1i. Note that in some cases groups are combined; this is either because there was no statistical difference in length of stay (for example, operation type D with respect to delay), or because there was not enough data to perform a split (for example, operation type C, ASA grade IV, with respect to delay). For operation types F and G no statistical fits could be found and therefore empirical values were sampled from, see Table 6.8.1ii for a summary.

Table 6.8.1i: Lognormal fits for length of stay, by operation type, delay and ASA grade

Operation type | Delay | ASA grade	μ	σ	Min	Theoretical		Empirical	
				Mean	S.D.	Mean	S.D.
B | ≤ 2 days | I&II	2.624	0.895	3	23.6	22.8	22.9	18.6
B | ≤ 2 days | III	2.914	0.998	3	33.3	39.6	31.8	32.5
B | ≤ 2 days | IV	2.712	1.089	1	28.2	41.1	25.1	26.7
B | > 2 days | I&II	2.543	0.974	6	26.4	25.7	24.1	15.8
B | > 2 days | III	3.075	0.901	5	37.5	36.4	37.7	40.2
B | > 2 days | IV	3.261	1.020	7	50.9	59.4	43.9	32.1
C | ≤ 2 days | I&II	2.057	0.967	1	13.5	15.5	14.3	20.6
C | ≤ 2 days | III	2.241	1.001	4	19.5	20.4	20.2	27.1
C | > 2 days | I&II	2.611	0.808	5	23.9	18.1	22.7	19.0
C | > 2 days | III	2.579	1.119	7	31.7	39.0	31.8	33.8
C | All | IV	2.631	1.174	5	32.7	47.7	28.3	26.9
D | All | I&II	2.644	0.948	6	28.1	26.6	27.6	25.6
D | All | III&IV	3.122	0.917	3	37.6	39.7	35.9	30.5
E | ≤ 2 days | all	2.932	0.823	1	27.3	25.9	26.8	22.9
E | > 2 days | all	3.145	0.867	3	36.8	35.8	36.1	34.8

Table 6.8.1ii: Summary statistics for length of stay (days) for operation types F and G

Operation type	Mean	S.D.	Minimum	Maximum	Skewness	Kurtosis
F	17.8	15.2	2	102	3.1	12.8
G	27.8	30.7	3	119	2.3	4.7

The extra length of stay seen in the delayed groups will in part be attributable to the extra delay and not due to any knock-on effects that the delay has on hospital stay after the operation has taken place. Despite this, investigation here is still worthwhile; even if the longer length of stay is caused wholly by the extra delay, amending the proportion of delayed patients (for example) will still have an effect on the system.

Inter-arrival times were modelled by sampling from empirical values, as per Section 6.3.5, and patients assigned an operation type, operative delay grouping and ASA grade on arrival

based on historical data, as displayed in Figure 6.8.1iii. Variation is introduced by attaching a Poisson distribution to the delay percentages using previously explained methodology.
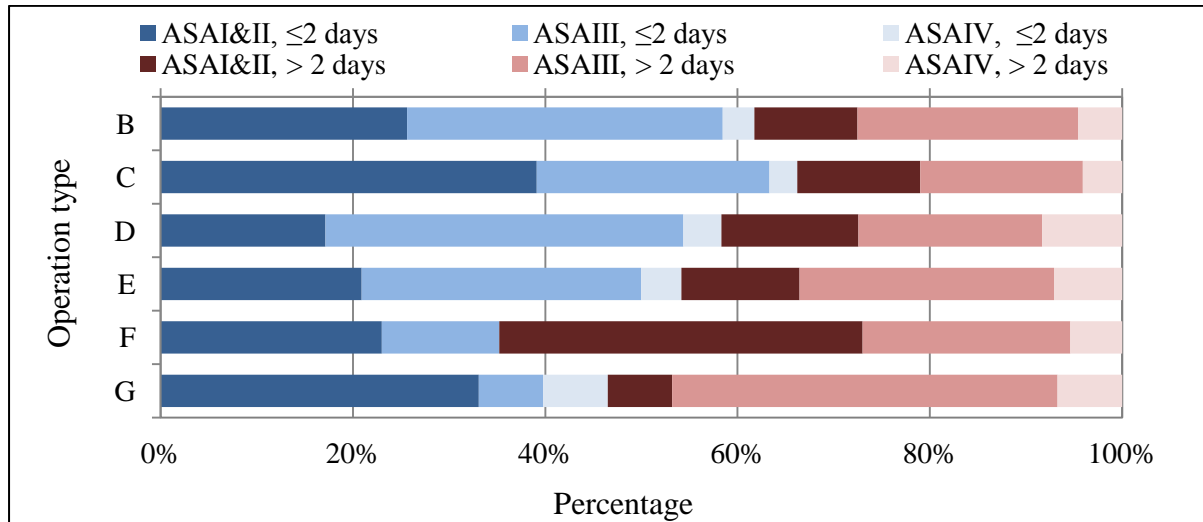
Figure 6.8.1iii: Patient distribution by ASA grade and delay category within operation type

Discharge destination was modelled similarly to Model I but could not be considered by operation type as well as admission source due to data limitations. However, the admission source was dependent upon operation type and thus indirectly influenced the discharge destination.

The same rigorous set-up exercises were undertaken for Model II as were performed for Model I with regard to the determination of the warm-up period, run length and number of replications. This was complemented by thorough validation and verification procedures.

6.8.2   What-if scenario: investigating the effect of delay

The main outcome of interest is to look at how varying the number of patients who are delayed impacts upon the system. The percentage of delayed patients within each operation group was varied while the remaining five groups were left unchanged. Finally, all groups were varied simultaneously. The amendments were made on the actual percentage value, this was then increased and decreased by each integer up to and including 35; this value was chosen as it ensured that no percentage ever exceeded 100 or was less than 0. The delay

percentage was then set to its new value, and the remaining percentage values were then adjusted proportionally. Clearly the percentage change for one group will therefore be affecting a different *number* of patients compared with the same percentage change in another group.

(a) Ward impact

Firstly, the effect that these changes have upon the ward is investigated, primarily with respect to mean bed occupancy. Results are displayed in Figure 6.8.2i.
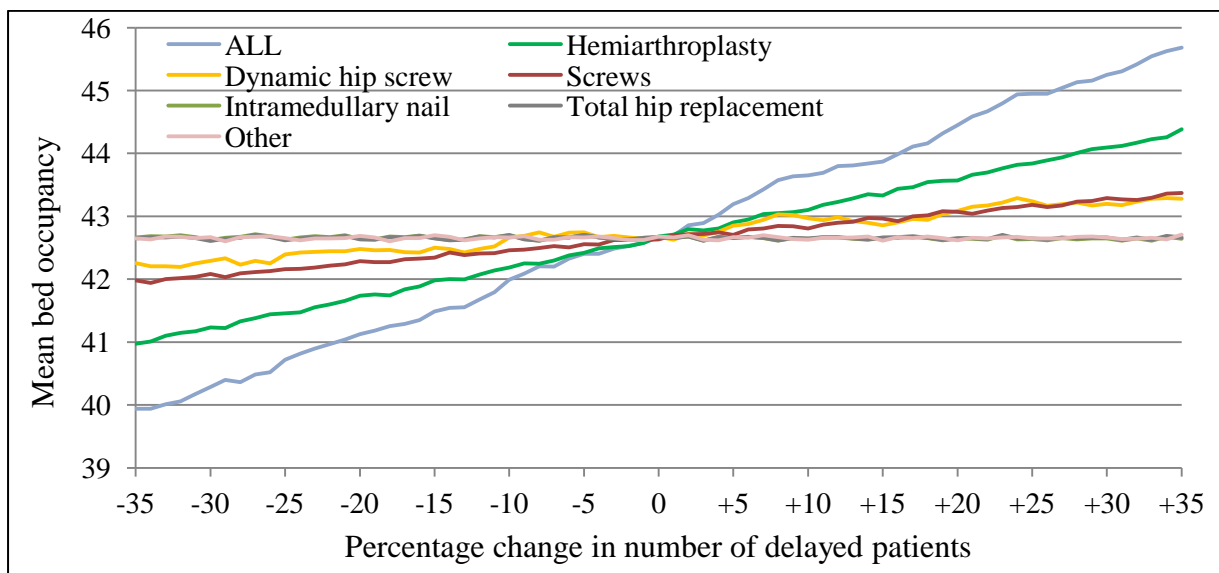
Figure 6.8.2i: The impact on bed occupancy by varying the percentage of delayed patients

The interest here is really in the particulars of this graph. Recall that patient numbers in the hemiarthroplasty and dynamic hip screw groups were very similar so the percentage changes here are affecting almost exactly the same number of people. With this in mind, the difference in results between these two groups is quite stark. Concentrating efforts on reducing the delay for hemiarthroplasty patients will have a considerably larger effect on the ward than for dynamic hip screw patients. Indeed, reducing the percentage of delayed patients in this group from the current value of 45.9% to 15.9% (a crude decrease of 35%, the largest decrease considered here), would mean that average bed occupancy will fall by over 1.5 beds, a saving of over 500 bed days per year. On average, to obtain this progression in

reality it would mean improving service for approximately 50 patients each year, where an improvement in service is defined as changing a delayed patient to a non-delayed patient.

Consider also the results seen for the operation type of screws. Patients in this group account for considerably fewer patients than in the dynamic hip screw group, but see better results. An improvement in the percentage delayed here is therefore referring to much fewer patients. The simulation model has therefore shown that by concentrating on an improvement in this patient group will lead to better results in terms of bed occupancy.

These results are obtained due to the differing length of stay profiles within each operation type group. Since ASA grade was taken into account, those groups with a higher proportion of sicker patients will in turn see this effect on length of stay and, ultimately, bed occupancy. A similar rationalisation can be applied to delay.

(b) Impact on discharge

The impact of delay on patient outcome, measured in terms of acute discharge destination, is also considered. A similar graph is now displayed, but with results relating to the average number of deaths. Unfortunately due to limited data, discharge destination could not be explicitly split by operation type as well as admission source and delay and therefore similar results will be gained from similar-sized groups. However, discharge was still influenced by operation type in terms of the discharge destination being affected by admission source and delay probabilities, which in turn were determined by operation type.

The greatest improvements are gained within the two largest groups, namely dynamic hip screw and hemiarthroplasty patients. If an improvement (in terms of fewer delayed patients) could be achieved across all operation types, the percentage of patients dying on the ward is shown to drop from around 12% to 10%, equating to approximately ten lives saved per year.

It is also important to consider the other acute discharge destinations as well as death. The percentage change from the number of patients currently discharged to each of the four destinations is given in Table 6.8.2iii. This is for changing the number of delayed patients across all operation types.

The results show that based on 500 admissions per year, if the crude percentage change in the number of delayed patients is decreased by 20 (for example), then 25 extra patients per year will be discharged home. A change in the opposite direction however, with a crude increase of 20 percent, will result in around five more patients per year being discharged to a care home, with same rise seen for healthcare institutions. This is not only less preferable to the patient, but provides more pressure on the National Health Service and Social Services.
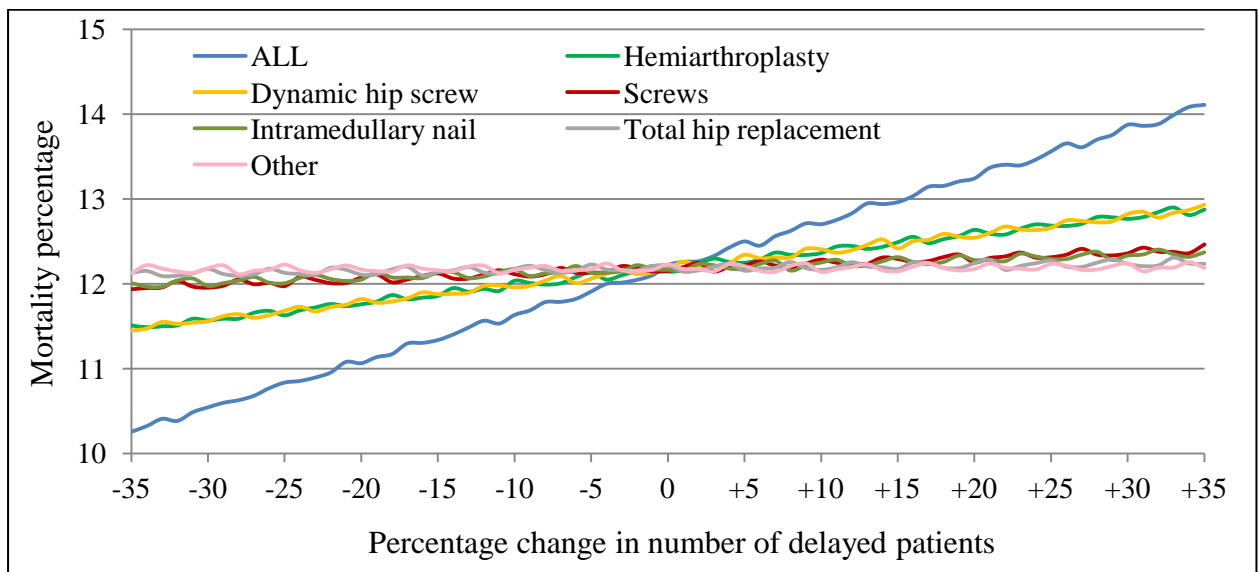
Figure 6.8.2ii: The impact on mortality by varying the percentage of delayed patients

Table 6.8.2iii: The impact on acute discharge destination by varying the percentage of delayed patients

Percentage change in the number of delayed patients	% change from current situation (acute discharge destination)			
	Home	**Care home**	**Healthcare institution**	**Died**
-30%	7.65	-1.80	-1.82	-13.72
-20%	5.04	-1.50	-0.96	-9.47
-10%	2.57	-0.69	-0.52	-4.81
+10%	-2.22	0.48	0.56	3.95
+20%	-4.58	1.04	1.06	8.37
+30%	-7.15	1.06	1.70	13.55

6.9   Recommendations

Using results obtained from running Model I and Model II of the hip fracture ward, a number of recommendations can be made into the most effective ways to improve throughput and occupancy levels.

Statistical investigation showed that for ASAI&II patients, post-operation length of stay is also reduced if surgical delay, defined as more than 48 hours between admission and operation, can be avoided. It is thus suggested that if any changes could be made to the system whereby any patients could undergo surgery more promptly, the greatest impact will be realised by focussing on ASAI&II patients. Results showed that by ensuring all ASAI&II patients receive surgery within 48 hours, 440 bed days and 1.2 lives could be saved per year. Other investigations showed concentrating efforts on reducing delay for ASAIII patients would give greater improvement than for ASAIV patients.

Of course, it is recommended that unnecessary delay is reduced as much as possible across all patients, regardless of their ASA grade. A decrease in pre-operation length of stay will therefore be directly reduced, while indirectly reducing length of stay post-operation. (Post-operation length of stay was shorter for non-delayed patients for each ASA grade, but did not reach statistical significance for ASAIII or ASAIV and thus was not incorporated into the model.) Results from the simulation show that completely reducing surgical delay would result in approximately four fewer beds occupied on average and a reduction in mortality rates from 12% to 10%.

It is additionally recommended that by targeting hemiarthroplasty patients (type E) for an improvement in treatment (reduction in delay), a better outcome will be realised than for other patients. 500 bed days would be saved each year if the 50 hemiarthroplasty patients who are currently delayed, instead reached surgery within two days. The smaller cohort who receive an operation using screws (type C) also yields good results.

Finally, it is advised that any variation in the system should be controlled as much as possible. Controlling for this variation can lead to improved outcomes, even if it superficially may appear than no improvements have been made to treatment of patients (no change in *mean* delay, for example).

6.10   Chapter summary

This chapter has focussed on building a simulation model of all trauma fractured neck of femur patients admitted under the care of Cardiff and Vale University Health Board. While comprehensive statistical analyses were completed prior to building the simulation model (Chapters 3 to 5), the task of formulating and building the model provided further useful insight into important factors relating to this patient cohort.

Due to the thorough investigations completed into important factors which affect the trauma hip fracture ward and its patients, there is confidence not only in the quality of this simulation model, but also that the correct classification variables were used. Rigorous analyses were performed in terms of the determination of the warm-up period, run length and the number of replications to run before any later investigations were undertaken. Additionally, both models were validated and verified and found to be fit for purpose.

A variety of what-if scenarios were performed once it was deemed to be fit the purpose for which it was designed. Recall that the first model has more flexibility in terms of the central variable of delay to operation, and as a consequence was considered in finer detail. The issue of an ageing population has also been discussed and the likely impact of this on the hip fracture ward has been documented. This idea was then amalgamated with one considered earlier, that of varying the number of patients who are delayed, in order to assess the impact on various performance indicators of an increase in admissions as well as better patient management. The second model provided additional useful information relating to operation type.

CHAPTER 7: THE TRAUMA THEATRE

7.1   Introduction

This chapter focuses primarily on the data available for operating theatres at the University Hospital of Wales (UHW) and, in particular, the emergency trauma theatre. The reason for this is that this is the theatre in which almost all trauma hip fracture operations take place. These surgeries are only performed in another theatre on very few occasions and thus the impact of this on the system is assumed to be negligible and is therefore ignored.

An understanding of the workings of the trauma theatre, and indeed the general operating theatre department at the UHW, was made possible via two main methods. Firstly, detailed discussion with staff members, primarily with the Theatre Manager (who is also a consultant anaesthetist) and two data managers, allowed for insight into how and why data is recorded. Secondly, a morning was spent shadowing the Theatre Manager in surgery, observing operations at all stages (pre-, peri- and post-) across a variety of different theatres. This allowed for further useful discussion with surgeons, operating department practitioners (ODPs) and other medical staff.

7.1.1   TheatreMan software

Theatre activity data is collected in the operating theatres at the UHW using the software package TheatreMan, which is "designed to report on all activity that is captured during the patient episode for clinical and management purposes and includes real-time patient data capture" (Trisoft 2009a). There are a range of modules available within this program, including those related to advanced scheduling, forecasting and list management and treatment, as well as various reporting and documentation options. Key benefits of implementing this software include the promotion of better patient care, increased ease of identifying bottlenecks and delays in the system and providing staff audit information (Trisoft 2009b). This software has been implemented across many NHS Trusts and accounts of successful functioning have been reported (Fairley 1991, Trisoft 2006).

The data fields available for the purpose of this study include patient information, operation information and detailed timings regarding the patient's journey from the ward, during the operation, and through to the recovery ward. Patient information includes age and hospital number, which is each patient's unique identifier.

Operation information includes OPCS-4 code, operation description, surgeon and location. OPCS-4 codes are a classification of interventions and procedures given by the Office of Population Censuses and Surveys. They consist of a letter followed by three numerical figures, which are separated by a period (.) between the second and third digits. The letters denote 24 chapters of classification and each chapter represents a different part or system of the body. The relevant chapter code here is *W*, which represents *Other Bones and Joints*. A localised code, in the form of a single letter, may also be appended if required (NHS 2005).

A screenshot of the treatment module of the TheatreMan software is given in Figure D7.1.1a of the Appendix. Patient-specific and surgeon information is censored for privacy reasons. Various pieces of information can be seen, including the operation performed (both the OPCS-4 code and a textual description), the theatre in which the operation took place and a number of timings which map the patient's journey through the theatre.

7.1.2 Data extraction

A selection of different databases was provided by the theatres data team at the UHW, relating to theatre pathway timings, cancellations and utilisation. Note that these databases did not necessarily cover the same time period and that also the times at which they became available was staggered throughout the course of this study.

In liaison with medical and administrative staff at the UHW, a complete list of fractured hip OPCS-4 codes was compiled, so that the relevant operations could be extracted from the dataset of all trauma theatre patients. Initially, theatre information relating to hip fractures only was available and thus the primary work undertaken, and hence the primary analyses described in this chapter, concentrate on this patient group. In total, 21 different types of hip fracture operation would be considered, see Appendix B, Table B7.1.2a, for a complete list of these operations. This original database included OPCS-4 codes, surgeon information and

pathway timings. The remaining databases became available later and are described as they are introduced throughout this chapter.

7.1.3 Data validation

Many of the fields seen in the TheatreMan screenshot are entered by medical staff in real time. High workload in a stressful environment such as an operating theatre may lead to difficulties in entering accurate and complete data and evidence of this was seen throughout the data validation process. One validation measure undertaken was checking whether the successive intervals were recorded in chronological order; this was a simple test to do but resolving errors was a little trickier as it was not necessarily obvious which of the times was recorded incorrectly. For simplicity, any times not recorded chronologically meant the removal of that time and the preceding time.

Other validity issues included duplicate entries; on several occasions patients were found to be entered twice into the database but with slightly conflicting times or with different operation codes. In the latter case, it was suggested that this may be due to the patient undergoing more than one type of operation, but in the majority of cases this could be regarded as a mistake. In the former case, one observation usually included many more entries than the other and was therefore kept.

Once all of the data was cleaned and erroneous values were removed, or corrected if possible, it was deemed that it was then fit for purpose and could be explored and investigated further. Because of the issues explained previously, the value of n will vary throughout the following report due to missing data and obvious data entry errors being excluded.

7.2   Theatre pathway

The time fields recorded for each patient, mapping their journey from ward to recovery, via theatre, are as follows:

Asked for:	Porter requested to fetch the patient from the orthopaedic ward.
Sent for:	Porter fetches the patient.
Arrived:	Patient arrives at the loading bay, part of the operating theatre suite.
Into anaesthetic room:	Patient enters the anaesthetic room.
Anaesthetic start:	Anaesthesia procedure is started.
[Into theatre:	Patient arrives enters the operating theatre.]
Operation start:	Surgical procedure is started ('knife to skin' time).
Operation finish:	Surgical procedure finishes.
Out of theatre:	Patient leaves the operating theatre.
Into recovery:	Patient arrives at the recovery ward.
Out of recovery:	Patient leaves the recovery ward and returns to an orthopaedic ward.

One time field is parenthesised since although it is available within TheatreMan, it was not recorded in each of the databases provided and is therefore not always used; leaving a total of either eleven or ten time recordings available. From these times, which are all recorded in the format *hh:mm* on a 24-hour clock, ten or nine time intervals can be created to represent the length of time, in minutes, that each process takes.

Some of these time intervals can be viewed as delays. As an example, the time between *asked for* and *sent for* is, in theory, an avoidable delay; if a porter was available to fetch the patient as soon as they were requested, then this would always return a time of zero for this time interval. However, discussion with staff involved in this process on a daily basis led to the advice that these delays are known by the surgeon; that is, the surgeon will ask for their patient with the knowledge that this will not happen immediately, so they will ask for the patient before they are actually ready for them knowing that there will be a certain length of time before the patient arrives at the theatre area. From this perspective, it is more desirable

that a patient waits in the loading bay rather than a surgeon waits in the operating theatre, wasting valuable theatre time and resources.

Initially, *into theatre* was not available, leaving nine time intervals available for further scrutiny. Of these, it was decided that the possibility of considering three as dependent upon type of operation should be contemplated. The remaining six could therefore be analysed using all hip fracture trauma data and did not need to be broken down further by operation type. The three intervals dependent upon type of operation are as follows: *anaesthetic start – operation start*, *operation start – operation finish*, *into recovery – out of recovery*; or, the time taken for the anaesthesia procedure, the time taken to perform the operation and the time the patient spends on the recovery ward. These are looked in Section 7.2.2.

7.2.1   Independent time intervals

A total of 1136 patients were extracted from the trauma theatre database for the original data analysis (hip operations only). First consider the time taken to complete the six intervals not dependent upon operation type or any other influencing factors.

It was recommended by the Theatre Manager at the UHW that times of zero are unrealistic for some intervals; it is simply physically impossible for some of the transfers from one state to another to be instantaneous. It was decided that two intervals would need to be amended, namely *sent for – arrived* and *out of theatre – into recovery*.

For the former, times less than or equal to five minutes were deemed infeasible. The wards used for hip fracture patients are at least three floors away and it would not be possible for a porter to collect the patient and then arrive at the loading bay in less than five minutes. Similarly, the recovery ward is along a corridor from the operating theatre and it was decided that this journey cannot be done in less than two minutes, particularly with a large trolley to be taken into consideration. Though fairly rudimentary, it was decided to simply ignore any times in these intervals which do not suit the criteria above, as they must be invalid entries. This reduces the number of entries available for analysis by around half for the *out of theatre – into recovery* category and although this is unfortunate, it is better than including known incorrect values in the analysis. Summary statistics for each of these six independent time

intervals, ignoring the invalid entries as described previously, are now given. (The digit *2* is appended to the two intervals which have had entries removed.)

Table 7.2.1i: Summary statistics for time intervals not dependent on operation type (minutes; anaes. – anaesthetic)

Time interval	N	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurt-osis
Asked for – sent for	1045	6.31	6.17	0	51	2.07	5.96
Sent for – arrived 2	944	16.17	8.43	6	153	7.27	99.57
Arrived – into anaes. room	1136	12.13	9.92	0	97	2.21	9.37
Into anaes. room – anaes. start	1135	3.92	4.92	0	65	3.71	28.53
Operation finish – out of theatre	1128	7.45	8.62	0	185	9.89	171.02
Out of theatre – into recovery 2	565	6.74	9.47	3	125	7.53	71.75

In order to model the trauma theatre, it may be required to find distributions for each of these intervals. Stat::Fit was used to attempt to find statistical distributions for each of these six time intervals but all null hypotheses were rejected. This was true for each of the Anderson-Darling, Kolmogorov-Smirnov and Chi-square goodness-of-fit tests in every case at the 5% significance level, for each continuous distribution which Stat::Fit checks against. For verification, this exercise was repeated in SAS and the same conclusions were reached.

Another approach considered here was the Hyperexponential distribution. This distribution is the sum of n non-identical Negative Exponentially distributed random variables, where $n \geq 2$, and is a special case of the more general phase-type distributions, which are well-suited to the field of healthcare (Fackrell 2009).

More specifically, the probability density function of the random variable X, as represented by n Negative Exponentially distributed random variables, is given by $f_X(x) = \sum_{i=1}^{n} p_i f_{Y_i}(y)$.

.

Here Y_i is random variable with a Negative Exponential distribution with rate parameter λ_i; that is $f_{Y_i}(y) = \lambda_i e^{-\lambda_i y}$ for $i = 1,...,n$. The probability that X will take the form of the Negative Exponential distribution given by Y_i is given by p_i, so that $\sum_{i=1}^{n} p_i = 1$.

A defining feature of this distribution is that its coefficient of variation is greater than one or, analogously, the standard deviation is greater than the mean. It is therefore immediately obvious that it will not be suitable for some of the time intervals presented so far. For the remainder, fits were attempted to be found for the Hyperexponential distribution with two, three, four and five phases. This was done using Solver, an add-in available for Microsoft Excel, where the parameters to estimate are the rates λ_i and the probabilities p_i; $i = 1, ..., n$. Solver is an optimisation software, where the aim is to find certain values of cells within a spreadsheet, in which the decision variables are located, that optimise a certain objective. Both maximisation and minimisation problems can be dealt with using this method. Hard constraints can also be specified if required. The method employed here involved minimising the sum of squares between the empirical and fitted probabilities by amending the aforementioned parameters. However, still no fits could be found.

On further inspection, various other issues with the data were found. Since the data is recorded by theatre staff in real time, optimal accuracy is difficult. As an example, the distribution of times recorded for hip fracture patients from the time they arrived at the operating theatre suite to the time they entered the anaesthetic room is given, see Figure 7.2.1ii.
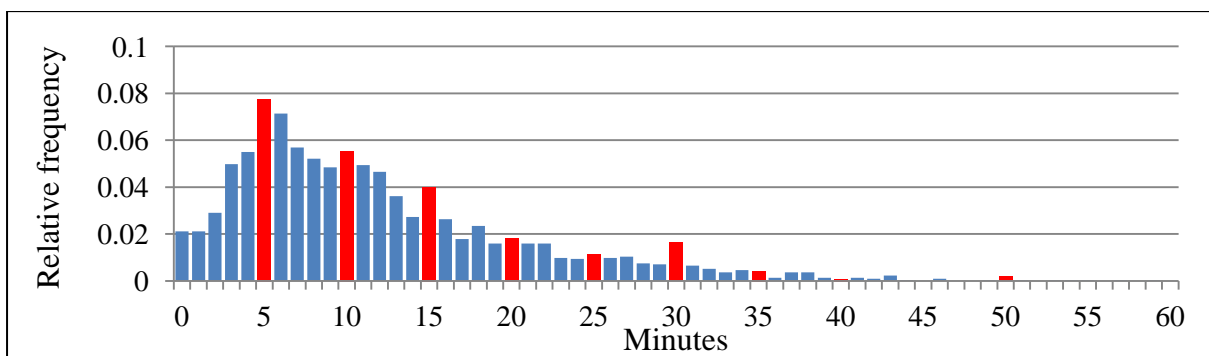

Figure 7.2.1ii: Distribution of time spent in the interval *arrived – into anaesthetic room* for patients undergoing trauma hip surgery on a minute-by-minute basis

214

It can be seen that there are peaks at multiples of five minutes, in comparison to the surrounding bars – these are highlighted in red. Although the actual times that are recorded are when these two events occur, and not the time taken for interval, it appears that values may have been rounded or estimated in order to facilitate ease of entry into the TheatreMan system. This awkwardness of shape goes some way to explain why the distribution-fitting exercise has proved to be problematic.

Another issue is that some intervals have a high frequency of zero values. As an example, the time taken between entering the anaesthetic room and the anaesthetic procedure starting is considered. It can be seen in Figure 7.2.1iii that this is the case here and under these circumstances the bimodal shape is a major cause of difficulties encountered in attempting to fit a statistical distribution. The probabilities have been calculated twice, firstly including the zeroes, then excluding them. To combat this issue, this amended distribution could be used; firstly, it would be decided probabilistically whether or not a value of zero is taken and if not, then the second distribution is sampled from. This method of removing zeroes was tried for each of the intervals under consideration in this section but again no fits could be found.
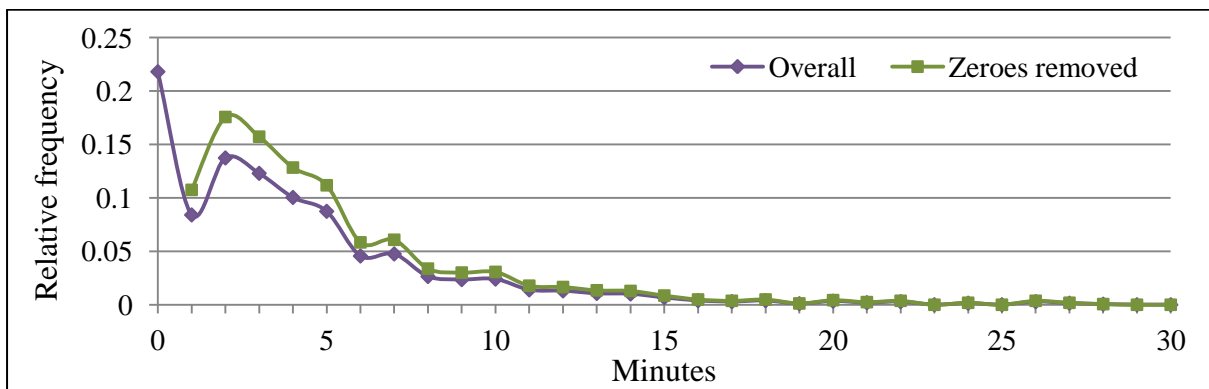
Figure 7.2.1iii: Distribution of time spent in the interval *into anaesthetic room – anaesthetic start* for patients undergoing trauma hip surgery

One other way to tackle both of the issues discussed here is to combine several values into one interval; for example, 0 – 4 minutes, 5 – 9 minutes, etc. This is likely to solve the issue of peaks as displayed in Figure 7.2.1ii and possibly also the second problem discussed, but it is not particularly desirable due to loss of information. It would also pose problems to any

distributional fitting exercise. The graphical distribution of the four time intervals not presented thus far can be seen in Figure D7.2.1a.

7.2.2   Non-independent time intervals

To confirm that it was correct to perform separate analyses on three of the time intervals for each type of operation, it was first checked to see whether times did differ between operation types. Only operation types with ten or more entries were included in the analysis.

For each of the three time intervals, the *p*-value given by the Kruskal-Wallis test was found to be less than 0.0001, leading to a rejection of the null hypothesis there that is no difference in the time taken for each respective interval among the operation types. These intervals are likely to be dependent on a number of other factors too, some of which are discussed forthwith.

The American Society of Anaesthesiologists (ASA) grade of a patient is likely to have an effect on their performance in each of these three time intervals (see Section 2.2.4 for a description of ASA grade). The time taken to perform anaesthesia is expected to be shorter for a much healthier patient compared with one who is sicker, as there are more factors affecting the procedure and a higher risk of complications and/or death for patients with a higher ASA grade (Aitkenhead 2005). However, the ASA grade scoring system assumes that a patient's age has no relation to their physical fitness, which is not necessarily true. As an example, an octogenarian patient with an ASA grade of I is not as fit and healthy as a teenager with an ASA grade of I. However, in view of the fact that the majority of patients here are elderly, this may not pose too much of a problem. Merging the TheatreMan data with the Cardiff Hip Fracture Survey data would mean that ASA grade could be obtained for the hip fracture patients in the TheatreMan dataset using each patient's unique identifier as the merging variable. However, at the time of this study not enough data was available from the Cardiff Hip Fracture Survey dataset which recorded this information and thus this idea could not be taken further, but is a possible avenue to explore in the future.

Another factor which may influence the time taken to complete these intervals is the experience level of the clinical team. See Section 7.3.1 for an in-depth investigation into

whether or not the level of experience of the surgeon impacts upon surgery procedure time. No such information was available regarding the anaesthetist for each operation, but there is some evidence to suggest that their experience level does influence the anaesthetic procedure time and outcome (Byrne and Jones 1997, DeAnda and Gaba 1991).

The merging exercise proposed previously would also provide the opportunity to investigate if any other patient characteristics, aside from ASA grade, affect their time in theatre. For example, one study found that male gender and younger age were associated with longer operating times for primary hip replacement surgery (Småbrekke et al. 2004).

These intervals are not looked at in more detail here but are considered both later in this chapter and in Chapter 8.

7.3   Preliminary analyses

Two separate preliminary analyses regarding the time taken to complete the surgical procedure for hip fracture patients were undertaken on the request of the Theatre Manager at the University Hospital of Wales. While these were standalone ad-hoc analyses, the results were also useful from a modelling perspective in terms of identifying important and unimportant issues or parameters relating to the trauma theatre.

7.3.1   Effect of surgeon experience

Several studies have investigated the performance of physicians and have suggested that increasing age (of the physician) is related to inferior patient outcomes, especially when combined with a decrease in patient volume (Eva 2002, Niteesh et al. 2005), but that older physicians may have better diagnostic skills despite being less aware of newer medicines and techniques (Eva 2002). Specifically for surgeons, patient mortality has been shown to increase with surgeon age across some, but not all, procedures (Hartz et al. 1999, Waljee et al. 2006). While these studies provide useful insight, death during surgery for hip fracture patients is very low (Parvizi et al. 1999, Wachtl et al. 2003). It has also been shown that the qualifications of the surgeon have no impact upon patient mortality for hip fracture (Siegmeth et al. 2005), thus the primary question to answer here does not relate to mortality, but whether experience of the surgeon has any impact upon the time taken for the surgical procedure.

The condition of bone tissue in the femur has been suggested as the most important factor influencing length of operation in hip revision surgery (Frey 2010), while longer operations have been shown to increase the chances of infection (Leong et al. 2006, Malik et al. 2004). Here the issue to consider is whether the experience of a surgeon has an impact upon the time taken to perform the operation; clearly evidence shows that a shorter operation is desirable. Surgeons are classified as experienced if they are at consultant level and are classified as less experienced, or trainees, otherwise.

In order to remove bias regarding the type of operation, univariate statistical analyses were performed for each type of hip surgery (as classified by OPCS-4 code and where sufficient data would allow) in order to determine whether there were significant differences between the length of operation and surgeon experience. In each case, the Wilcoxon test was used and

218

all values were calculated using SAS. Procedures were only considered if there were ten or more observations for both levels of surgical experience, leaving seven procedure types available for further scrutiny.

Table 7.3.1i: Results of length of operation (minutes) by operation type and surgeon experience level

OPCS-4 code	Experience level	n	Mean	S.D.	p-value
W19.1C	Consultant	26	38.1	32.9	0.0102
	Trainee	92	43.2	22.7	
W19.1D	Consultant	10	47.5	16.0	0.8464
	Trainee	68	49.7	23.8	
W19.1E	Consultant	60	45.2	21.8	0.0279
	Trainee	263	50.5	21.9	
W46.1C	Consultant	28	63.0	21.8	0.0005
	Trainee	70	74.4	18.8	
W46.1E	Consultant	47	75.5	35.2	0.0551
	Trainee	53	64.6	18.2	
W46.1F	Consultant	10	50.9	15.7	0.0273
	Trainee	48	65.9	15.9	
W47.1B	Consultant	35	36.2	14.5	< 0.0001
	Trainee	116	48.4	18.0	

Using a significance level of 5%, the level of experience of the surgeon suggests an impact on operation time in five of the seven procedures, with operations performed by consultant surgeons being shorter. One interesting case is that of the procedure W46.1E (hemiarthroplasty): the results for this operation suggest that trainee surgeons are in fact quicker than consultant surgeons, although these differences were not found to be statistically significant.

Matters and problems relating to the training of surgeons is well-documented; with issues relating to capacity (Crofts et al. 1997), obligatory responsibility, workload and unrest between trainees and trainers (Chikwe et al. 2004, Murday et al. 2000) all discussed in the

literature. It is also worthy of note that problems exist in particular with the training of orthopaedic surgeons (BOA 2002). Clearly the training of new surgeons is a complex issue.

While these results are interesting from a statistical and clinical viewpoint, it was decided that this matter would not be taken further for this project. The emergency theatre is modelled in Chapter 8 and while it may be interesting to investigate the effect on the theatre of altering the level of experience of the surgeons (to assess the impact of all operations being undertaken by consultants, for example), the discussion here has shown that this is not really a realistic avenue to explore. It is simply infeasible to make managerial changes of this type in a timely fashion.

7.3.2 Effect of timing of surgery

Another piece of exploratory data analysis performed involved the investigation of the timing of surgery on the length of operation. It was thought that there may be differences between those performed before 17:00 and those performed after 17:00; these are henceforth categorised as day and night cases respectively.

There have been few studies in the literature that have evaluated surgical measures or outcomes associated with operative start time, while those found tend to concentrate primarily on clinical outcome. For example, one study found that there were no differences in overall mortality and morbidity rates for operations performed at night (22:00 – 06:00) compared with those performed during the day (06:00 – 22:00) for laparoscopic cholecystectomies and appendectomies, as well as no difference in the median length of operation for both types of surgery (Yaghoubian et al. 2010). However, another study found that time of day of surgical start time had a moderately strong association with mortality for non-emergency cases starting in the time period 21:30 – 07:30, while the effect on morbidity was also the strongest in this overnight cohort (Kelz et al. 2009). Similarly, worse outcomes have been found for overnight resuscitation (23:00 – 06:59) after cardiac arrest compared with daytime resuscitation (Peberdy et al. 2008), as well as an increased risk of morbidity for non-emergent general and vascular surgical procedures for operations starting after 16:00 compared with those starting between 07:00 and 16:00 (Kelz et al. 2008).

Specifically for orthopaedic surgery, it has been shown that after-hours (16:00 – 06:00) operative times were shorter compared with a daytime (06:00 – 16:00) group, for tibial and femoral nail fixation groups. Despite this, minor complication rates and the number of unplanned reoperations were higher in the after-hours group, leading to the overall conclusion that an increase in allocated orthopaedic surgery time during the day would be desirable (Ricci et al. 2009).

For hip surgery, it has been shown that some types of operation (dynamic hip screw, intramedullary nails) took significantly longer at night (18:00 – 06:59), while there was no difference for hemiarthroplasties. There were also no differences found between the day and night groups with respect to mortality at one month, one year and two years, as well as in other noted complications (Chacko et al. 2011). Similarly it was concluded that patients perform equally well if they are operated on at night (after 21:00) compared with during the day; with no differences reported between the two groups with respect to complication rates and non-statistically significantly differences reported with respect to mortality (Dorotka et al. 2003a). It has also been reported that duration of surgery and the incidence of peri-operative complications for total hip arthroplasty may be greater for later surgery start times, but the statistical differences are small and thus it is concluded that they are unlikely to be significant in a clinical setting (Peskun et al. 2012). This is supported by a Dutch study which found that there was no increased risk of complications or mortality for pertrochanteric fractures operated on outside of working hours (17:00 – 08:00) (Bosma et al. 2010).

In order to eliminate any bias imposed by surgeon experience, as discussed in Section 7.3.1, cases were split by surgeon type as well as operation type. The time category was determined by the start time of the operation and weekends were excluded from calculations. As previously, the statistical analysis was only completed if there were ten or more observations in each group.

In each case, when classified by start time of operation (day or night), after an initial split by surgeon experience and operation type, no differences in length of operation were found at a significance level of 5%. As an example, consider the largest group available: operation type W19.1E (open reduction with internal fixation intertrochanteric fracture / dynamic hip screw) with trainee surgeons. Removing weekend cases left 106 and 45 observations for the day and night groups respectively. The corresponding mean lengths of operation were 51.2 and 49.7

minutes, with standard deviations of 23.6 and 22.1 minutes. The data is displayed in a standard box-and-whisker plot in Figure 7.3.2i where the whiskers represent the 5th and 95th percentiles.
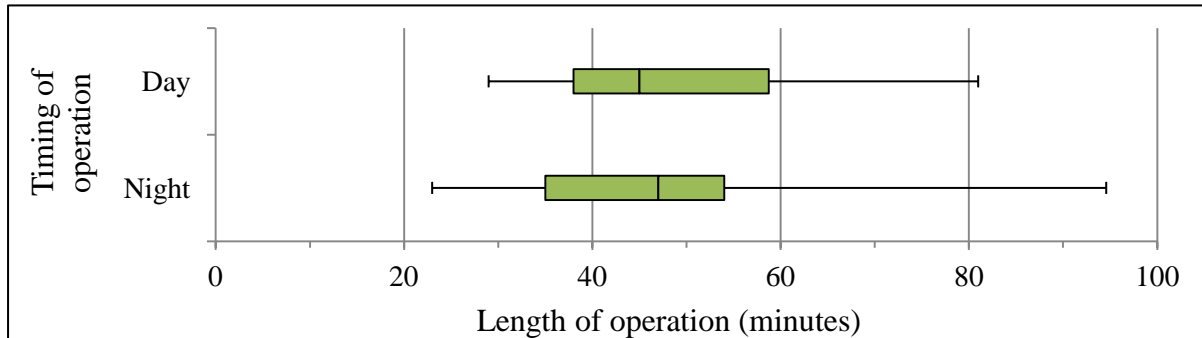
Figure 7.3.2i: Comparison of length of operation by surgical start time: operation type W19.1E, trainee surgeons, weekends excluded

The *p*-value in this case was 0.8486, supporting the notion suggested by the summary statistics and Figure 7.3.2i of no difference in length of operation when classified by surgical start time.

Mortality and morbidity analysis is not possible here using the TheatreMan data since the data collection does not extend beyond the patient's pathway through the theatre. However, for some of the later data collected on the ward there is information on the time of the operation. An operation during the day is thus classified as starting between 07:00 and 17:00. A split is made here by ASA grade in order to create more homogenous groups in terms of medical fitness, leaving 357, 576 and 93 observations available for ASA grades I&II, III and IV respectively; this was for patients for whom the operation time and acute outcome (in terms of survival) was known. Results of the percentage of deaths on the acute ward within each group are given in Table 7.3.2ii. The quoted *p*-value relates to a Chi-square test of independence between surgical start time and death on the acute ward.

Table 7.3.2ii: Percentage of acute ward deaths by surgical start time, split by ASA grade

ASA grade	Surgical start time		p-value
	Day	Night	
I&II	4.7	7.4	0.3417
III	14.2	13.6	0.8726
IV	28.6	26.1	0.8178

It can be seen that the percentage of deaths in each surgical start time group is relatively similar across each of the three ASA grade categories, suggesting that there is no time-of-day effect on mortality. This is further supported by the Chi-square test of independence performed in each case.

These investigations indicate that the timing of the operation for hip fracture patients has no significant effect on both the length of the operation and acute ward outcome. This could have been something to include in the modelling of the trauma theatre – for example, in terms of a what-if scenario regarding the operative timing of this cohort of patients – but evidence here suggests that doing so would not be of any particular value. For this reason, while results found here are informative, this specific area is not pursued further.

7.4 All trauma patients

Throughout the course of this project, a more comprehensive dataset became available which included all operations performed within the emergency trauma theatre at the UHW. A particular area of interest was to look in more detail at spinal operations. These operations are typically very complex and thus can occupy the theatre for some considerable time. If these operations could be performed elsewhere, then more theatre time would be available for trauma hip operations, as well as any other emergencies which present themselves. The type of operation performed will therefore be in one of three general categories: *Hip*, *Spinal* or *Other*. There will of course be a huge variation in the type of operation performed in the *Other* category, but further breakdown is not required.

In some cases patients will undergo more than one operation at a time and while these operations are often of the same type, this is not always the case. The database used records which operation was the main procedure for each patient. Where this information was unavailable, if one of the procedures was either a hip fracture or spinal operation, then this was recorded as the primary procedure.

A total of 8975 procedures were performed over the time period for which data was available. This data spanned a total of 1365 days (195 weeks, approximately 3.7 years). This was a total of 7935 theatre episodes; the average number of procedures per episode per patient being 1.06, 1.20 and 1.15 for the main procedure types of hip, spinal and other respectively. This equates to 1.13 procedures per patient episode overall. A breakdown by type of primary procedure is given in Table 7.4i, along with row and column percentages in parentheses. Six episodes are excluded since the operation code was not recorded.

It can be seen that a relatively small percentage of patients have three or more procedures during a single theatre episode and so further classification beyond this is not necessary. This category is made up entirely of patients having three, four of five procedures, with the exception of one patient who underwent eleven procedures at one time.

Summary statistics on the length of operation are given in Table 7.4ii. It may be surprising to see minimum values of zero in some cases, but these can be justified by a patient deteriorating during the anaesthetic procedure and therefore the operation itself being cancelled, for example.

Table 7.4i: Summary of the number of procedures performed per patient episode by type of primary procedure

Primary operation type	Number of procedures			Total (%)
	1	2	3 or more	
Hip	1971	122	11	2104 (26.5)
Spinal	324	55	11	390 (4.9)
Other	4739	601	95	5435 (68.6)
Total (%)	7034 (88.7)	778 (9.8)	117 (1.5)	7929

Table 7.4ii: Summary statistics for operation time (minutes; procs – procedures)

Type	No. of procs	*n*	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurt-osis
Hip	1	1961	62.44	31.65	0	284	1.37	3.39
	2	121	103.38	55.13	1	291	1.02	1.27
	3+	11	164.27	107.92	39	364	0.59	-0.83
Spinal	1	321	87.46	51.11	5	327	1.25	2.23
	2	55	123.55	64.62	5	304	0.52	0.15
	3+	11	212.18	58.49	135	322	0.61	-0.18
Other	1	4728	46.16	36.62	0	352	1.84	6.16
	2	601	60.02	51.45	0	394	1.86	4.87
	3+	95	136.02	99.92	11	424	0.86	-0.11

It appears that there are differences in operation length between the three types. Indeed statistically the differences are significant not only between types but also within each type, when split by the number of procedures; $p < 0.0001$ in each case. Box-and-whisker plots are a useful display aid to show these differences. Figure D7.4a of Appendix D shows the distribution of operation times for each of the three operation types, broken down by the number of procedures performed per theatre episode. The whiskers display the 1st and 99th percentiles. Any values outside of this range are displayed with a dot.

For hips, it is immediately obvious that as the number of procedures performed increases, the time taken to complete the operation also increases, as does the variation in the operation time. This last observation however may be due to the smaller numbers seen in the groups with two and three or more procedures. Despite the same pattern being seen in terms of more

procedures take longer to perform, as is intuitive, the difference in the variation of operation length between the three spinal groups is less obvious. More procedures are again shown to have a longer length of operation for other operations too. There is also more variation in operation time for these operations, although this diminishes if the coefficient of variation is considered instead of the standard deviation.

Next the time taken to perform the anaesthetic procedure is considered; summary statistics are given in Table 7.4iii. This anaesthetic time is classified here as the combination of two original intervals: the time taken from the patient entering the anaesthetic room to the anaesthetic procedure starting, plus the time taken from the anaesthetic procedure starting to when the patient enters the operating theatre. Box-and-whisker plots are given in Figure D7.4b; note that one value (386 minutes, *Other* category) is excluded for display purposes.

Table 7.4iii: Summary statistics for anaesthetic time (minutes)

Type	No. of procs	*n*	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurt-osis
Hip	1	1960	24.93	13.28	0	148	1.84	7.72
	2	121	23.28	16.65	1	133	3.10	15.80
	3+	11	17.82	10.65	1	37	0.16	-0.36
Spinal	1	323	21.03	13.94	0	99	1.78	5.55
	2	55	21.84	18.61	0	110	2.63	9.28
	3+	11	34.27	17.68	10	76	1.16	2.46
Other	1	4725	15.88	11.61	0	386	8.58	227.87
	2	601	16.45	12.74	0	195	5.92	70.12
	3+	95	19.08	17.24	0	114	2.45	9.89

While differences can be seen between the groups, these are less pronounced than with the length of operation. Testing for these differences, it is shown that while each of the three types of operation are not only different from each other but that there are also differences within each type of operation for hip and spinal operations, with respect to the number of procedures performed. However, further pairwise investigations showed that these differences were caused by the *3+* group in both cases. Since these groups are small, it was

decided to combine all groups within the three operation types. All tests were carried out at the 5% level of significance.

It still makes sense to keep the division of groups by the type of operation; not only is this evident from the data but it also confirmed medically. Some of the operations in the *Other* category will have a simple local anaesthetic, while a spinal operation is most likely to require a more complicated anaesthetic procedure; for major spinal surgery, anaesthesia presents a number of challenges (Raw et al. 2003). These differences are shown graphically in Figure 7.4iv (again with one observation omitted), while summary statistics for anaesthetic time are given in Table 7.4v.
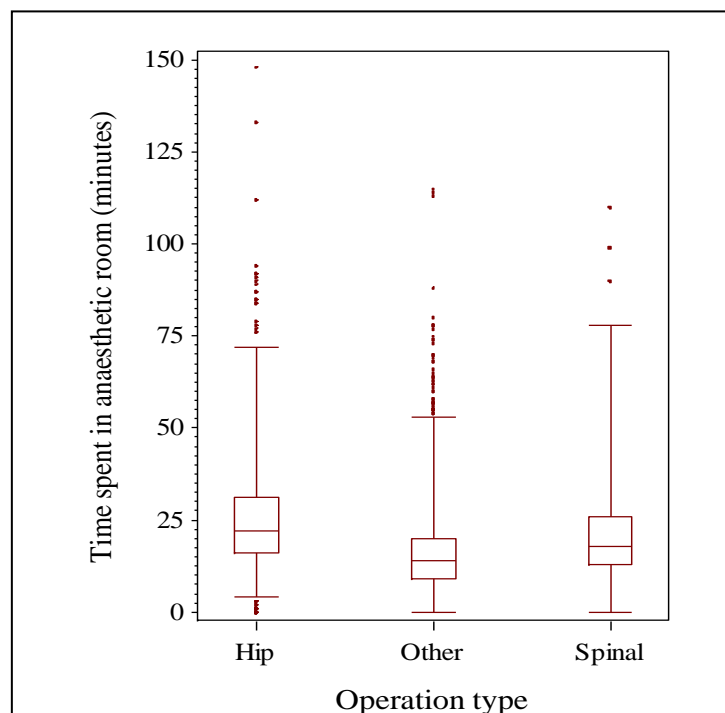
Figure 7.4iv: Distribution of anaesthetic time by type of operation

Table 7.4v: Summary statistics for anaesthetic time (minutes), by type of operation

Type	n	Mean	S.D.	Min	Max	Skewness	Kurtosis
Hip	2092	24.80	13.50	0	148	1.96	8.75
Spinal	389	21.52	14.91	0	110	1.99	6.63
Other	5421	16.00	11.86	0	386	7.96	192.98

7.4.1 Correlation between intervals

In addition to the two investigations presented in Sections 7.3.1 and 7.3.2, another final issue to investigate is the relationship between values recorded for different intervals. This is completed solely for hip operations and the entire group is looked at without any further categorisation.

Correlation is a bivariate analysis that measures the strength of association between two variables. Output includes the correlation coefficient, ρ, which will lie in the range [-1, 1]; a value of $\rho = 1$ signifies perfect positive correlation, while $\rho = -1$ signifies perfect negative correlation. Although an absolute value of one therefore means that there is a perfect degree of association between the two variables, it does not imply cause and effect. Note also that this method measures the degree of *linear* association between two variables.

Spearman's rank-order correlation test is a non-parametric test used to measure the degree of association between two variables, which does not impose any assumptions about the variables under consideration. (The widely-used Pearson correlation test assumes that both of the variables being used come from a Normal distribution which is not the case here.) Spearman's test is based on the ranks of the data values, rather than the actual data values themselves.

Comparing all original time intervals against each other would result in the evaluation of $^{10}C_2 = 45$ different correlation coefficients (ten intervals were available in this case due to the inclusion of *into theatre* in the dataset used). Additionally, combining adjacent time intervals means that there would be even more to consider.

Instead four new time blocks are created based on combinations of the existing intervals, as follows:

Stage A – Pre-theatre (*asked for – into anaesthetic room*)

Stage B – Anaesthetic procedure (*into anaesthetic room – operation start*)

Stage C – Operation time (*operation start – operation finish*)

Stage D – Recovery (*operation finish – out of recovery*)

Results are given in Table 7.4.1i, note that recovery time was not available for much of the dataset. Results are given vertically in the order: ρ, p-value against the null hypothesis that $\rho = 0$, n.

Table 7.4.1i: Correlation results between theatre times for trauma hip surgery

	B. Anaesthetic procedure	C. Operation time	D. Post-operation
A. Pre-theatre	0.1014 < 0.0001 2072	0.0299 0.1738 2071	0.0311 0.1023 1118
B. Anaesthetic procedure		0.1392 < 0.0001 2093	0.0562 0.0602 1118
C. Operation time			0.2073 < 0.0001 1119

Half of the comparisons made indicate a significant correlation (A *vs*. B, B *vs*. C, C *vs*. D), while the other half indicate no significant correlation. All significant values of ρ were positive, as would be expected, but the maximum value found was 0.2073 (C against D), indicating a weak correlation. Figure 7.4.2ii is a scatterplot of times for block C against block D, which clearly displays this weak relationship. The axes have been curtailed for display purposes meaning that six observations are missing from the presented plot.
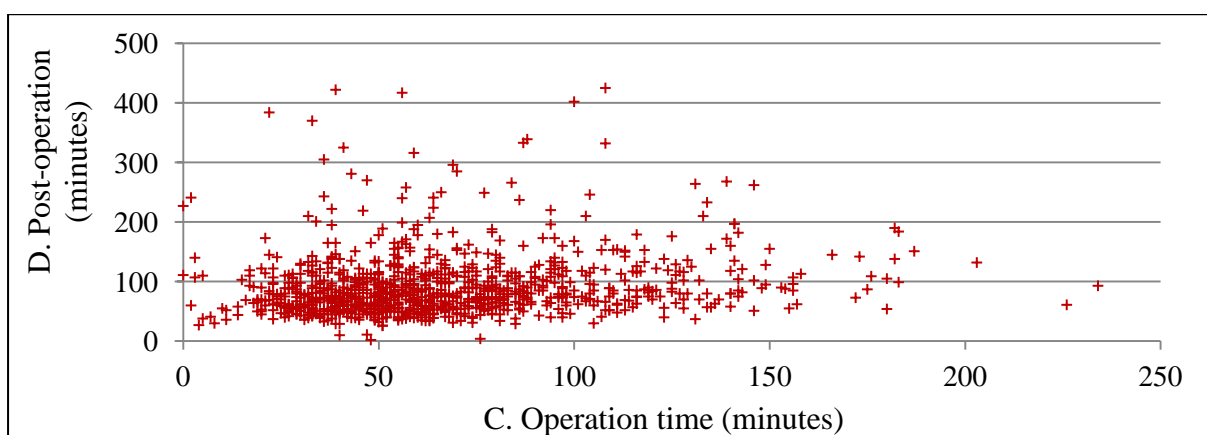

Figure 7.4.1ii: Relationship between operation time and post-operation time (minutes) for trauma hip surgery

7.4.2   Theatre utilisation

The task of allocating theatre time for orthopaedic trauma surgery while attempting to maintain an optimum (or high) level of theatre utilisation is difficult due to the unpredictability of trauma, both in terms of demand and case duration. It is not, however, only the instability of the system which causes problems in this area; one American study has identified eight key areas in which improvements can be made in the management of operating theatres which "could be directly linked to increased revenue, patient safety benchmarks, and potentially, staff satisfaction" (Girotto et al. 2010).

A large proportion of a hospital budget is represented by the operating theatre suite. One estimate is that 30.1% of all healthcare outlays are related to surgical expenditures (Muñoz et al. 1994) – and thus maximum utilisation is necessary to ensure optimum cost–benefit (Jan et al. 2003).

The daily planned busy time for the trauma theatre at the UHW is 11.5 hours, which is calculated as the time from when the first patient enters the anaesthetic room to when the last patient leaves the theatre. This 11.5 hour slot covers the time from 8:30am to 8:00pm. There are a multitude of reasons to explain why sessions may start or finish early and/or late, as well as why the total theatre busy time may surpass or fall short of the allocated time available, not least because of the theatre under scrutiny accommodates emergency patients and there is the inherent unpredictability that this brings.

Tardiness relating to the start time of theatre lists is not uncommon. For example, one study has shown a delayed start time of over thirty minutes in more than half of cases for the trauma theatre list (Rethnam et al. 2009), while another found that each orthopaedic trauma session started 18.8 minutes late on average, with just 8.2% starting early (Delaney et al. 2010). Average start delays of 18 minutes (Durani et al. 2005) and 26.5 minutes (Ricketts et al. 1994) have also been reported by two large orthopaedic centres in London. Late starts may be attributable to the anaesthetic staff, theatre staff or surgeons (Ricketts et al. 1994) and/or due to the delay in transferring the first patient from the ward to the theatre suite (Delaney et al. 2010).

Data for the trauma theatre at the UHW (over the period September 2005 – May 2009) shows that 79% of theatre sessions started late, with the average start time 26 minutes later than

scheduled. The discrepancy in the end time is more balanced; 55% of sessions were shown to finish early, while on average the sessions finished 18 minutes earlier than scheduled.

These results varied considerably by day and some summary values by day of the week are given in Table 7.4.2i. The differences on which these values are based are calculated as the number of minutes elapsed between the allocated time and the actual recorded time; positive values thus indicate lateness while negative values indicate earliness.

Table 7.4.2i: Differences between allocated and actual times for trauma theatre start and end times (minutes)

Day	Start time difference		End time difference	
	Mean	**S.D.**	**Mean**	**S.D.**
Sunday	46.7	26.4	-27.7	80.0
Monday	19.8	30.1	-12.0	79.6
Tuesday	22.4	31.1	-27.7	81.2
Wednesday	15.6	26.7	-7.8	67.6
Thursday	11.1	21.7	-7.9	76.6
Friday	20.1	24.5	-22.1	69.6
Saturday	47.1	26.5	-23.4	68.6

Weekend theatre sessions started on average 47 minutes late compared with an average of 18 minutes late during the week. Weekend sessions were also more likely to finish early; on average finishing 26 minutes early compared with weekday sessions which finished an average of 15 minutes early. However, there is some considerable deviation to be seen in these results; with standard deviations of 30.0 minutes and 75.2 minutes for start and end time differences respectively across all days.

The reason for late starts becomes apparent after scrutiny of the TheatreMan dataset. Overall the average time that the first patient of the day was asked for was 8:28am and on only 63% of days was the first patient asked for before the planned start time of 8:30am. On average the first patient arrived at the theatre suite twenty minutes later, at 8:48am, and the first patient had only arrived by 8:30am in 32% of cases.

Theatre utilisation is now investigated and discussed. One study used simulation to conclude that, without patient delays and staff overtime, an operating theatre utilisation of 85–90% is the optimum that can be achieved and thus provides the best cost–benefit (Tyler et al. 2003). This is consistent with The Bevan Report, a large scale audit in hospitals in the United Kingdom which concluded that an acceptable standard value for operating theatre utilisation was 90% (NHS Management Executive 1989). Thus theatre utilisation is not necessarily expected to be 100% and indeed it has been recommended that it should not be at this value.

Observed values quoted in the literature vary across studies. One study found that the trauma theatre was opened for 8.71 hours out of the available 12 hours (73%), but despite this underuse of the theatre, an average of 86 minutes of trauma surgery happened outside of the trauma theatre each day (Collantes et al. 2008). Another survey found that operating theatre utilisation was 81% over a six month period, while end utilisation was 78.8% (Delaney et al. 2010). End utilisation is discussed further in Section 7.7.

The average time that the theatre was busy for per day for the trauma theatre at the UHW was 10.8 hours, standard deviation 1.35 hours (81 minutes), which equates to a mean value of 93.7% utilisation. The observed minimum busy time was 4.7 hours (41% utilisation) and the maximum was 14.8 hours (129% utilisation), with the 50th percentile at 10.9 hours. The distribution of theatre busy time is given in Figure 7.4.2ii.
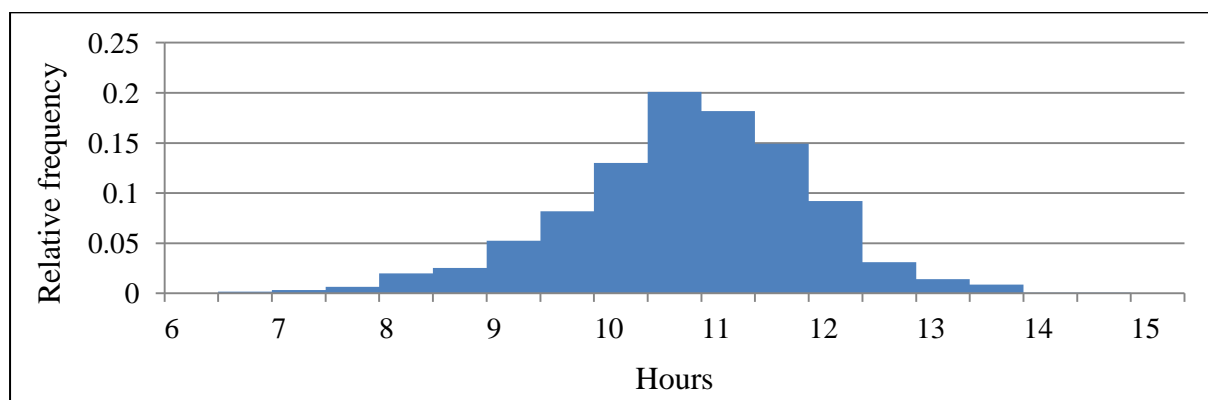

Figure 7.4.2ii: Distribution of theatre busy time for the trauma theatre

7.5   Cancellations

The NHS has identified three key causes as to why operations may be cancelled; hospital non-clinical, hospital clinical and patient reasons (NHS 2008), the occurrence of which are stated to be "distressing and inconvenient for patients" (Commission for Health Improvement 2003).

Cancellation data was available for the trauma theatre at the UHW from September 2005 to May 2009. There were 53 distinct reasons recorded why operations were cancelled, ranging from the patient being unfit for surgery to there being no blood available. These were then reclassified into the three causes defined by the NHS. 54.0% of all cancellations were due to hospital non-clinical reasons, 33.6% due to hospital clinical reasons and 1.9% attributable to reasons relating to the patient. The reason was unclear or not recorded in the remainder (10.5%) of all cases.

Non-clinical hospital reasons are thus the most prominent cause for a cancellation. Of these, cancellations due to lack of time account for the majority of cases; this is the cause recorded for 39.6% of all cases (and 44.3% of all cases for which the cancellation reason was known). The high frequency of these and the actuality that this is the cause that is the most tackleable mean that it is the one concentrated on here. Many (if not all) of the others cannot easily be influenced by a change in how the theatre is run. Reasons for running out of time include unreliability of the theatre schedule (Klimek et al. 2008), bureaucracy (Bone and Hooker 2007) and the difficulty in estimating the length of time an operation will take to complete, despite some recent advances made in this area (Eijkemans et al. 2010).

Summary measures of the number of cancellations per day are given in Table 7.5i. As a reference, recall that 7935 procedures were completed over the same time period. The distribution of the number of cancellations per day is given in Figure 7.5ii; values recorded by the *y*-axis are relative to each group.

Table 7.5i: Summary measures for the number of cancellations per day

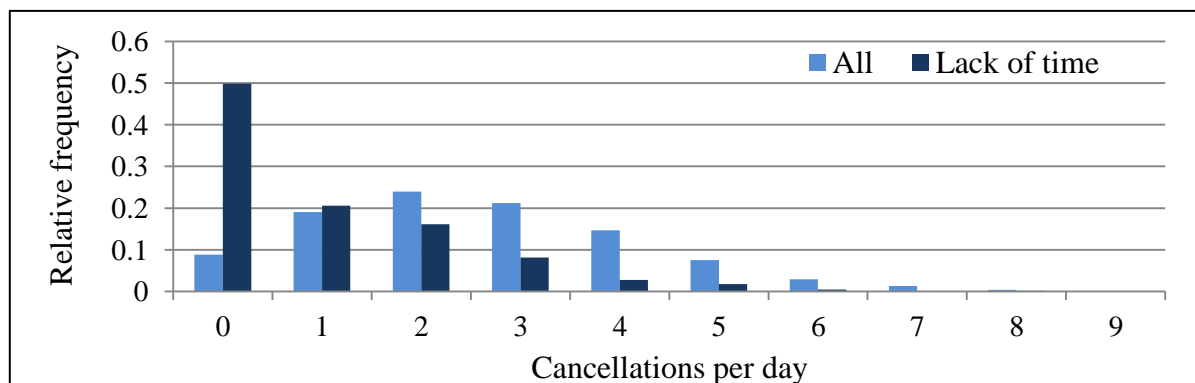| Reason for cancellation	Total	Cancellations per day				
		Mean	**S.D.**	**Median**	**Minimum**	**Maximum**
All	3520	2.57	1.63	2	0	9
Lack of time	1394	1.02	1.32	1	0	8

Figure 7.5ii: Distribution of the number of cancellations per day

Lack of time cancellations occur on around half of all days, while the modal number of daily cancellations across all groups is two. However it can be seen from Table 7.5i that on average just over one operation is cancelled due to lack of time each day. It is rather alarming that on two occasions eight operation were cancelled in a single day because of lack of time, but this may be explained by, for example, a large road traffic accident resulting in many people requiring urgent and unexpected surgery. Unfortunately there is no easy way to plan for this and so in cases such as these cancelled operations are unfortunate but necessary occurrences.

It would be useful to know what proportion of cancelled operations belong to each of the operation type groups. The type of procedure cancelled is recorded but there are no operation codes or standardised descriptions, which has resulted in 2691 different types of operation reported to be cancelled! Clearly this figure is unrealistic and the reason for it is inconsistency in the recording of operation types. It is well-documented that a better computerised system, which recognises words while maintaining ease of data input, would be advantageous in circumstances such as this (Jones et al. 2003).

As an example, consider right-sided dynamic hip screw (DHS) operations; just some of the many relevant recorded values here are RIGHT DHS, DHS RIGHT, DHS (R) HIP and RIGHT DYNAMIC HIP SCREW. By searching through the database for relevant keywords, it was possible to flag those which are hip operations but it is difficult to know whether all of the hip operations have been captured (due to spelling mistakes, for example). Despite these difficulties, this exercise was completed in order to obtain an estimate of the proportion of cancelled operations which were for a fracture of the hip.

The data suggests that around 19% of all cancellations were hip fracture patients, while around 25% of lack of time cancellations belonged to this patient group. This translates to 14% of all hip operations being cancelled. These are most likely to be underestimates due to the issues explained previously, as well as the knowledge that these patients are often moved to the end of the schedule in favour of other patients, meaning that they are more likely to be cancelled due to a shortage of theatre time. It cannot, however, be an overestimate.

7.6   Turnover times

The trauma theatre can never be fully operational; there will always be some turnover time between operations. The theatre must be cleaned after the previous operation and prepared for the forthcoming operation. The movement of equipment, medical apparatus, staff and indeed the patients undergoing surgery will occur in this time. This time will therefore be dependent on the operation that has just finished, as well as the next operation which will take place. Understanding the occurrence and magnitude of turnover times is important, not least because it has been suggested that the extra expense incurred to improve throughput (as measured by a decrease in turnover times) is more than offset by the financial gains of improved efficiency (Krupka and Sandberg 2006).

There is a separate room, which adjoins the theatre, in which the anaesthesic procedure is carried out. Turnover time can thus be defined in two ways: (i) the time between one patient leaving the theatre and the next patient entering the anaesthetic room; or (ii) the time between one patient leaving the theatre and the next patient entering the theatre. Using the three operation types seen earlier, there are 3^2 different permutations of sequence available. Table 7.6i gives the number of occurrences (and row percentages) of each of the sequences.

Table 7.6i: Frequency of ordering of operations in the trauma theatre

Preceding operation	Following operation		
	Hip	Spinal	Other
Hip	678 (38.4%)	38 (2.2%)	1051 (59.5%)
Spinal	50 (14.3%)	112 (32.0%)	188 (53.7%)
Other	1072 (24.3%)	124 (2.8%)	3213 (72.9%)

7.6.1   Anaesthetic room / operating theatre turnover

Summary statistics on the turnover times between these operations are now presented for the first definition of turnover time. The first column refers to the order of operation; for example, *SH* refers to a hip operation following a spinal operation. Values of *n* may differ

from those in Table 7.6i due to missing data. Negative values are considered to be valid here since simultaneous anaesthesia and surgical treatment can be given for two different patients, but values less than -30 minutes were ignored. The distribution of these times is displayed by means of a box-and-whisker plot in Figure 7.6.1ii, where the whiskers represent the 1st and 99th percentiles. The mean overall anaesthetic room / theatre turnover time was 23.4 minutes.

Table 7.6.1i: Summary statistics for anaesthetic room / theatre turnover times (minutes)

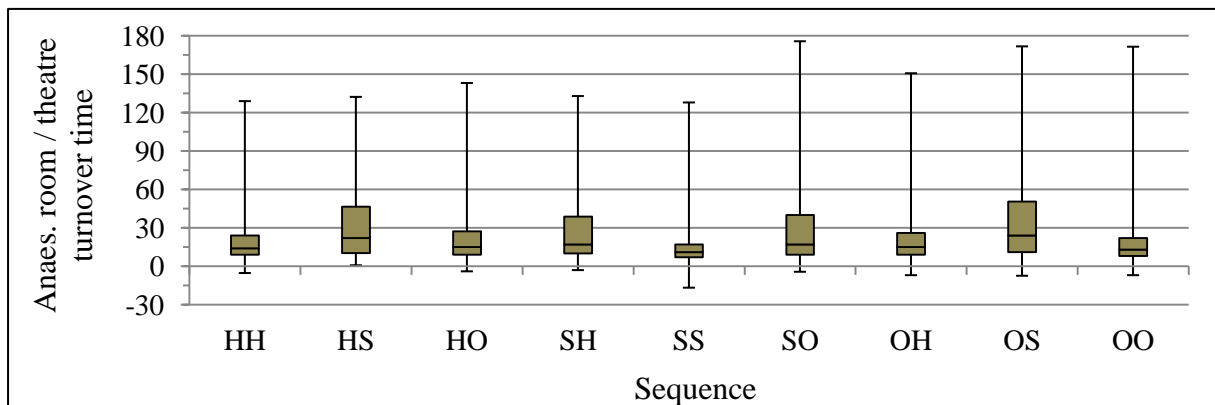| Sequence	*n*	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurtosis
H H	669	22.21	27.24	-20	345	4.76	38.32
H S	38	36.87	37.54	0	143	1.29	0.66
H O	1048	23.71	29.67	-28	315	4.35	29.28
S H	50	28.58	30.55	-5	154	2.13	5.48
S S	109	16.00	23.88	-20	155	3.89	18.06
S O	185	30.61	35.91	-25	193	2.28	6.19
O H	1064	24.66	37.59	-25	792	9.79	170.22
O S	123	35.19	34.93	-25	183	1.73	4.17
O O	3189	22.19	34.45	-30	488	5.60	45.09

Figure 7.6.1ii: Distribution of anaesthetic room / theatre turnover time (minutes)

Considerable differences between the turnover times by sequence of operations can be seen. It interesting to note that the turnover time between two spinal operations is smaller than for any other sequence, as well as having the lowest standard deviation, but when a spinal operation follows one of the two other types, the largest turnover times are seen.

237

7.6.2 Operating theatre turnover

Consider now the turnover time for the operating theatre. Negative values are not possible in this case since the treatment of patients cannot overlap in terms of theatre use. Instances where this situation was reported by the data were regarded as invalid.

It is clear that shorter and less variable turnover times tend to be seen when like operations occur in sequence, while the longest turnover times relate to when a spinal operation follows either hip or other surgery. The distribution of these times is displayed by means of a box-and-whisker plot in Figure 7.6.2ii. The mean overall theatre turnover time was 41.7 minutes.

Table 7.6.2i: Summary statistics for theatre turnover times (minutes)

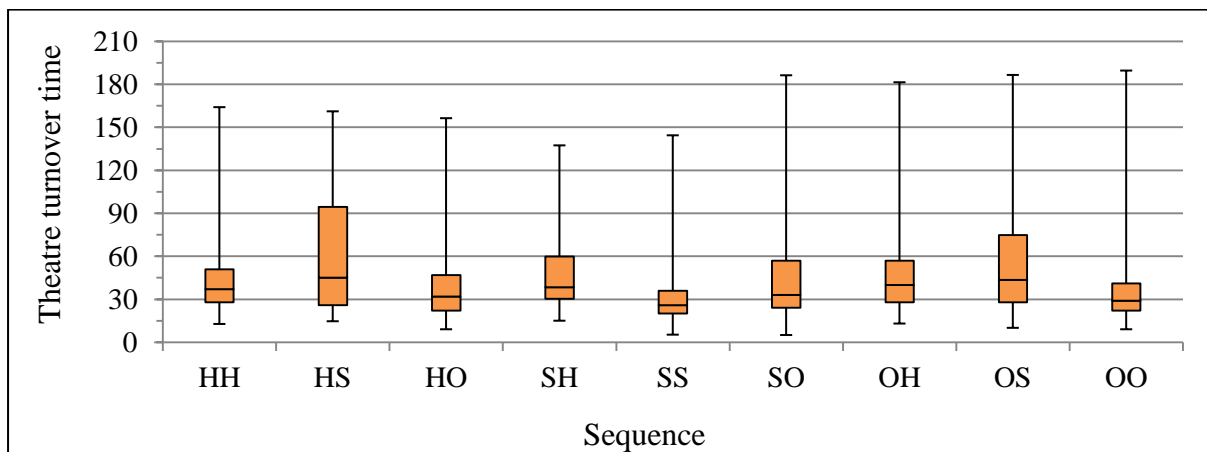| Sequence	n	Mean	S.D.	Mini-mum	Maxi-mum	Skew-ness	Kurtosis
H H	671	45.11	29.33	5	367	3.95	28.28
H S	38	60.00	42.38	12	177	1.02	0.04
H O	1046	40.28	30.95	4	330	3.85	24.41
S H	50	48.20	28.89	14	154	1.68	3.10
S S	109	33.78	25.48	3	167	3.01	11.27
S O	185	46.43	36.70	1	226	2.12	5.70
O H	1064	48.63	39.51	1	803	8.18	130.87
O S	122	55.91	38.27	7	204	1.53	2.69
O O	3183	38.28	35.51	2	501	5.19	38.95

Figure 7.6.2ii: Distribution of theatre turnover time (minutes)

7.7 Theatre efficiency

In Section 7.4.2 results regarding the utilisation of the trauma theatre were presented. However, utilisation alone does not necessarily accurately represent the efficiency of an operating theatre; one may achieve 100% utilisation but cancel several operations in order to prevent over-running of the schedule and thus cancellations should also be taken into account when considering theatre efficiency.

Let Fr_{schedt_util} be the fraction of theatre scheduled time utilised, Fr_{schedt_over} be the fraction of scheduled time over-running and Fr_{schedo_comp} be the fraction of scheduled operations completed. The following formula was devised to give a measure of theatre efficiency:

$$Fr_{schedo\_comp}\left(Fr_{schedt\_util} - Fr_{schedt\_over}\right)$$

which yields a value in the range [0, 1] (Pandit et al. 2007).

The only way a value exceeding 1 may be achieved is that if more patients are operated on than are originally scheduled. Strengths of this formula include that it incorporates more information about the running of the theatre than other approaches such as simply looking at utilisation or cancellation rates, where impressive results can be achieved by making large sacrifices in other areas.

This formula was applied to the trauma theatre data at the UHW and efficiency scores calculated as a percentage. It should be noted that a score exceeding 100% will not be possible here since the number scheduled per day is calculated by adding the number of cancellations to the number of theatre episodes completed. A total of 1333 days were available where all information required was valid and complete.

The mean efficiency score was calculated as 77.7% with a median of 80.3%. The maximum value was realised on two occasions, while a score exceeding 99% was achieved on 31 occasions (2.3% of total). The standard deviation was 15.6 percentage points and the minimum score was 16.8%. 187 (14.0% of total) of all lists scored an efficiency greater than 95% while 79 theatre lists scored an efficiency lower than 50% (5.9% of total).

By this measure, it appears that the theatre is performing relatively well although there is still some scope for improvement. However, the limitations of this formula must also be considered. One weakness is that it does not include any measure of efficiency *within* the theatre list (discussed previously in Section 7.6) and thus the period that the list is being used efficiently is overstated by an estimated 10-15% (Cook 2008). It also does not include any measure of tardiness relating to the start of the theatre list or additional booking of patients due to reduced theatre down time, so there is no reward for improving utilisation within the theatre in terms of booking and performing more operations in a timeslot of the same length (Sanders et al. 2008). Due to these drawbacks, while results are informative to a certain degree, this measure will not be used further.

End utilisation is calculated as the combination of anaesthesia and surgical time as a proportion of operating time (allocated theatre time less turnover time between operations) and has a nationwide target of 77% established by an Audit Commission survey (Audit Commission 2002). A year later a review of national results relating to operating theatres was published and huge variations in end utilisation between NHS Trusts were found, ranging from 41% to 103%, with an average of 73%. It was suggested that those Trusts with low end utilisation should increase their throughput and reduce capacity, while those exceeding 100% should consider increasing the capacity of their theatres. Orthopaedic surgery was shown to have the third best median end utilisation score, while the national interquartile range for end utilisation for trauma surgery was 58.9-86.4% (Audit Commission 2003).

The Wales Audit Office produced a report focussing in particular on NHS day surgery in Wales. Results for Cardiff and Vales NHS Trust showed an end utilisation rate of 64%, which compares favourably against the Welsh average of 57% (Wales Audit Office 2006). This compares with an average of 55% for day surgery in England (Healthcare Commission 2005).

Results for the trauma theatre at the UHW show an average end utilisation of 87.4%, suggesting that increased capacity may be required. Some considerable variation in results was also evident, with a standard deviation of 20.6 percentage points. 79.9% of sessions reached the Audit Commission target of 77% end utilisation.

7.8   Chapter summary

Various detailed analyses have been presented throughout this chapter and a number of concepts relating to the trauma theatre have been established. The TheatreMan database management system has been introduced and detailed results of the data obtained has been presented. Some of the issues with the data have also been discussed, including inaccurate readings and issues relating to fitting distributions of the created time intervals.

Two detailed data exploration exercises were undertaken regarding hip surgery; namely whether surgeon experience and timing of the operation have any impact upon length of operation and, in the latter case, acute outcome. It was found that surgeon experience did influence operation length in the majority of cases; consultants were found to be quicker than trainee surgeons. However, it was also decided that this would not be built into any simulation model since decisions of this kind at a strategic level are rather infeasible here; and there can be no consultant surgeons in the future without them being trainee surgeons first! Timing of the operation was found to have no effect on the length of operation or acute outcome, thus also would not be incorporated into a simulation model.

Results relating to theatre efficiency, utilisation and cancellations were also presented. It was found that the trauma theatre was performing relatively well in terms of utilisation, but that there was, on average, just over one operation cancelled per day due to running out of time. The theatre often started late and, despite these results being consistent with those reported in the literature, it has been identified as a problematic area.

Turnover times for the trauma theatre were presented. These were calculated for nine separate groups and for two different definitions: anaesthetic room / theatre turnover and theatre turnover. In some cases the theatre turnover times were found to be particularly large, identifying this as an area in which potential improvements could be made.

Finally, there was some discussion of the concept of theatre efficiency and targets and results for end utilisation were presented. With regard to this measure, the trauma theatre under study appeared to be performing relatively well but there was still room for considerable improvement.

CHAPTER 8: MODELLING THE TRAUMA THEATRE

8.1 Introduction

A thorough overview of the trauma theatre at the UHW was given in Chapter 7. This included summaries of various measures and outputs and some statistical analyses were also completed.

Using findings and conclusions drawn from this work, a discrete event simulation (DES) model of the trauma theatre was built. The formulation of the model, together with evidence of satisfying validation and verification procedures and thus producing a model which accurately represents the real life system, are hence explained. Key outcomes here include cancellations and theatre utilisation.

Cancellations can cause considerable distress for patients, particularly if, for example, they are unnecessarily starved in preparation for an operation which ultimately does not happen that day. There may also be inconvenience to hospital staff and/or a waste in resources.

An appropriate balance in theatre utilisation is also desirable. High utilisation leads to a pressured system and over-worked staff, and potentially paying costly overtime rates to staff. Low levels of utilisation indicate a waste in valuable resources and an inefficient system.

8.2   Model formulation

A pictorial representation of the trauma theatre model is given in Figure 8.2i, and explained forthwith. This model was built using VBA for Excel.

Late starts have been shown to be a particular affliction associated with the trauma theatre at the UHW, see Section 7.4.2, thus tardiness needs to be incorporated into the model. The theatre is scheduled to start at time A, but actually starts at time B. Note that an early start can occur (i.e. $B < A$), but this is not common.

At time B, the first patient enters the anaesthetic room to be anaesthetised. Referring to Section 7.2, this is equivalent to the *into anaesthetic room* time field. When this procedure is completed, they immediately enter theatre at time C (*into theatre*) and undergo surgery, starting at time D (*operation start*), so that between times C and D the patient does occupy the theatre, but their operation has not yet started. The operation takes places between times D and E. Theatre exit time occurs between E (*operation finish*) and F (*out of theatre*).

Between F and G is theatre turnover time, so that when the second patient enters theatre at G, they have already been anaesthetised. Their anaesthesia time can in fact extend to before F (but after C), as long as the anaesthetic room has been prepared.
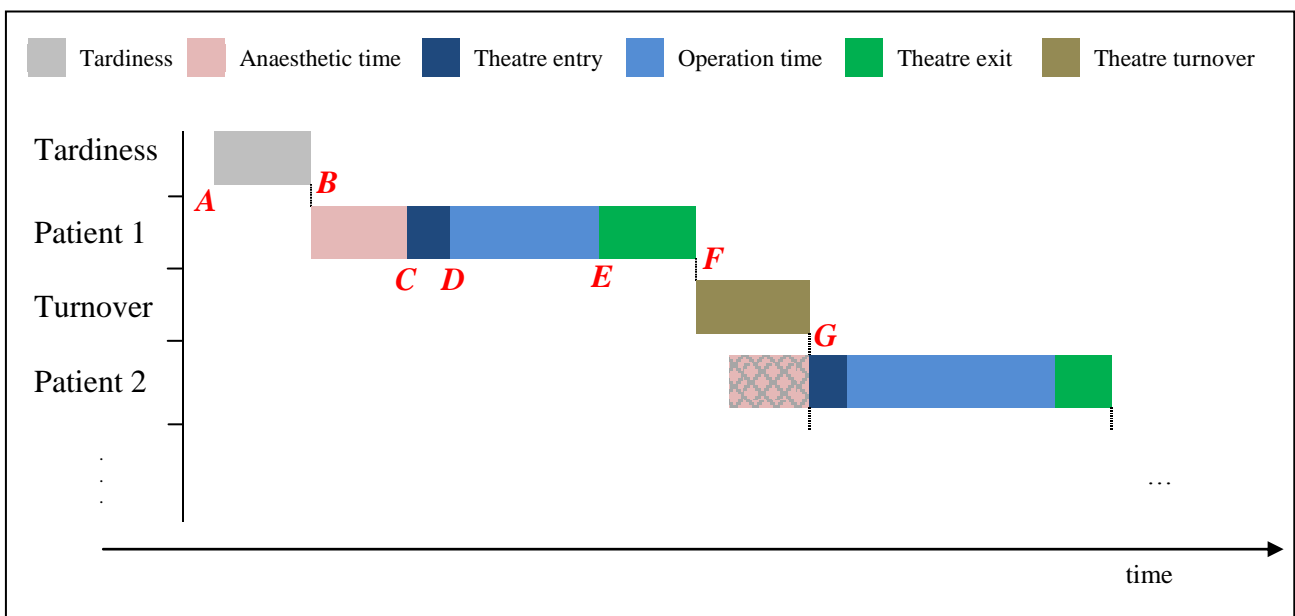

Figure 8.2i: Pictorial representation of trauma theatre model

Due to these complications, the theatre is modelled as above, where anaesthetic room/theatre turnover and anaesthetic time beyond the first patient is not considered. This is instead incorporated into theatre turnover (F to G). This is indeed what happens in reality; there would never be two patients simultaneously occupying the trauma theatre and theatre turnover incorporates anaesthetic time of the next patient. Including anaesthetic time of patients after the first patient would unnecessarily add calculations into the model without improving its outputs or appropriateness of representing the real world situation.

Using the inputs derived and defined in Sections 8.2.1–5, a theatre schedule was created which accurately models what currently happens in the trauma theatre. This covers all scheduled operations, not just those that actually take place, since cancelled operations are an important factor to consider. Once it was determined that an accurate schedule and means of cancelling operations had been derived, an appropriate run length and the number of runs was determined. Using the pre-determined schedule, a number of what-if scenarios were then applied to determine any changes which would be observed should the schedule be changed. By using the same schedule each time, direct and fair comparisons can be made.

8.2.1 Scheduled operations

The number of scheduled operations per day, calculated as the number performed plus the number cancelled due to lack of time, was found to follow day-dependent Binomial distributions, see Table 8.2.1i. The day-dependency was due to a higher number of longer, spinal operations being performed on Mondays and Thursdays, for example. On average Wednesdays tended to have a higher number of scheduled operations and so was also segregated.

First operation type was also found to be day-dependent, and as a result, so was the ordering of subsequent operations. Again this was due, for example, to a higher proportion of spinal operations being performed on Mondays and Thursdays, and alike operations being more likely to be scheduled in succession. The type of first and each following operation were determined using the data. For more information on ordering of operations, see Section 7.6.

Table 8.2.1i: Binomial fits for number of scheduled operations per day

Day	*n*	*p*	Theoretical		Empirical	
			Mean	S.D.	Mean	S.D.
Mon, Thurs	12	0.636	6.4	1.7	6.4	1.6
Tues, Fri	13	0.536	7.0	1.8	7.0	1.7
Wed	15	0.605	7.6	1.9	7.5	1.7
Weekend	13	0.596	6.5	1.8	6.5	1.9

8.2.2   Anaesthetic time

Anaesthetic time is only considered for the first patient. It was previously found (Section 7.4) that this time was dependent upon operation type but not the number of procedures performed. Further investigation showed that anaesthetic time could be modelled by the Gamma distribution for each of the three operation types; summary measures are given in Table 8.2.2i and are displayed graphically in Figure D8.2.2a of the Appendix.

Table 8.2.2i: Gamma fits for anaesthetic time (minutes)

Type	*α*	*β*	Min	Theoretical		Empirical	
				Mean	S.D.	Mean	S.D.
Hip	3.924	6.320	0	24.8	12.5	24.8	13.5
Spinal	3.038	7.083	0	21.5	12.3	21.5	14.9
Other	3.227	4.937	0	15.9	8.9	16.0	11.9

8.2.3   Operation time

Operation time was shown to be dependent upon main operation type and the number of procedures performed (Section 7.4). Further investigation showed that this time could be modelled by the Lognormal distribution in two cases and by the Gamma distribution in five cases, as displayed in Tables 8.2.3i and 8.2.3ii and Figures D8.2.3a and D8.2.3b. No fits are given for two operation types, hip with three or more procedures and spinal with three or more procedures, each only had eleven data points and therefore data was sampled from.

Table 8.2.3i: Lognormal distributions for operation time (minutes)

Type	No. of procs	μ	σ	Min	Theoretical		Empirical	
					Mean	S.D.	Mean	S.D.
Hip	2	4.509	0.541	1	106.2	61.3	103.4	55.1
Other	2	3.767	0.832	0	61.1	61.1	60.0	51.5

Table 8.2.3ii: Gamma distributions for operation time (minutes)

Type	No. of procs	α	β	Min	Theoretical		Empirical	
					Mean	S.D.	Mean	S.D.
Hip	1	4.086	15.279	0	62.4	30.9	62.4	31.7
Spinal	1	2.376	34.711	5	87.5	53.5	87.5	51.1
Spinal	2	2.988	39.668	5	123.5	68.6	123.6	64.6
Other	1	1.736	26.590	0	46.2	35.0	46.2	36.6
Other	3+	1.414	88.399	11	136.0	105.1	136.0	99.9

8.2.4 Tardiness and theatre turnover

Tardiness relating to start time was previously discussed in Section 7.4.2. In addition to weekends starting noticeably later than other days, it was also found that theatre tended to start more promptly on Thursdays. Tardiness was found to follow a Gamma distribution for each of the three day groupings of weekend, Thursday and other days. Estimators are given below where an excellent fit can be seen in each case.

Preceding and following operation types must be considered for theatre turnover time. The Lognormal distribution was found to accurately model turnover time for each of the nine sequence possibilities. Results are given in Table 8.2.4ii and Figure D8.2.4a.

Table 8.2.4i: Gamma fits for theatre start time tardiness (minutes)

Day	α	β	Min	Theoretical		Empirical	
				Mean	S.D.	Mean	S.D.
Mon-Wed, Fri	4.523	11.305	-32	19.1	24.0	19.5	28.2
Thurs	2.843	11.080	-20	11.5	18.7	11.1	21.7
Weekend	6.943	9.784	-21	46.9	25.8	46.9	26.5

Table 8.2.4ii: Lognormal distributions of theatre turnover time (minutes)

Sequence	μ	σ	Min	Theoretical		Empirical	
				Mean	S.D.	Mean	S.D.
H H	3.521	0.564	5	44.7	24.3	45.1	29.3
H S	3.518	0.897	12	62.4	56.0	60.0	42.4
H O	3.361	0.656	4	39.7	26.2	40.3	31.0
S H	3.242	0.827	14	50.0	35.7	48.2	28.9
S S	3.221	0.638	3	33.7	21.8	33.8	25.5
S O	3.563	0.722	1	46.8	37.9	46.4	36.7
O H	3.688	0.547	1	47.4	27.4	48.6	39.5
O S	3.604	0.814	7	58.2	49.6	55.9	38.3
O O	3.359	0.628	2	37.0	24.4	38.3	35.5

8.2.5   Theatre entry and exit

Theatre entry time is defined as the time between the patient entering the theatre and the operation starting. Theatre exit time is defined as the time between the operation finishing and the patient leaving theatre. Theatre entry and exit was a little trickier to model as no distribution function could be found to accurately fit the data, so instead data was sampled from. A summary is given by operation type in Table 8.2.5i.

Table 8.2.5i: Summary of theatre entry and exit times (minutes)

Operation type	Theatre entry		Theatre exit	
	Mean	S.D.	Mean	S.D.
Hip	13.7	7.0	7.5	7.1
Spinal	8.7	6.0	7.3	7.6
Other	13.1	8.0	11.8	7.8

8.2.6   Cancellations

Lack of time cancellations were modelled by comparing how often an overrun of the theatre was allowed with how often an overrun was required. Other types of cancellation are not

considered. Theatre use beyond the daily allocation of 11.5 hours is classified as an overrun. Days where theatre usage was less than 11.5 hours and there were no lack of time cancellations were classified as no overrun required, all other days were classified as requiring an overrun.

The data showed that an overrun was required 68% of the time, and occurred 45% of the time, so that in 67% of cases where an overrun was required, it was permitted. This was included in the simulation by taking a random value v from the interval [0, 1] once total theatre time exceeded 11.5 hours and there were still scheduled operations remaining. If $v \leq r_l$, the overrun limit initially set to 0.67, then the overrun was allowed and the next operation would go ahead. If $v > r_l$, an overrun would not be permitted and all outstanding operations would be cancelled. This was repeated for each operation if there was more than one outstanding operation in the schedule at 11.5 hours, but the limit of 0.67 was scaled by a factor of four for each subsequent operation after the first one allowed after 11.5 hours; that is, it was changed to 0.67^4 (= 0.20) for the second operation, 0.67^8 (= 0.04) for the third operation, and so on. It was found that this gave values approximately in line with the data using the logic explained previously.

8.2.7 Initialisation bias, run length and number of replications

The model set-up and validation process need not be quite as rigorous as was completed in Chapter 6. In this case there is a terminating system and so no warm-up period is required.

Run length needs to be long enough to be confident that sufficient time has been covered for the model to accurately represent the trauma theatre. A run length of 2000 days was used. It was found that this gave an accurate representation of the system while not compromising on runtime.

Methodology to determine the number of replications to perform was described in Section 6.3.11. Cumulative mean results with 95% confidence intervals for two key outputs of this simulation model, the number of daily lack of time cancellations and total theatre busy time, are now presented. These graphs were constructed up to 1000 replications but results are excluded for display purposes. Additionally graphs were inspected for other measures and the same pattern seen. Precision results, where p_r gives the precision at replication r, are

given for four key measures in Table 8.2.7iii and displayed in Figure D8.2.7a. Precision is obtained within 1% for each of these measures by using 100 replications. Increasing the number of replications beyond this does not yield a notably higher level of precision, so *r* was set to 100 for all subsequent runs of the model.
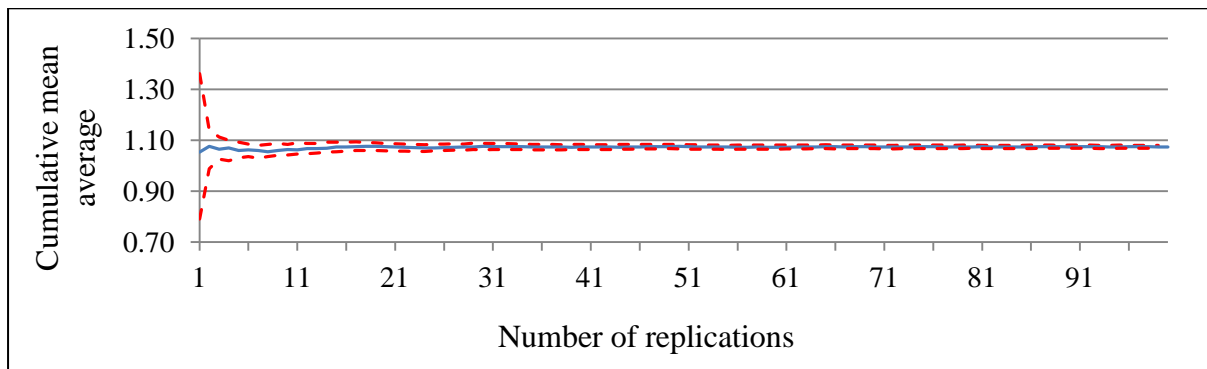

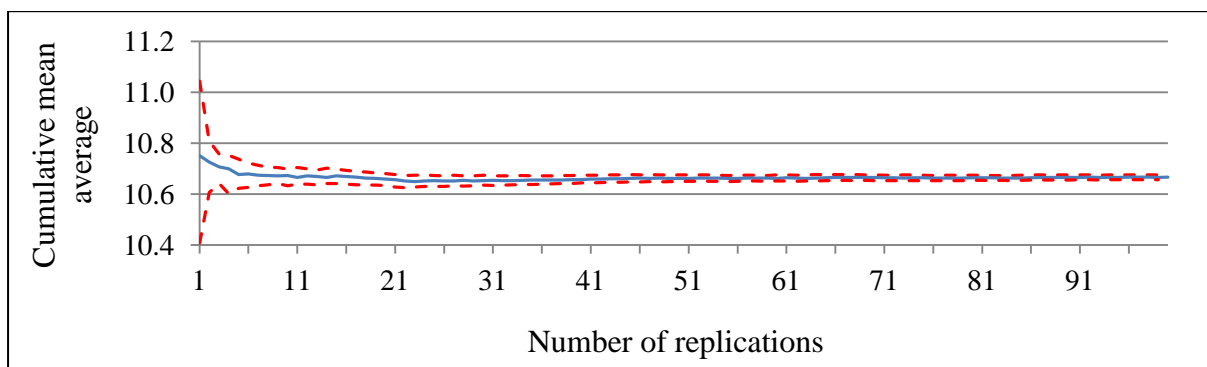**Figure 8.2.7i:** Results of average number of cancelled operations with respect to the number of replications to perform

Figure 8.2.7ii: Results of average theatre busy time (hours) with respect to the number of replications to perform

Table 8.2.7iii: Precision values obtained for various measures at different values of *r*

Measure (mean of)	Precision value, p_r (%)				
	r = 10	*r* = 50	*r* = 100	*r* = 150	*r* = 200
Cancellations	2.21	0.80	0.56	0.48	0.43
Busy time	0.29	0.12	0.09	0.07	0.06
Percentage hip operations	1.38	0.46	0.32	0.26	0.22
Percentage hip cancellations	2.99	1.15	0.75	0.60	0.51

8.3   Validation and verification

Validation was completed via the comparison of various model outputs and the data. It is important to look at these measures in detail in this way since the data was broken down in several different ways in order to formulate the model, which could have led to inaccuracies in overall numbers/proportions.

A multitude of outputs were looked at and all were found to be accurate when compared with the data. Some examples of these are now presented; a comparison of the number of scheduled and performed operations per day in Figure 8.3i and tardiness in Figure 8.3ii. An overall summary is given in Table 8.3iii. Two values are italicised to indicate that they were estimated using the data; see Section 7.5, where it was also noted that these values are likely to be underestimates of the true value.
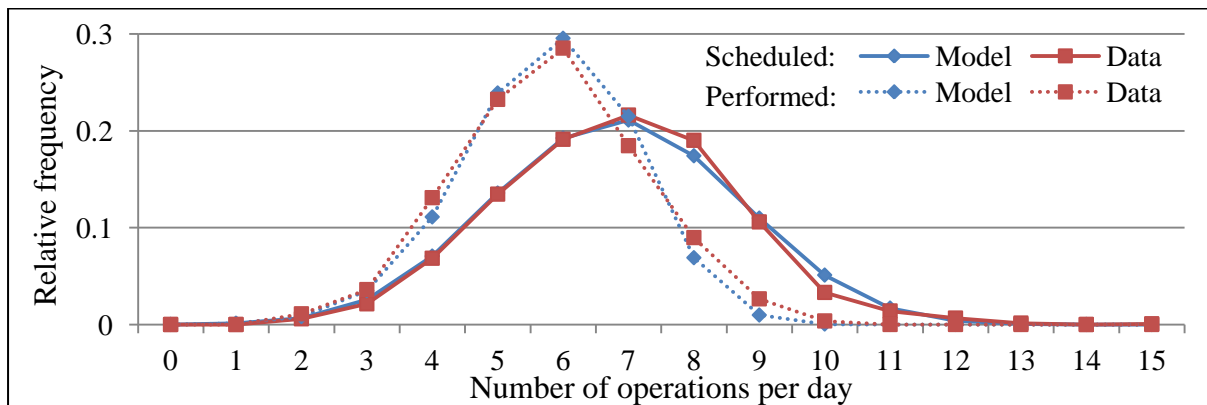

Figure 8.3i: Comparison of number of scheduled and performed operations per day.
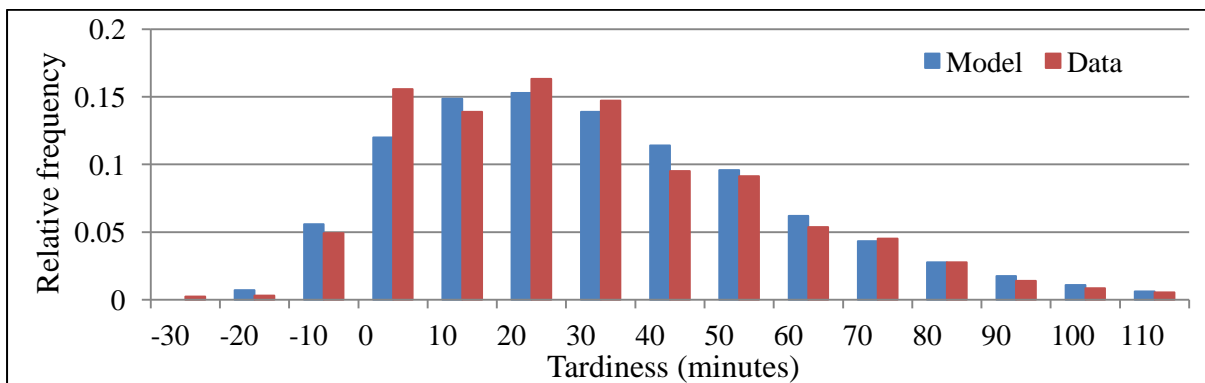

Figure 8.3ii: Comparison of tardiness

Table 8.3iii: Summary of measures, model versus data (*estimated*)

Measure (mean of)	Model	Data
Number of scheduled operations per day	6.83	6.84
Number of performed operations per day	5.76	5.81
Number of cancelled operations per day	1.07	1.02
Theatre busy time (hours)	10.7	10.8
Theatre utilisation	92.8%	92.8%
Percentage of performed operation types: Hips Spinal Other	24.2% 5.1% 70.6%	26.5% 4.9% 68.6%
Percentage of cancellations that are hips	29.3%	*25.3%*
Percentage of scheduled hips that are cancelled	18.5%	*14.2%*

Theatre usage is being represented to a very high level of accuracy, with theatre busy time being within 0.1 of an hour. The number of cancelled operations per day is slightly overstated, but only by approximately 0.05 per day, which does not give much cause for concern. The model estimates that 15.7% of all operations are cancelled due to lack of time, compared with 14.9% from the data.

Overall these results show that the model accurately represents the trauma theatre. Note that comparisons made later when the system is amended via what-if scenarios relate to these baseline modelled values and not empirical values.

The process of verification was completed as per Section 6.4.

Once the model was sufficiently validated and verified, the simulation was amended to model a number of what-if scenarios to investigate possible changes to the running of the trauma theatre.

8.4   What-if A: Change to start time tardiness

The data showed that the average tardiness relating to start time was 26 minutes. If this tardiness could be reduced, or even eradicated, then a more efficient system would be seen. This change was implemented into the model by taking proportions of the original modelled value for tardiness; both positive and negative changes in efficiency are considered. Taking a proportion of 0% thus means that the theatre starts on time each day.

The change to the total number of cancellations and theatre hours per year is given in Table 8.4i. By starting punctually each day, approximately 50 cancellations and 86 theatre hours can be saved per year. Even just by reducing tardiness by 25% (75% of original), approximately 14 cancellations and 22 theatre hours can be saved per year. This is a fairly considerable saving bearing in mind that this translates to just eliminating just 6.5 minutes off lateness on average. Specific results for the impact upon hip cancellations are shown in Figure 8.4ii. Hip cancellation rate refers to the percentage of scheduled hip operations that are cancelled (and not the proportion of cancellations that are hip patients).

Table 8.4i: Effect of a change in tardiness on yearly cancellations and theatre time (hours)

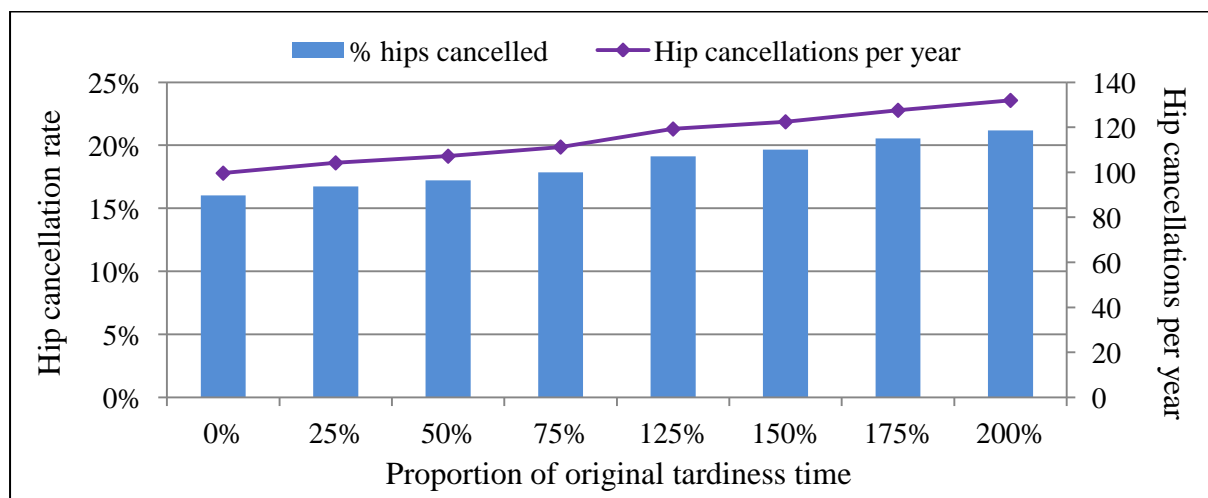| Change to	Proportion of original tardiness							
	0%	25%	50%	75%	125%	150%	175%	200%
Cancellations	-49.8	-37.1	-25.6	-14.0	13.8	26.2	41.3	56.6
Theatre time	-86.1	-64.6	-43.8	-22.5	18.5	35.4	50.7	66.8

Figure 8.4ii: Effect of a change in tardiness on hip cancellations

8.5 What-if B: Change to turnover time

While theatre turnover time can never be completely reduced in reality, the impact of an improvement in efficiency relating to this time is investigated here. This was investigated by both changing the turnover time by a certain percentage, and setting turnover time to a fixed value. Turnover times are also considered indirectly in Section 8.6.2, where operations of the same type are scheduled consecutively.

8.5.1 Percentage change in turnover time

For this first case, the original modelled value was increased or decreased by a certain percentage and this value then taken as the new turnover time. Results are displayed in Figure 8.5.1i.
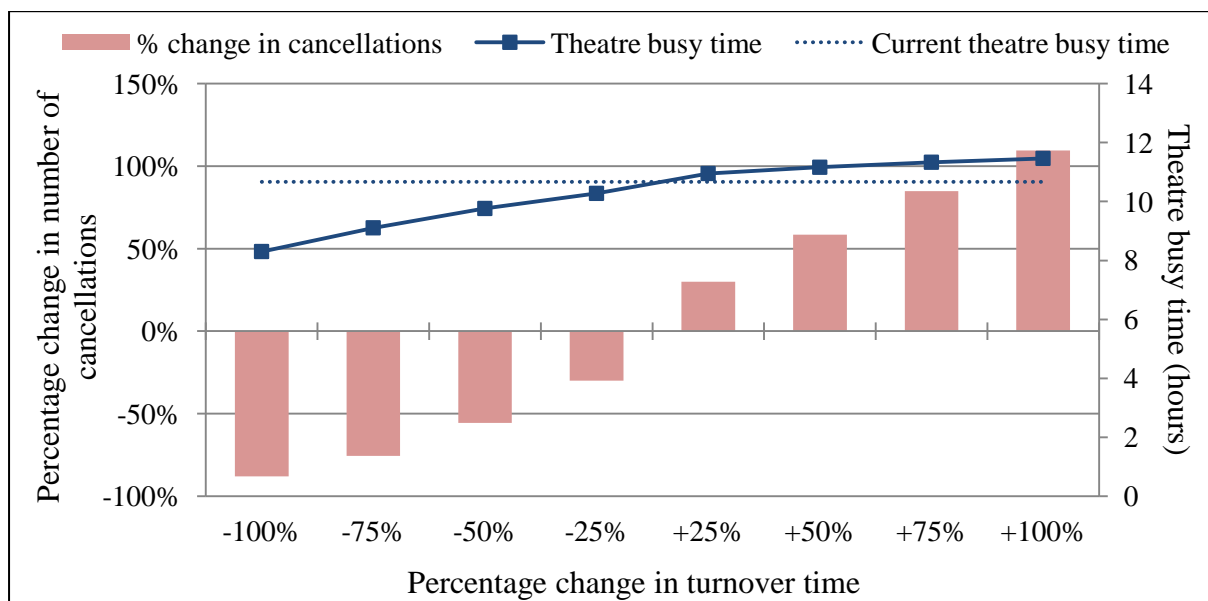

Figure 8.5.1i: Impact of a percentage change in turnover time

While a reduction of 100% (that is, all turnover times reducing to zero) is unlikely, results are included for completeness. Note that doing this does not mean that cancellations are completely avoided, but they are reduced by 88% to approximately 0.1 per day on average, while doubling the turnover time would more than double the average number of daily cancellations, to approximately 2.2 per day. The best case scenario sees the percentage of

hips that are cancelled reduce to 2%, while the worst case scenario sees 39% of all hip operations being cancelled.

The average theatre turnover time was 42 minutes, so a 25% reduction equates to trimming approximately ten minutes on average from the turnover time. This would save 117 cancellations and 142 hours of theatre time per year. The percentage of all scheduled operations that are cancelled would decrease to 11.0% from 15.7%, and to 12.8% from 18.5% for scheduled hip operations. In this case theatre utilisation would reduce to 89% from 93% and would therefore be within the guidelines of The Bevan Report (see Section 7.4.2).

A 25% increase would lead to 118 additional annual cancellations and an extra 103 hours of theatre time in total per year. The percentage of all scheduled operations that are cancelled increases to 20.4% and to 24.0% for hip operations, an extra 34 hip cancellations per year.

8.5.2   Fixed turnover time

Two cases are considered here: A, the same fixed value for all turnover times, and B, fixed turnover time values but with some consideration of operation type. In general, turnover times between operations of the same type are shorter, so for case B, if the turnover time was set to, say, x minutes, then turnover time between operations of the same type was set to $0.5x$ minutes. Zero turnover is again included for completeness.

The impact on the trauma theatre if either of these scenarios could be achieved is evident. In particular consider the greater impact if case B was achieved, see Figure 8.5.2i and Table 8.5.2ii (* time halved for turnover between operations of the same type, scenario B). Table 8.5.2ii gives the value change in both the number of operations performed per day and the number of theatre busy hours per day; for example, a fixed turnover time of zero minutes means that the average extra operations that can be performed per day is 0.95, with a reduction in the theatre usage of 2.36 hours per day.

For scenario B, the fixed turnover time needs to be increased to 60 minutes (30 minutes between same type operations) before the current overall cancellation rates and theatre usage are seen.
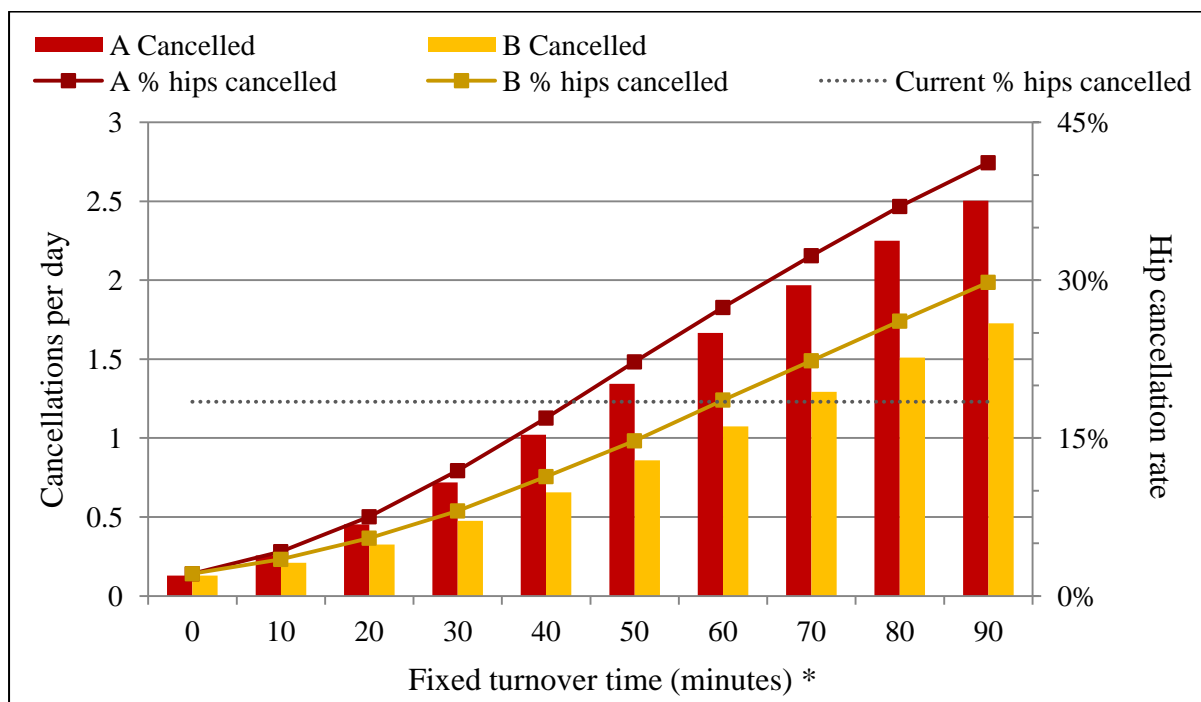
Figure 8.5.2i: Impact of changing turnover to a fixed value for two scenarios

Table 8.5.2ii: Impact of a fixed turnover time (minutes) on number of operations performed per day and theatre usage for two scenarios

Fixed turnover *	Number of operations		Theatre busy time (hours)	
	Scenario A	**Scenario B**	**Scenario A**	**Scenario B**
0	0.95	0.95	-2.36	-2.36
10	0.82	0.86	-1.57	-1.82
20	0.62	0.75	-0.90	-1.31
30	0.36	0.60	-0.36	-0.89
40	0.05	0.42	0.05	-0.52
50	-0.27	0.21	0.36	-0.21
60	-0.59	0.00	0.60	0.05
70	-0.90	-0.22	0.78	0.26
80	-1.18	-0.44	0.93	0.43
90	-1.43	-0.65	1.06	0.58

8.6 What-if C: Re-ordering/re-allocation of operations

In this section, some changes are made to the theatre schedule. The same overall schedule is still used so that fair comparisons can be made, but the ordering and allocation of operations is amended to assess three different approaches.

8.6.1 Priority to hip patients

Anecdotal evidence suggests that hip patients are often moved to the end of the theatre list. This what-if scenario tests the effect of moving all hip patients to the start of the schedule in order to minimise the likelihood that they are cancelled.

Firstly, consider the scenario where priority is given to the first hip patient on the schedule and this patient moves to the start of the list. Days without any scheduled hip operations thus remain unaffected and any hip patients further down the list are also not moved. This is defined as policy A.

Secondly, for policy B, all hip patients are given priority. Any scheduled hip patients move to the start of the theatre list in the order that they were originally scheduled, followed by any remaining scheduled patients, also in their original order.

Results collated from making these changes to the simulation model showed little difference to overall cancellation rates. There was a minor decrease for both policies, which was greater for policy B, but these differed little from the current situation; however, a change was seen in the proportion of hip patients which made up the cancellations, see Figure 8.6.1i. There was also very little impact to theatre busy time.

29.3% of all cancellations are hip patients in the current situation, which decreases to 24.8% for policy A and 5.2% for policy B. Of all scheduled hip operations, currently 18.5% are cancelled, which would reduce to 15.5% and 3.2% for policies A and B respectively.

Hip cancellations have thus not been completely eradicated by either of these policies, since in some cases hips would dominate the schedule or particularly lengthy operations may occur. However, results show that policy B in particular would go some way to improving the amount of hip operations that are cancelled.
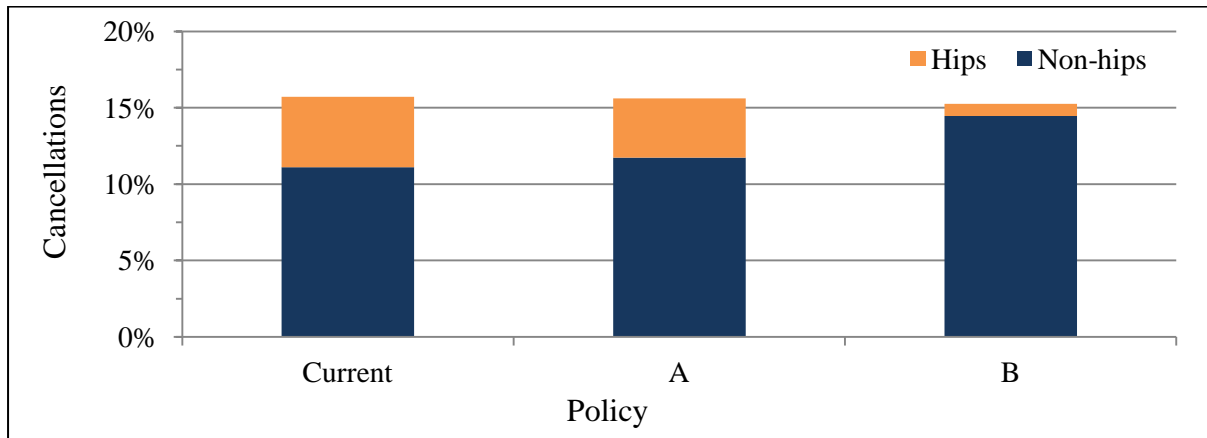
Figure 8.6.1i: Cancellation results for two policies of prioritising hip patients

8.6.2 Alike operations in succession

Although the data suggests that there may be some effort made to schedule alike operations consecutively, this does not always happen. Longer turnover times tend to occur when unalike operations follow each other, thus the result of simply scheduling alike operations together should result in a shorter theatre busy time. Since there are three operation types, there are $3! = 6$ combinations of ordering which may occur.

Some cancellation results are given in Figure 8.6.2i, where *HSO* refers to the ordering of hip operations, followed by spinal operations, followed by other operations, for example. The lowest cancellation rate of 14.2% was given by the ordering *OHS*, with the highest of 16.0% given by *HSO*. The data shows that the longest average turnover between two operations is seen when a spinal operation follows a hip operation, which in some part explains this result. In addition to *HSO*, *SHO* also gave a marginally worse cancellation rate than what is currently seen, with 15.8% compared with 15.7% currently; all other orderings gave an improvement.

As expected, fewer hip cancellations are seen when they are prioritised, and this result is also observed when they come second to spinal operations. For the two cases where hips come first, just 0.7% of all scheduled hip operations are cancelled. This increases slightly to 1.5% for the ordering *SHO*. For *SOH*, *OHS* and *OSH* respectively, the percentages of hip operations that are cancelled are 39.9%, 35.0% and 40.3%, which account for approximately 10% of all scheduled operations in each case.
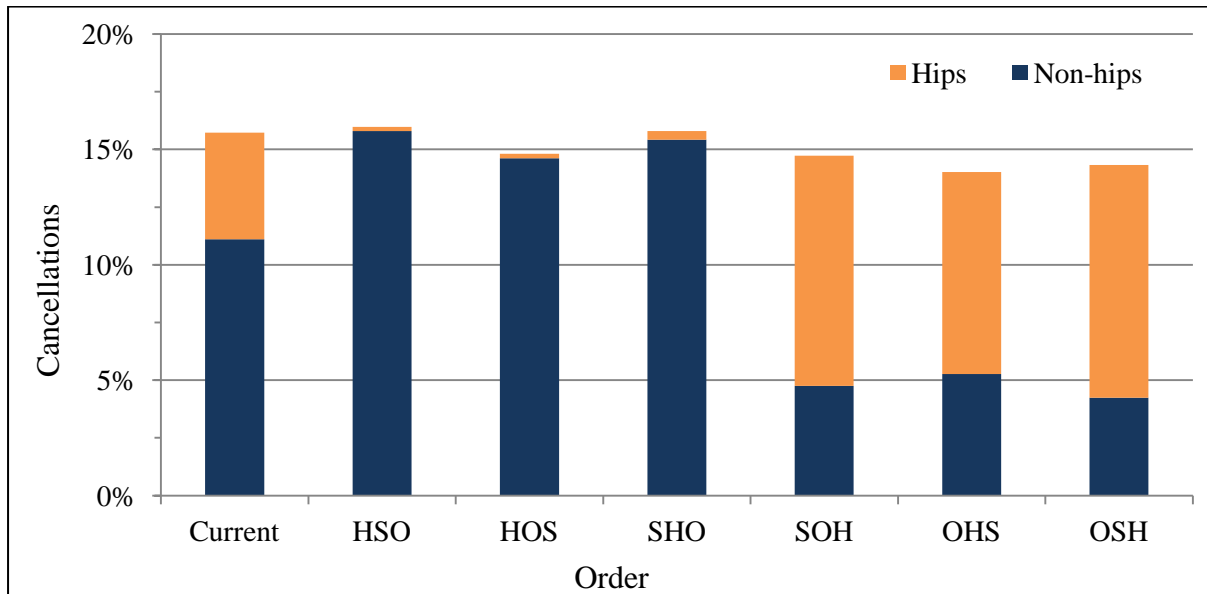
Figure 8.6.2i: Cancellation results for scheduling alike operations together

Total theatre time varies very little across these orderings. Therefore in the cases where fewer cancellations are achieved, a more efficient theatre system is evident; the theatre is still used for the same amount of time but more operations are performed.

The best ordering of *OHS*, in terms of cancellations saved, would result in 42 fewer cancellations per year; however this is an increase of 103 hip cancellations and a decrease of 145 non-hip cancellations, so hip patients would bear the burden of this improvement. The orderings which prioritise hip patients would almost eradicate hip cancellations; a decrease of approximately 96% (110 cancellations) is seen. However, non-hip operations would be affected by this; for *HSO*, non-hip cancellations would increase by 116 per year and for *HOS* by 87 per year. The worst outcome for hip patients would be *OSH*, where the number of hip cancellations per year would more than double, although non-hip cancellations would reduce by 62% as a result.

Choosing which of these options, if any, to implement would ultimately depend upon staffing, resources and preferences of hospital staff; a stakeholder of hip fracture care may have differing opinions to a spinal surgeon, for example.

8.6.3 Removal of spinal operations

There has been some discussion amongst staff at the UHW of moving spinal operations to another theatre. These operations can be particularly time-consuming and re-allocating them to another theatre will mean a reduction in the demand on the trauma theatre.

By removing all spinal operations, the average theatre usage per day would reduce slightly from 10.7 hours to 10.3 hours, just 20 minutes on average per day, with theatre utilisation dropping to 89.9%. The number of lack of time cancellations would reduce by approximately 37 per year for all operation types, with hip lack of time cancellations reducing by 14 per year.

It can be seen that removing spinal operations from the trauma theatre has led to a fairly modest change in outcomes, but it must be remembered that these operations account for only approximately 5% of all operations. One useful result is that it can be seen that should this approach be implemented, average theatre utilisation would just about fall to below the recommended threshold of 90%.

8.7 What-if D: Change to theatre allocation

The theatre is currently allocated to be used for 11.5 hours (690 minutes) per day. At this point in the model, if there were still outstanding operations, a decision was made to determine whether or not an overrun would be required using the cancellation strategy described in Section 8.2.6. Changes to both the theatre allocation and the cancellation strategy are now investigated.

8.7.1 Change to allocated hours

Firstly, a change in the number of scheduled theatre hours per day is made and results relating to cancellations and theatre busy time are presented in Table 8.7.1i. Both increases and decreases to daily scheduled theatre hours are considered and the overrun limit r_l is kept at 0.67.

Table 8.7.1i: Effect of changing trauma theatre allocated hours

Allocated hours	Percentage of cancelled operations		Change in theatre hours per day
	All	**Hips**	
10	23.6%	27.6%	-0.87
10.5	20.8%	24.4%	-0.56
11	18.1%	21.4%	-0.27
11.5 (current)	15.7%	18.5%	0.00
12	13.6%	16.0%	0.23
12.5	11.5%	13.7%	0.45
13	9.7%	11.6%	0.65
13.5	8.2%	9.6%	0.82
14	6.8%	8.0%	0.97
14.5	5.6%	6.7%	1.11

The model shows that even if the allocated time is increased by 180 minutes to 14.5 hours per day, the change in the number of used hours per day will only increase by approximately 67 minutes. A change of this magnitude will therefore lead to a waste in expensive resources. However, just one additional hour per day, resulting in an extra 27 minutes of used theatre

time, would lead to 54 fewer cancellations per year, 15 of which are hip patients. A decrease in theatre allocation is not recommended; a reduction of just 30 minutes would see an extra 59 cancellations annually, 18 of which are hips, just for a reduction in theatre busy time of 16 minutes. Additionally, theatre utilisation would increase to 94.5%.

8.7.2 Change to cancellation strategy

Consider first the scenario where overruns are not allowed, so that $r_l = 0$. In this case, if the next operation (and the turnover time preceding it) will mean that the theatre overruns, that operation and any other outstanding operations are cancelled. The impact of doing this on cancellations and theatre usage is shown in Figure 8.7.2i and Table 8.7.2ii respectively, for a variety of theatre allocations.

Current cancellation rates are seen at approximately an allocation of 12.5 hours, an increase of one hour on the current allocation. Therefore if overruns are prohibited, the theatre allocation would need to be increased by an extra hour per day just to achieve current cancellation rates.
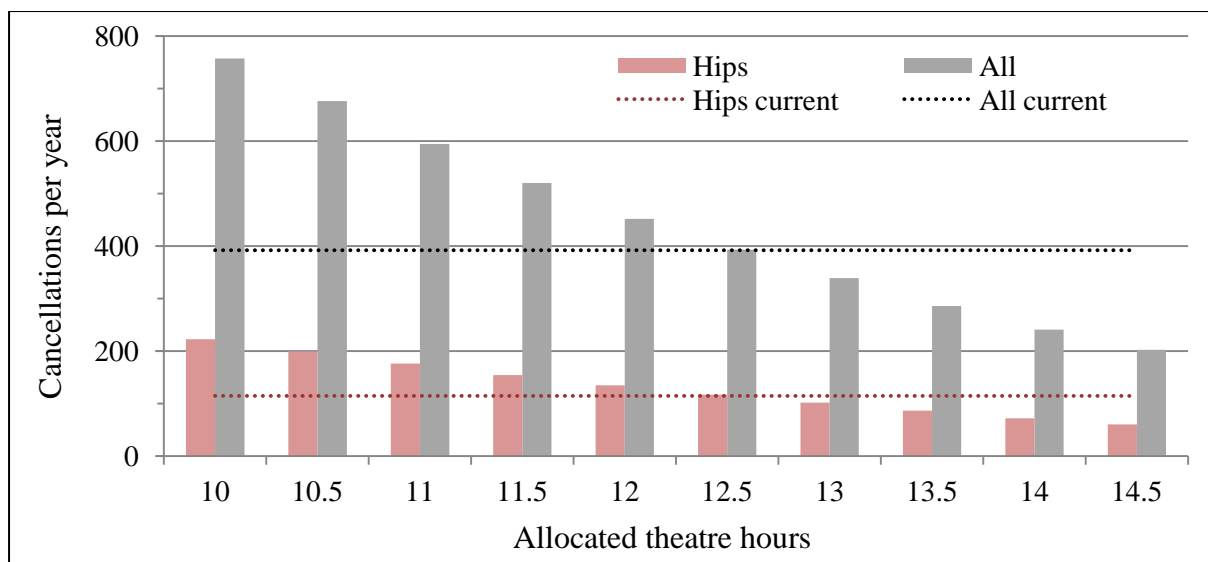

Figure 8.7.2i: Impact of not allowing overruns on cancellations

Table 8.7.2ii: Impact of not allowing overruns on theatre usage

Allocated hours	Change in theatre hours per day	Theatre utilisation
10.0	-1.72	89%
10.5	-1.35	89%
11.0	-1.00	88%
11.5	-0.68	87%
12.0	-0.37	86%
12.5	-0.08	85%
13.0	0.16	83%
13.5	0.40	82%
14.0	0.60	80%
14.5	0.79	79%

An increase to 14.5 hours sees the current cancellation rates approximately halve, but this would mean that the theatre is utilised just 79% of the time, which would be a significant waste of resources. Using any of the allocations considered here but with a policy of not allowing overruns would lead to a reduction in utilisation.

Setting $r_l = 1$ gives, in effect, an opposite policy of dealing with cancellations. Instead, all overruns are permitted whenever they are required and no cancellations are allowed. This is also the same as allowing a theatre usage of 24 hours. If this strategy was implemented, average theatre usage would increase by 17% to 12.4 hours, a utilisation of 52% of 24 hours, or 108% of 11.5 hours.

The value of r_l is now varied, but theatre allocation kept at 11.5 hours. $r_l < 0.67$ represents a more stringent overrun policy on what is currently used; overruns are allowed less often and therefore cancellations become more frequent. $r_l > 0.67$ thus represents a more lenient system whereby overruns are allowed more often, resulting in fewer cancellations but a more busy theatre. Results are displayed in Figure 8.7.2iii.
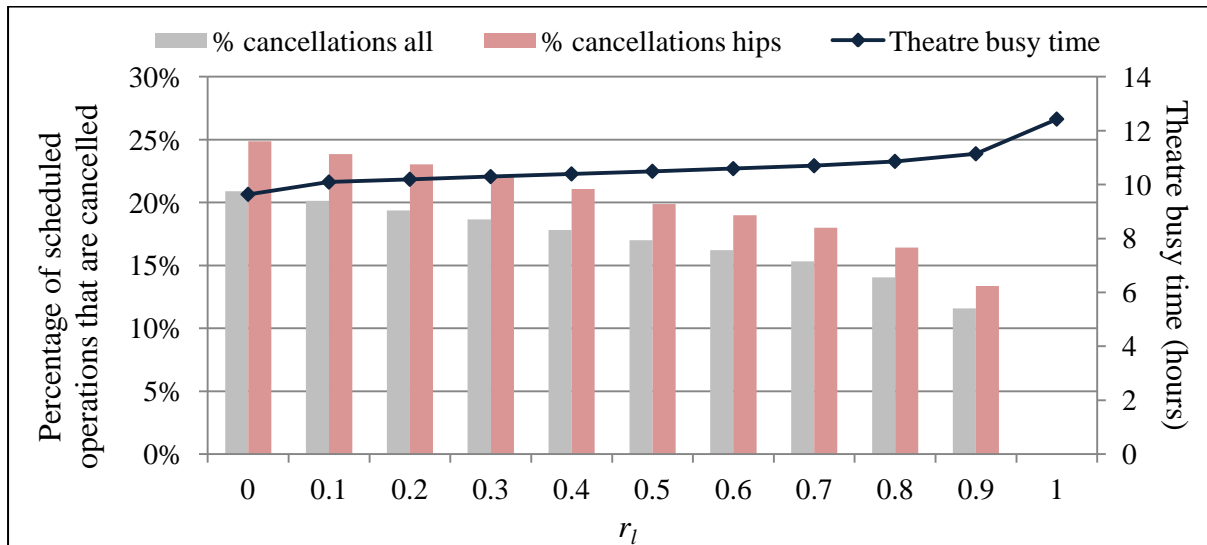
Figure 8.7.2iii: Impact of a change in r_l on cancellations and theatre usage

There is clearly a trade-off to be gained here; while increasing r_l results in fewer cancellations, the theatre becomes more likely to overrun and utilisation increases. For example, at $r_l = 0.9$, while just 12% of scheduled operations are cancelled, the theatre is operating at 97% utilisation and 57% of the time will exceed the limit of 11.5 hours. Average theatre utilisation is within 90% when $r_l < 0.4$.

The relationship between r_l and theatre allocation is now investigated. Nine scenarios are considered in total; each combination of a theatre allocation of 10.5, 11.5 and 12.5 hours and cancellation limits of 0.25, 0.5 and 0.75. Current allocation and an increase and decrease of one hour are thus both considered.

Firstly, the percentage change in the number of cancellations is presented in Figure 8.7.2iv, where green blocks show a decrease and red blocks show an increase from the current amount.

If theatre allocation was kept at 11.5 hours but overruns allowed only half of the time they were required, then there would be an 8% increase in the number of cancellations, a total of 32 extra affected patients per year. Setting $r_l = 0.5$ alongside a reduction in theatre time to 10.5 hours would see a much greater increase in cancellations of 42%, affecting an extra 163 patients per year. The most 'generous' scenario considered here, an allocation of 12.5 hours

263

and r_l = 0.75, where overruns are permitted 75% of the time that they are required, gives a 31% reduction in cancellations, the equivalent of 123 fewer cancelled patients per year.
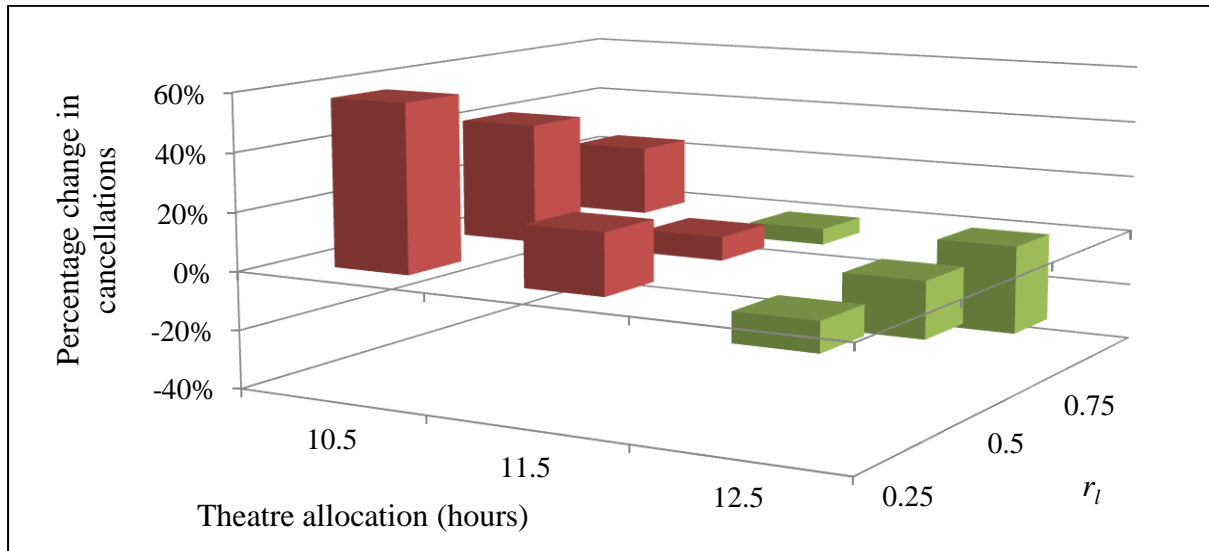

Figure 8.7.2iv: Impact on cancellations of a change to theatre allocation and r_l

A change in theatre usage will be another consequence of making these changes and some results are given in Table 8.7.2v. Cells contain average theatre usage in hours and average theatre utilisation for each combination of scheduled hours and r_l.

Table 8.7.2v: Impact of changing r_l and scheduled theatre hours on theatre usage

r_l	Theatre allocation (hours)		
	10.5	**11.5**	**12.5**
0.25	9.6	10.2	10.8
	92%	89%	86%
0.50	9.9	10.5	11.0
	94%	91%	88%
0.75	10.2	10.8	11.2
	98%	94%	90%

It is suggested that an allocation of 10.5 hours would not be advisable, since not only would there be a considerable increase in cancellations, utilisation would be above 90% even if only 25% of required overruns were allowed.

Utilisation is at most 90% when theatre allocation is 12.5 hours for each of the three cancellation limits simulated. Increasing r_l brings more uncertainty to the system, since overruns would be allowed more often and thus last-minute changes to staffing arrangements may need to be made, for example. Therefore it would be up to hospital managers if they wanted a more stable system where overruns are allowed less often, or one which has greater instability but fewer cancellations.

8.8 What-if E: Increase in arrivals

A change in the demand on the theatre is now investigated via changing the arrival rate. Changes to the *p* parameter of the Binomial distributions used to decide the number of operations scheduled per day were made in order to increase and decrease the overall demand by various percentages. Cancellation and utilisation results are presented in Figure 8.8i.
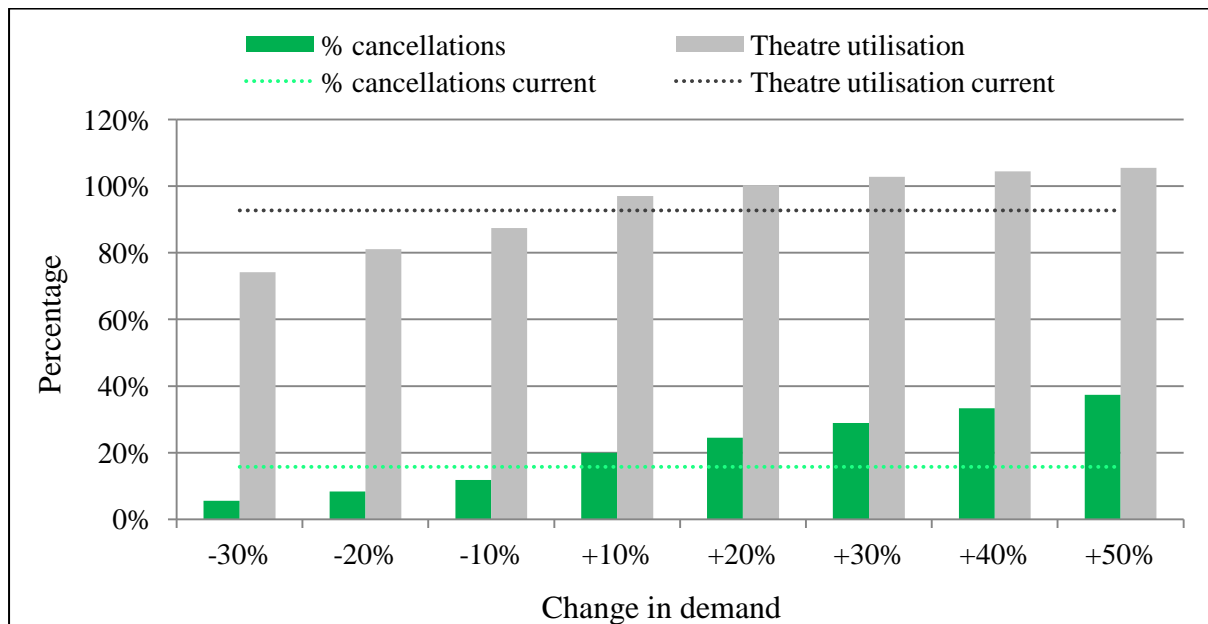

Figure 8.8i: Impact of a change in demand on cancellations and theatre utilisation

If the number of arrivals increases by 20%, average theatre utilisation will reach 100%. The percentage of all scheduled operations which are cancelled increases from 15.7% to 24.5%, which translates to approximately doubling the average daily number of cancellations (1.07 to 2.00). An increase in demand of 50% does not see a huge change to theatre utilisation due to the allocation and overrun strategy, but the percentage of cancelled operations increases to 37.4%, and 41.2% for hip operations. This translates to an average of 3.8 cancellations per day, 1.1 of which are hip operations. This must be considered should the demand on the trauma theatre increase in the future by, for example, increasing the allocated hours per day or implementing a change to the overrun strategy.

A decrease in demand is also considered. While the drop in the number of cancellations would be welcomed, the trade-off of the drop in utilisation must also be considered. For

example, the theatre would be empty 26% of the time should the demand fall by 30%, resulting in an expensive waste of resources. A 10% decrease in demand would see the utilisation fall to below the 90% guideline.

The most logical way to accommodate a change in demand is to change the open hours of the theatre. Consider the case where a change in trauma theatre demand is met with an equivalent change in theatre allocation, so that for example, if the demand increases by 10%, the scheduled open hours of the theatre is also increased by 10%. Changes to cancellations and theatre utilisation as a result of this are displayed in Table 8.8ii.

Recall that currently 15.7% of all operations and 18.5% of hip operations are cancelled. It is interesting to see that a reduction in demand and theatre time would lead to an increase in the proportion of scheduled operations that are cancelled. This is complemented by a reduction in theatre utilisation. The actual number of cancelled operations, however, would still decrease.

If the demand was to increase and the open hours changed accordingly, a reduction in the proportion of cancelled operations would be seen, but there would be an increase in theatre utilisation. Despite the reduction in the cancellation percentage, the number of cancelled operations would in fact increase in all cases except for a 10% increase in demand/allocation. This would see a very marginal decrease of 2% in the actual number of cancellations per year.

Table 8.8ii: Impact of a corresponding change in theatre demand and allocation

Percentage change in demand/allocation	Cancellations percentage		Theatre utilisation
	All	Hips	
-30%	20.2%	24.9%	91.6%
-20%	18.4%	22.3%	91.6%
-10%	17.0%	20.3%	92.0%
+10%	14.6%	16.9%	93.4%
+20%	13.6%	15.7%	94.0%
+30%	12.7%	14.5%	94.6%
+40%	11.9%	13.5%	95.2%
+50%	11.1%	12.4%	95.8%

8.9   A system designed around hip patients

Finally, an idealised trauma theatre system is simulated. Since hip patients are the primary focus of this research, this system is designed primarily to accommodate these patients. Some results gained from the previous sections are used in order to decide various parameters within the system, so that hip patients receive the best possible treatment while still working within the boundaries of a feasible and realistic scenario.

The first decision to be made is the ordering of patients. Clearly a system focussed on hip patients will always give priority to these patients, so all hip patients scheduled on any day will move to the start of the list. Ordering alike operations together was shown to usually decrease overall cancellation rates, with the order hip-other-spinal giving better results than hip-spinal-other. The former option means a maximum of one turnover is required between a spinal operation and another type of operation, which is preferable. Thus the ordering hip-other-spinal will be used here.

No tardiness relating to start time is allowed. Turnover times between operations are set to 75% of the original modelled value. Of course it would be more desirable to reduce these even further in pursuit of an ideal system, but the simulation still needs to represent a realistic and implementable system. Reducing turnover by more than this value could impose more difficult and less reasonable targets to theatre staff, but since a reduction of 25% is only a saving of approximately ten minutes per turnover, it is deemed realistic. It was also shown that by reducing turnover by this amount, improvements were still seen.

Theatre allocation is kept at 11.5 hours per day. While fewer cancellations are seen for longer hours, the theatre currently operates to these times and therefore it is unchanged to cause minimal prospective upheaval. If the number of scheduled hours was altered, changes to staffing would be needed not only in the trauma theatre, but also to portering services, recovery ward staff and potentially across the wider hospital. The value of r_l is set to 0.5, meaning that half of all required overruns will be allowed. This brings more control to the theatre by reducing the amount of unplanned hours.

Spinal operations are not removed from the theatre lists. It was shown that doing this resulted in relatively little impact and so it is not deemed necessary. Additionally, changes

made to tardiness, turnover times and ordering of operations will result in a more efficient system, freeing up additional hours and thus making it unnecessary to remove patients.

Some results are presented in Figure 8.9i, comparing this model with the current system, where better results are seen in each case.

Average theatre busy time in the new model is 9.4 hours per day, giving a mean daily reduction in busy time of 75 minutes. Theatre utilisation would fall to 82%. This is within recommended guidelines but also not so low to result in an inefficient system. Having an average utilisation under the recommended value of 90% also provides a buffer so that, if zero tardiness or the 25% reduction in turnover was not possible, for example, the guideline may still not be breached.

The percentage of cancelled operations falls from 15.7% to 9.6%, overall representing an improvement of 152 fewer cancellations per year. The hip cancellation percentage in Figure 8.9i gives the proportion of all cancellations which are hip patients; this falls from 29.3% to almost none. The hip cancellation rate gives the percentage of scheduled hip patients that are cancelled, falling from 18.5% to, again, almost none. Cancellations in the new model are almost exclusively non-hip patients.
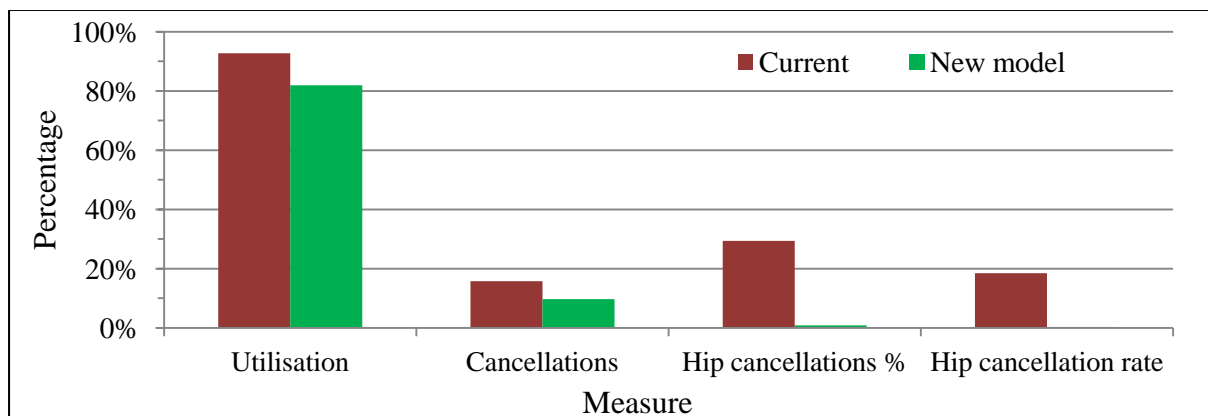

Figure 8.9i: Effect of designing a trauma theatre system around hip patients

This example provides just one idealised model of designing a trauma theatre which favours hip fracture patients. It is stressed that in doing this a system has been created which not only vastly improves outcomes for hip patients, but also provides better results across all patients.

8.10   Chapter summary

A simulation model of the trauma theatre at the UHW has been presented in this chapter. Validation and verification procedures showed that the model represented the real system with sufficient accuracy. The model was then used to explore a number of what-if scenarios, in order to investigate the effect of changing the running of the theatre. A summary of some of the scenarios tested, alongside results of the key outputs of lack of time cancellations, as a percentage of all scheduled patients, and theatre busy time and utilisation, is presented in Table 8.10i.

Recommendations to be made on the basis of these results will ultimately depend upon whether priority can be given to hip patients. Clearly the best results, from the point of view of a hip fracture patient/stakeholder, are seen when a system is designed around hip patients, see Section 8.9, where hip patients were given priority, turnover was reduced and all theatre sessions started on time. However, it was shown previously that not all changes need to be made to see beneficial results; for example, simply by scheduling all hip patients first, a considerable reduction in cancellation rates is seen for these patients, but this is generally compensated for by an increase in cancellation rates for other operations. However, mostly beneficial results are seen when alike operations are scheduled sequentially, so it is stressed that this should be achieved whenever possible.

An increase of one hour to the theatre allocation, from 11.5 to 12.5 hours per day, will mean that utilisation falls within the 90% guideline. It is not recommended that allocation is decreased since utilisation is already above 90% and any decrease will see utilisation increase further. A change to the cancellation strategy has also been documented. It is recommended that an increase in theatre allocation could be coupled with a stricter, more controlled, strategy where, say, only half of all required overruns are permitted.

Finally, attention is drawn to the scenario where an increase or decrease in demand is seen. Results show that simply matching the change in demand with a change to theatre open hours will not give the same results as what is currently seen. An increase in demand, for example, is actually shown to increase theatre utilisation, even if the increase is matched with an equivalent increase to theatre allocation, but cancellation rates are reduced.

Table 8.10i: Summary of results, trauma theatre model (* half between alike operations)

Scenario	Lack of time cancellations (%)		Theatre busy time (hours)	Theatre utilisation (%)
	Hips	**All**		
Current	18.5	15.7	10.7	92.8
Change to start time tardiness				
0% of current	16.0	12.2	10.5	91.0
200% of current	21.2	18.3	10.9	94.6
Change to turnover time				
0% of current	2.1	1.9	9.1	79.2
200% of current	38.7	32.7	10.7	92.7
Fixed to 30 minutes	11.9	10.5	10.3	89.6
Fixed to 30 minutes *	8.1	7.0	9.8	85.0
Fixed to 60 minutes	27.4	24.4	11.3	97.9
Fixed to 60 minutes *	18.6	15.7	10.7	93.2
Priority to hip patients				
First patient only	15.5	15.6	10.7	92.7
All hip patients	3.2	15.3	10.7	93.0
Alike operations in succession				
HSO	0.7	16.0	10.6	92.8
HOS	0.7	14.8	10.6	92.4
SHO	1.5	15.8	10.6	92.3
SOH	39.9	14.7	10.6	92.5
OHS	35.0	14.0	10.7	92.7
OSH	40.3	14.3	10.7	92.7
Change to theatre allocation				
10.5 hours	24.4	20.8	10.1	96.6
12.5 hours	13.7	11.5	11.2	89.2
13.5 hours	9.6	8.2	11.5	85.3
14.5 hours	6.7	5.6	11.8	81.4
Change to cancellation strategy				
$r_l = 0$, no overruns allowed	24.9	20.9	10.0	87.1
$r_l = 0.5$	19.9	17.0	10.5	91.3
$r_l = 1$, all overruns allowed	<0.1	<0.1	12.4	108.1
Change in demand				
Decrease of 20%	8.4	10.2	9.3	81.1
Increase of 20%	24.5	27.8	11.5	100.3
Increase of 50%	37.4	41.2	12.1	105.5
Change in demand and allocation				
Decrease of 20%	22.3	18.4	8.4	91.6
Increase of 20%	15.7	13.6	13.0	94.0
Increase of 50%	12.4	11.1	16.5	95.8
System designed around hip fracture patients	0.3	9.6	9.4	81.9

CHAPTER 9: A THEORETICAL APPROACH TO MODELLING THE TRAUMA THEATRE

9.1 Introduction

The trauma theatre at the UHW was previously modelled in Chapter 8 using simulation. In the next two chapters, the same theatre is modelled but using an analytical mathematical approach.

Results from queuing theory are utilised in order to model patients (customers) as they arrive, are operated on (served) and ultimately exit the system. Using queuing theory to model the theatre complements and extends previous investigations using simulation. It provides a robust means of investigating the impact that making changes to the system would have, so that high confidence can be placed on results using proven mathematically methodology. Over the next two chapters, some existing results from queuing theory are used and developed, while new, specific formulations are also presented where a tailored model was designed to represent the trauma theatre at the UHW.

9.1.1 The Erlang distribution

The Erlang distribution is a continuous probability distribution which is widely used in queuing theory. It was originally developed by A. K. Erlang for the field of telephone traffic engineering, specifically to examine the number of telephone calls which could be made simultaneously to operators in a switching station (Erlang 1920).

Consider k independent identically distributed random variables, each of which follows the Negative Exponential distribution with the same parameter α, so that $f(x_i) = \alpha e^{-\alpha x_i}$, $i = 1, ..., k$. Consider the general case of these k events occurring in series; the time spent in the ith phase is represented by x_i, while the probability that x_i is the time taken to complete this interval is given by $f(x_i)$. The Erlang distribution, sometimes referred to as the Erlang-k distribution, is the sum of these independent variables.

Note that it is not necessarily a requirement that these are physical phases, but that this formulation may be used a mathematical device to represent the time taken to complete a certain process. Also note that each phase must be completed before exit from the system can occur.

Let $t = \sum_{i=1}^{k} x_i$ be the total time taken to complete all phases. The probability density function (PDF) of the total time to complete the k phases is given by $g(t)$, where

$$g(t) = \frac{\alpha(\alpha t)^{k-1} e^{-\alpha t}}{(k-1)!}, \quad t \geq 0.$$

For continuity with other functions, let μ^{-1} be the mean of this distribution. Rearranging for α and substituting, the PDF for the Erlang distribution with k phases, which may be denoted by E_k, can be expressed as

$$g(t) = \frac{k\mu(k\mu t)^{k-1} e^{-k\mu t}}{(k-1)!}, \quad t \geq 0.$$

The variance is given by $\frac{1}{k\mu^2}$. Additionally note that $\alpha = k\mu$.

As k increases, the Erlang distribution becomes less skewed and begins to resemble the Normal distribution; consequently the distribution becomes more concentrated about the mean. This can be seen in Figure 9.1.1i where μ is kept constant at 0.5.
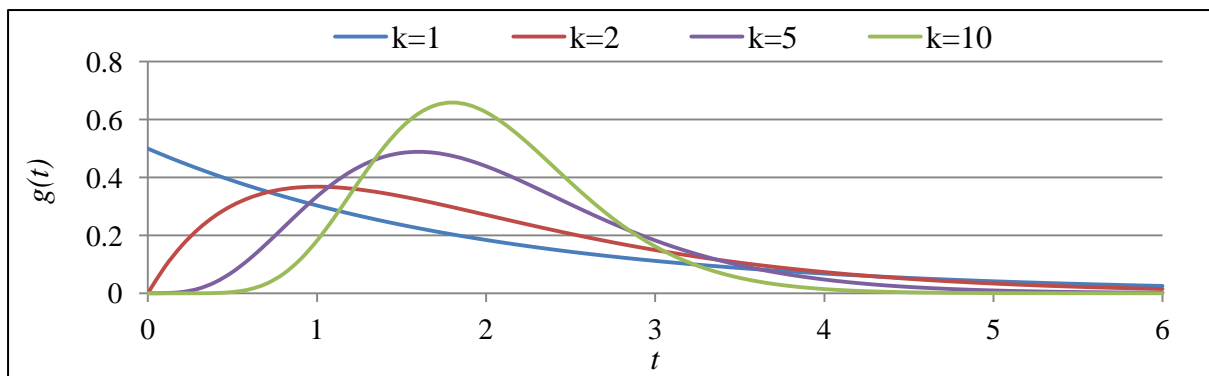

Figure 9.1.1i: The changing shape of the Erlang distribution as k is varied

9.1.2 Relationship to other distributions

Recall that the Erlang-k distribution represents k independent random variables, each of which follows an identical Negative Exponential distribution. Therefore if $k=1$, then E_k simply collapses to the Negative Exponential distribution with parameter μ ($=\alpha$). Since k represents the number of phases, it must be a positive integer. If this criterion is generalised to let k be real, then this is equivalent to the Gamma distribution, using the Gamma function ($\Gamma(k) = \int_0^\infty t^{k-1} e^{-t}\, \mathrm{d}t$) instead of the factorial function in the denominator. One final relationship is that if $k\mu = 0.5$, then E_k simplifies to the Chi-square distribution with $2k$ degrees of freedom.

9.2 Extension of the Erlang-*k* distribution

The previous formulation assumed an identical distribution for each phase. This assumption is now relaxed; each phase still follows a Negative Exponential distribution but the rate parameter of this distribution can vary between phases.

9.2.1 Case $k = 2$

Consider the simplest case of two phases, where the time spent in the first phase is given by x_1 and the time spent in the second phase is given by x_2. The probability that the times taken to complete these phases are x_1 and x_2 are given by $f(x_1)$ and $f(x_2)$ respectively, so that

$$f(x_i) = \alpha_i e^{-\alpha_i x_i}; \ x_i \geq 0, \ i = 1, 2.$$

Consider the Laplace transform of a function $f(t)$, defined for all real numbers $t \geq 0$, and denoted by the function $F(s)$, where the parameter s is a complex number. Then

$$F(s) = L_s\{f(t)\} = \int_0^\infty f(t)e^{-st} \, dt. \quad \text{Now consider the Laplace transform of } f(x_i), \text{ as}$$

previously defined, so that

$$L_s\{f(x_i)\} = \int_0^\infty \alpha_i e^{-\alpha_i t} e^{-st} \, dt = \alpha_i \int_0^\infty e^{-(\alpha_i + s)t} \, dt$$

$$= \frac{\alpha_i}{-(\alpha_i + s)}[e^{-\infty} - e^0] = \frac{\alpha_i}{\alpha_i + s}, \quad i = 1, 2.$$

Now let t be the time taken to complete the two phases, so that $t = \sum_{i=1}^{2} x_i$, and let $h(t)$ represent the PDF of the total time to complete the two phases. Using the Convolution Theorem, which states that $L_s\{f \cdot g\} = L_s\{f\} \cdot L_s\{g\}$ for two functions f and g, yields

$$L_s\{h(t)\} = L_s\{f(x_1)\} . L_s\{f(x_2)\}$$

$$= \frac{\alpha_1 \alpha_2}{(\alpha_1 + s)(\alpha_2 + s)} = \alpha_1 \alpha_2 \left(\frac{1}{\alpha_2 - \alpha_1}\right)\left(\frac{1}{\alpha_1 + s} - \frac{1}{\alpha_2 + s}\right).$$

To find $h(t)$, this is inverted and the linearity property of Laplace transforms is utilised;

$$h(t) = L_s^{-1}\left\{\frac{\alpha_1\alpha_2}{(\alpha_1+s)(\alpha_2+s)}\right\} = \left(\frac{\alpha_1\alpha_2}{\alpha_2-\alpha_1}\right)\left[L_s^{-1}\left\{\frac{1}{\alpha_1+s}\right\} - L_s^{-1}\left\{\frac{1}{\alpha_2+s}\right\}\right]$$

$$= \left(\frac{\alpha_1\alpha_2}{\alpha_2-\alpha_1}\right)\left[e^{-\alpha_1 t} - e^{-\alpha_2 t}\right].$$

9.2.2   Case $k = 3$

Now consider the case of three phases, where the times spent in each phase are x_1, x_2 and x_3, so that $f(x_i) = \alpha_i e^{-\alpha_i x_i}$, $x_i \geq 0$, $i = 1, 2, 3$.

Again letting $h(t)$ represent the PDF of the total time to complete all three phases and following the method seen previously, the following result is obtained:

$$L_s\{h(t)\} = L_s\{f(x_1)\}.L_s\{f(x_2)\}.L_s\{f(x_3)\}$$

$$= \frac{\alpha_1\alpha_2\alpha_3}{(\alpha_1+s)(\alpha_2+s)(\alpha_3+s)}.$$

Rearranging gives

$$h(t) = L_s^{-1}\left\{\alpha_1\alpha_2\alpha_3\left[\frac{A}{(\alpha_1+s)} + \frac{B}{(\alpha_2+s)} + \frac{C}{(\alpha_3+s)}\right]\right\}$$

where $A = \dfrac{1}{(\alpha_2-\alpha_1)(\alpha_3-\alpha_1)}$, $B = \dfrac{1}{(\alpha_1-\alpha_2)(\alpha_3-\alpha_2)}$, $C = \dfrac{1}{(\alpha_1-\alpha_3)(\alpha_2-\alpha_3)}$.

Using the linearity property for Laplace transforms yields the following result:

$$h(t) = \alpha_1\alpha_2\alpha_3\left[A.L_s^{-1}\left\{\frac{1}{\alpha_1+s}\right\} + B.L_s^{-1}\left\{\frac{1}{\alpha_2+s}\right\} + C.L_s^{-1}\left\{\frac{1}{\alpha_3+s}\right\}\right]$$

$$= \alpha_1\alpha_2\alpha_3\left[\frac{1}{(\alpha_2-\alpha_1)(\alpha_3-\alpha_1)}e^{-\alpha_1 t} + \frac{1}{(\alpha_1-\alpha_2)(\alpha_3-\alpha_2)}e^{-\alpha_2 t} + \frac{1}{(\alpha_1-\alpha_3)(\alpha_2-\alpha_3)}e^{-\alpha_3 t}\right].$$

9.2.3 General k

In general, consider k random variables each of which can be represented by the Negative Exponential distribution with a PDF of

$$f\left(x_i\right) = \alpha_i e^{-\alpha_i x_i}; \; x_i \geq 0, \; i = 1,\ldots,k, \; \alpha_i = k_i \mu_i.$$

Let $t = \sum_{i=1}^{k} x_i$, then the probability density function of the sum of the k random variables is given by

$$h\left(t\right) = \prod_{i=1}^{k} \alpha_i \left\{ \sum_{j=1}^{k} \left(\frac{1}{\prod_{i=1,i\neq j}^{k} \left(\alpha_i - \alpha_j\right)} e^{-\alpha_j t} \right) \right\} \qquad t \geq 0; \; \alpha_{i,j} \geq 0, \; i, j = 1,\ldots,k.$$

This may be adjusted to bring in a minimum value for each of the stages, where the minimum value for j^{th} phase is given by $\min_j$, so that

$$h\left(t\right) = \prod_{i=1}^{k} \alpha_i \left\{ \sum_{j=1}^{k} \left(\frac{1}{\prod_{i=1,i\neq j}^{k} \left(\alpha_i - \alpha_j\right)} e^{-\alpha_j\left(t-\min_j\right)} \right) \right\} \qquad t \geq \min_j; \; \alpha_{i,j} \geq 0, \; i, j = 1,\ldots,k.$$

9.3 Application to theatre process

This mathematical formulation lends itself well to the process by which a patient goes through the operating theatre at the UHW. There are several stages to be completed, each of which must be completed before entry to the next stage can occur. All stages must be completed in series before exit from the system is possible and each stage may be assumed to be independent from each other stage. Identification of an accurate PDF to represent theatre time could prove useful and this is the aim here. Additionally note that the Erlang distribution can be considered as a potentially appropriate distribution for each stage due to the mean exceeding the standard deviation in every case (see Table 9.3.2i).

However, it is not the simplest case of one Erlang distribution. Instead consider each stage in turn; if each stage can be represented its own distinct Erlang distribution, then these can be combined in series and the situation discussed in Section 9.2 is achieved. There will be many phases in series, each of which is assigned a Negative Exponential distribution, some of which will have equivalent parameters (and together form an Erlang distribution for that particular stage).

In order to pursue this method, Erlang distributions for the stages of the theatre process must be found. It should be noted that this investigation concentrates solely on hip fracture patients.

9.3.1 Software limitations

Previously the software package Stat::Fit had been used to fit theoretical distributions to empirical data. However, for this analysis, some limitations with this software were found.

Firstly, one can only specify the number of intervals that the data is split into. The size of each interval is then calculated by taking the range of the data and dividing by the number of intervals. Because of this, the intervals are likely to fall at untidy intervals, which is undesirable. Additionally, aside from specifying the number of intervals and altering the minimum value, the user has no control over any parameter estimation, which is sometimes useful in order to exercise control over investigating the effect of varying these parameters on other parameter estimates, moment estimates and the overall fit. For example, with the case

of the Erlang distribution, the fitted value of k is simply given by Stat::Fit, but the user cannot experiment with various different values in order to find a different fit.

Recall that the mean of the Erlang distribution with k phases is given by $\dfrac{1}{\mu}$ and the variance is given by $\dfrac{1}{k\mu^2}$, thus $k = \left(\dfrac{mean}{s.d.}\right)^2$, where $s.d.$ denotes the standard deviation of the distribution. This result may be useful when fitting the Erlang distribution and hence the option to try out different values of k could prove to be advantageous to the fitting process.

9.3.2 Methodology

The Solver add-in for Microsoft Excel (see Section 7.2.1) was utilised here to find appropriate Erlang distributions for each of the theatre stages for all hip operations. In order to achieve higher levels of accuracy and minimise the total number of phases, several combinations of intervals were tried and ultimately a total of four intervals were determined (which are a slight variation on those presented in Section 7.4.1):

>   Stage A – Pre-theatre (*asked for – into anaesthetic room*);
>   Stage B – Anaesthetic procedure (*into anaesthetic room – into theatre*);
>   Stage C – Theatre time (*into theatre – operation finish*);
>   Stage D – Recovery (*operation finish – out of recovery*).

By entering the cumulative density function of the Erlang distribution alongside associated empirical probabilities calculated from the data, then calculating the square of the difference between each of the fitted and empirical values, the parameters of k and μ within the Erlang function can be altered in order to minimise the sum of squared differences.

For comparison purposes, the parameter estimates given by both Stat::Fit and Solver are given in Table 9.3.2i, alongside the first and second moment estimates. Values of k attempted with Solver were initially $\lfloor k_t \rfloor$ and $\lceil k_t \rceil$, where k_t is the value of k given theoretically and calculated using the data (and entered as k in the *Data* row); that is, the square of the mean divided by the standard deviation. If no reasonable fit was found by either of these values, then other neighbouring integers were tried systematically. In two of

the four cases the same value of k was given by both Stat::Fit and Solver and these were equal to $\lceil k_t \rceil$ each time.

The Solver solutions were tested by using the Chi-square goodness-of-fit test, which is used to test if a sample of data come from a population with a specific distribution (Snedecor and Cochran 1991). In each case it was found that the distribution could be accepted statistically. This was supported by the graphical fit for each of the four stages; an example is given in Figure 9.3.2ii for Stage B. The three other stages displayed similar results.

The evidence is clear that better fits can be gained by using Solver and so in each case the parameter estimates given by this method were used going forward.

Table 9.3.2i: Comparison of Erlang fits for theatre stages A-D (minutes)

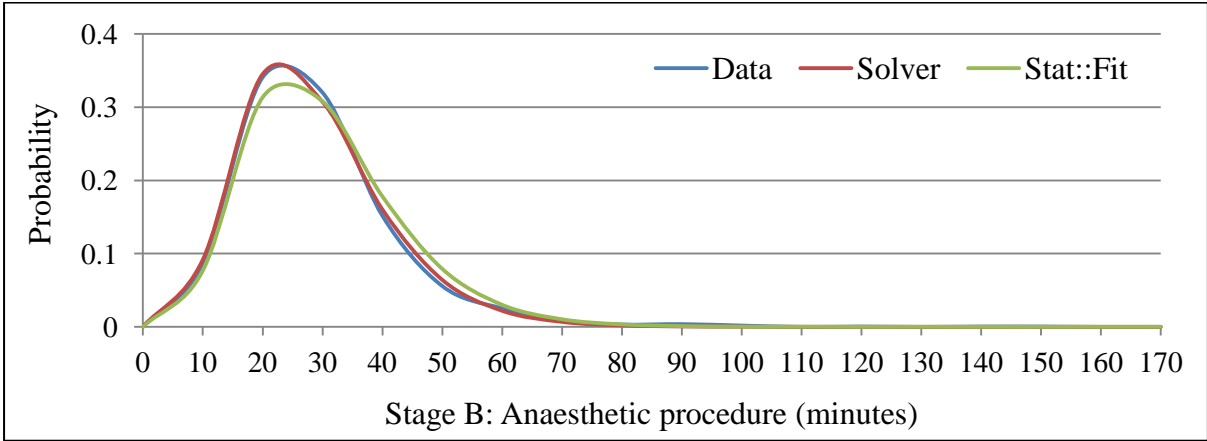| Stage	Method	k	μ	Minimum	Mean	Standard deviation
A	Solver	10	0.0318	6	37.5	10.0
	Stat::Fit	7	0.0328	6	36.5	11.5
	Data	9.50	-	6	37.6	12.2
B	Solver	4	0.0422	0	23.7	11.8
	Stat::Fit	4	0.0395	0	25.3	12.6
	Data	3.37	-	0	24.8	13.5
C	Solver	5	0.0133	0	75.3	33.7
	Stat::Fit	5	0.0119	0	84.2	37.6
	Data	4.13	-	0	79.1	38.9
D	Solver	5	0.0119	4	88.3	37.7
	Stat::Fit	4	0.0112	4	93.1	44.6
	Data	2.57	-	4	90.3	56.3

Figure 9.3.2ii: Graphical fits for the Erlang distribution, Stage B

9.3.3 Results

In summary, a total of four Erlang distributions have been fitted to the data and a summary of parameter estimates relating to these distributions is given in Table 9.3.3i.

Table 9.3.3i: Summary of fitted Erlang distributions for the theatre pathway (minutes)

Stage	k	μ	α	Minimum
A – Pre-theatre	10	0.0318	0.3178	6
B – Anaesthetic procedure	4	0.0422	0.1688	0
C – Theatre time	5	0.0133	0.0664	0
D – Recovery	5	0.0119	0.0593	4

Recall from Section 9.2.3 the formulation of the PDF for the sum of k random variables each following a Negative Exponential distribution. Here $k = 24$ in total but there are only four distinct Negative Exponential distributions. Since the α_i $(i = 1, .., k)$ values are not unique, the expression given in Section 9.2.3 will result in a zero in the denominator, and so a different approach is required to find a probability distribution to represent total theatre time. The mean and standard deviation, however, can be calculated and compared at this stage.

It follows from earlier that

$$L_s\{h(t)\} = \frac{\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1 + s)^{10}(\alpha_2 + s)^4(\alpha_3 + s)^5(\alpha_4 + s)^5}$$

281

where α_i represents the rate parameter for stage i, $i = 1,..,4$ (so that Stage 1 is Stage A using previous definitions). This formulation may be used to gain some inference about the quality of using this approach to represent the total theatre process time. For simplicity, $L_s\{h(t)\}$ will hereafter be denoted by $h*(s)$.

Firstly, some results of the application of the Laplace transform in the field of mathematical statistics are presented.

Let $f(x)$ be a PDF for a positive random variable X, so that f is positive and

$$\int_0^\infty f(x)\ dx = 1.$$

Then $f*(s)$ is defined as

$$f*(s) = \int_0^\infty e^{-st} f(t)\ dt = \int_0^\infty e^{-st}\ dF(t) = E\left[e^{-sX}\right].$$

It follows that the nth derivative of this is

$$f*^{(n)}(s) = \frac{d^n}{ds^n} E\left[e^{-sX}\right] = E\left[(-X)^n e^{-sX}\right].$$

Evaluating this expression at $s = 0$ yields

$$E[X] = -f*'(0)$$

$$E\left[X^2\right] = f*''(0)$$

$$\vdots \qquad \vdots$$

$$E\left[X^n\right] = (-1)^n f*^{(n)}(0).$$

The mean, μ, of f is therefore given by $-f*'(0)$, while the variance σ^2 is given by

$$E\left[X^2\right] - \{E[X]\}^2 = f*''(0) - \{-f*'(0)\}^2 = f*''(0) - \mu^2.$$

These results may now be applied to $h*(s)$ as presented previously. Differentiating with respect to s yields

$$h*'(s) = -\frac{10\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{11}(\alpha_2+s)^4(\alpha_3+s)^5(\alpha_4+s)^5} - \frac{4\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^5(\alpha_3+s)^5(\alpha_4+s)^5}$$

$$-\frac{5\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^4(\alpha_3+s)^6(\alpha_4+s)^5} - \frac{5\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^4(\alpha_3+s)^5(\alpha_4+s)^6}$$

and so the mean μ is obtained as

$$-h*'(0) = -\left\{-\frac{10}{\alpha_1} - \frac{4}{\alpha_2} - \frac{5}{\alpha_3} - \frac{5}{\alpha_4}\right\}.$$

Evaluating this expression at the relevant values of α_i, $i = 1,...4$, and adding on minimum values gives a mean value of 224.8 minutes, compared with an empirical value of 231.8 minutes. This empirical value is obtained by summing the empirical means of each of the four stages, while the overall mean was 227.5 minutes when taking an average of overall times; the slight discrepancy is due to data issues.

Differentiating again yields

$$h*''(s) = \frac{110\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{12}(\alpha_2+s)^4(\alpha_3+s)^5(\alpha_4+s)^5} + \frac{80\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{11}(\alpha_2+s)^5(\alpha_3+s)^5(\alpha_4+s)^5}$$

$$+\frac{100\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{11}(\alpha_2+s)^4(\alpha_3+s)^6(\alpha_4+s)^5} + \frac{100\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{11}(\alpha_2+s)^4(\alpha_3+s)^5(\alpha_4+s)^6}$$

$$+\frac{20\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^6(\alpha_3+s)^5(\alpha_4+s)^5} + \frac{40\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^5(\alpha_3+s)^6(\alpha_4+s)^5}$$

$$+\frac{40\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^5(\alpha_3+s)^5(\alpha_4+s)^6} + \frac{30\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^4(\alpha_3+s)^7(\alpha_4+s)^5}$$

$$+\frac{50\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^4(\alpha_3+s)^6(\alpha_4+s)^6} + \frac{30\alpha_1^{10}\alpha_2^4\alpha_3^5\alpha_4^5}{(\alpha_1+s)^{10}(\alpha_2+s)^4(\alpha_3+s)^5(\alpha_4+s)^7}$$

and

$$h*''(0) = \frac{110}{\alpha_1^2} + \frac{80}{\alpha_1\alpha_2} + \frac{100}{\alpha_1\alpha_3} + \frac{100}{\alpha_1\alpha_4} + \frac{20}{\alpha_2^2} + \frac{40}{\alpha_2\alpha_3} + \frac{40}{\alpha_2\alpha_4} + \frac{30}{\alpha_3^2} + \frac{50}{\alpha_3\alpha_4} + \frac{40}{\alpha_4^2}.$$

Evaluation of this expression and use of the appropriate formulae gives a theoretical standard deviation of 52.9 minutes, compared with an empirical standard deviation of 70.8 minutes (summing variances by stage) or 81.2 minutes when calculated overall.

Inverting $h^*(s)$ would facilitate finding an expression for $h(t)$; this methodology was pursued using the commercial computer algebra package MAPLE (MAPLE 1981-2010©). This resulted in an extremely long expression for $h(t)$ and substituting values of α_i $(i = 1,...4)$ proved problematic due to the high powers and small values involved; however, graphical comparisons can be formed using rounded expressions given by MAPLE. The probability distribution functions (PDFs) and cumulative distribution functions (CDFs) are now displayed. It can be seen that a reasonably good fit is found but that it becomes less accurate for longer total theatre times.
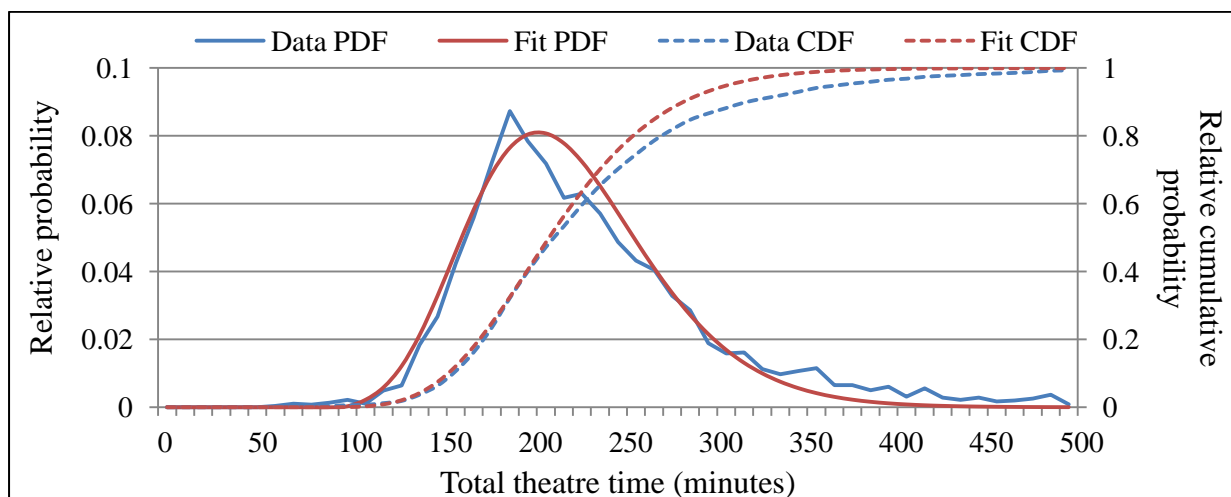
Figure 9.3.3ii: Comparison of theoretical and empirical values for total theatre time

9.3.4 Clinical time

The amalgamation of intervals B and C can be regarded as clinical time spent in the theatre suite and the time that clinical resources, such as the anaesthetic room, are being consumed. Let $h_c^*(s)$ be the Laplace transform of the PDF of clinical time, so that

$$h_c^*(s) = \frac{\alpha_2^4 \alpha_3^5}{(\alpha_2 + s)^4 (\alpha_3 + s)^5}.$$

Inverting this gives a more workable expression than the one obtained for total theatre time;

$$h_c(t) = \frac{1}{24} \alpha_2^4 \alpha_3^5 \left\{ \frac{840\left(-e^{-\alpha_2 t} + e^{-\alpha_3 t}\right)}{\left(\alpha_2 - \alpha_3\right)^8} + \frac{120t\left(3e^{-\alpha_2 t} + 4e^{-\alpha_3 t}\right)}{\left(\alpha_2 - \alpha_3\right)^7} + \frac{60t^2\left(-e^{-\alpha_2 t} + 2e^{-\alpha_3 t}\right)}{\left(\alpha_2 - \alpha_3\right)^6} \right.$$

$$\left. + \frac{4t^3\left(e^{-\alpha_2 t} + 4e^{-\alpha_3 t}\right)}{\left(\alpha_2 - \alpha_3\right)^5} + \frac{t^4 e^{-\alpha_3 t}}{\left(\alpha_2 - \alpha_3\right)^4} \right\}.$$

The theoretical mean and standard deviation are 99.0 and 35.7 minutes respectively compared with corresponding empirical values of 103.9 and 41.2 minutes. By these measures it appears that an accurate representation of clinical time has been found, which is further supported graphically, see Figure 9.3.4i.
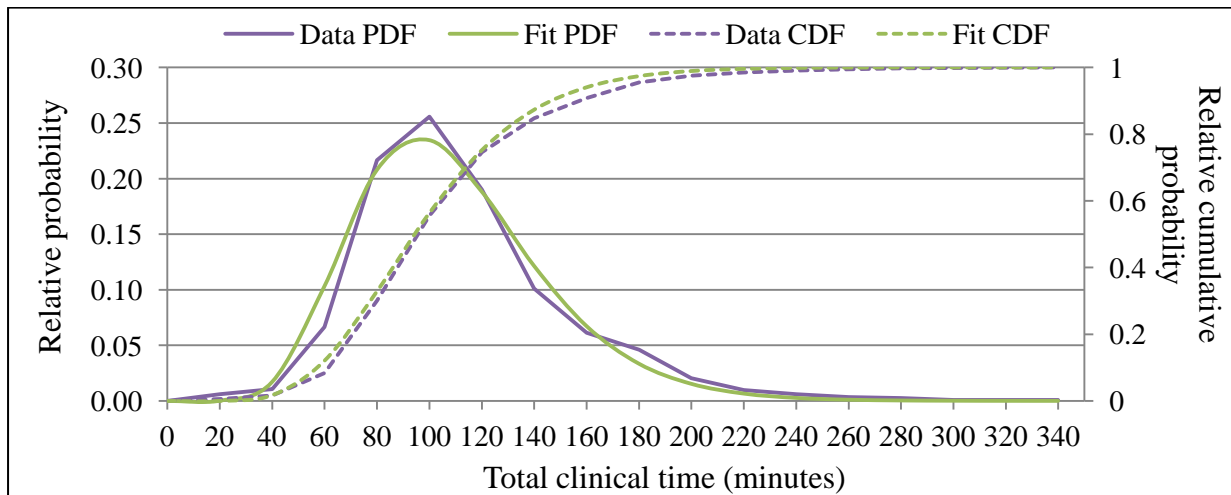
Figure 9.3.4i: Comparison of theoretical and empirical values for clinical time

9.4   The M | G | 1 queuing system

The three-category notation system to represent the characteristics of a queuing system was first proposed by D. G. Kendall (Kendall 1953) and has since been extended to include up to six factors, although five factors are more commonly used. The first factor represents the arrival process, the second represents the service time distribution and the third represents the number of servers in the system. The fourth and fifth, used later, represent the limit of the number in the system (or queue) and the queuing discipline. Here, random arrivals are assumed (denoted by M) and there is a general service distribution (denoted by G). These arrivals are served by one server, the anaesthetic room / operating theatre suite.

Random arrivals may be represented by the Negative Exponential distribution (particularly it is the inter-arrival time between successive arrivals that follows this distribution) with a given rate parameter λ, so that the mean inter-arrival time is given by $\dfrac{1}{\lambda}$. The Poisson distribution is intrinsically related to this; that is, the number of arrivals over a specified time period $[0, t]$ will follow the Poisson distribution with a mean parameter of λt.

It was previously stated that all arrivals did not follow a theoretical distribution and they were instead modelled using empirical values (see Chapter 6). Closer inspection of the Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit test statistics indicated that acceptance of the Negative Exponential distribution was one percentage point away (at a 5% significance level). This exercise was repeated excluding any patients who did not undergo surgery and acceptance was found, while it was also found that the number of arrivals per day did indeed follow a Poisson distribution with $\lambda = 1.270$; this fit is displayed in Figure 9.4i.
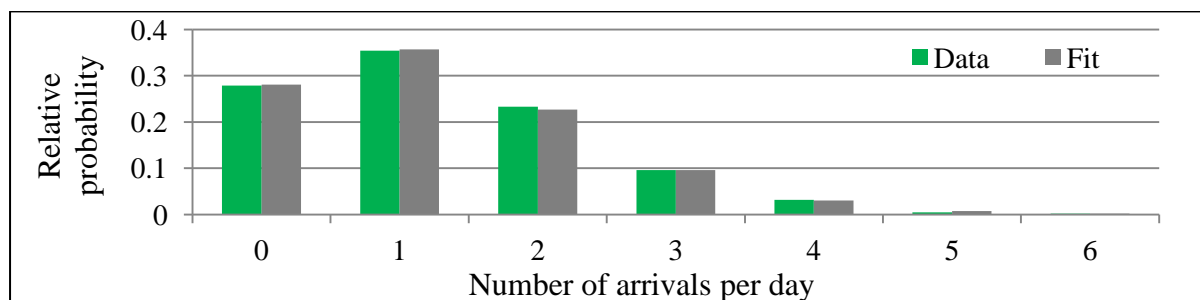

Figure 9.4i: Number of arrivals of trauma hip fracture patients requiring surgery per day against the Poisson distribution with parameter $\lambda = 1.270$

Let λ represent the arrival rate and μ represent the service rate (per unit time, so that mean service time is equal to $\dfrac{1}{\mu}$ units), while ρ, known as the utilisation factor or traffic intensity, is equal to $\dfrac{\lambda}{\mu}$. Also let the variance of the service time be represented by σ_s^2 so that the coefficient of variation of the distribution of service time is given by c_s where

$$c_s^2 = \frac{\sigma_s^2}{\left(1/\mu\right)^2} = \left(\sigma_s \mu\right)^2.$$

Note that $c_s^2 = 1$ when service times follow the Negative Exponential distribution and $c_s^2 = 0$ for constant service times, corresponding to changes to the queue discipline to M | M | 1 and M | D | 1, using standard notation.

To find summary results of an M | G | 1 system, only the arrival rate, service rate and coefficient of variation of service time are required. It is assumed that the system has infinite capacity and that customers are served on a first come first served, or first in first out (FIFO), basis. Additionally it is assumed that $\rho < 1$, otherwise the queuing delay becomes infinite.

The expression for the mean number of customers in the system, L, is a key result of queuing theory and is known as the Pollaczek-Khintchine formula (Khintchine 1932, Pollaczek 1930):

$$L = \frac{\rho^2\left(1+c_s^2\right)}{2\left(1-\rho\right)} + \rho.$$

Using Little's results (Little 1961), expressions for the mean time in the system $\bar{w}$, the mean number of customers in the queue L_q, and the mean wait in the queue $\bar{w}_q$, can be found as follows:

$$\bar{w} = \frac{1}{\lambda}L = \frac{1}{\mu}\frac{\rho\left(1+c_s^2\right)}{2\left(1-\rho\right)} + \frac{1}{\mu};$$

using $\bar{w} = \bar{w}_q + \dfrac{1}{\mu}$, $\bar{w}_q = \dfrac{1}{\mu}\dfrac{\rho\left(1+c_s^2\right)}{2\left(1-\rho\right)}$ and $L_q = \lambda\bar{w}_q = \dfrac{\rho^2\left(1+c_s^2\right)}{2\left(1-\rho\right)}.$

An important feature of these equations is that $\bar{w}_q$ and L_q both increase non-linearly as ρ increases, becoming very large as $\lambda \to \mu$ (as $\rho \to 1$), demonstrating the hazard of a stochastic system operating at a high level of utilisation.

Note that these results apply to a non-terminating system which has reached steady-state. If the observed arrival and service rates of hip fracture patients using the trauma theatre were inputted then excellent results (in terms of throughput efficiency) would be seen. However, it must be remembered that the theatre is used for other operations as well as hip fracture surgery and that additionally it is not scheduled to function on a 24-hour basis. In order to account for these issues, the value of λ is altered in the equations presented previously.

9.4.1   Altering the arrival rate

Let h_{hip} be the number of hours designated to hip surgery per day, then the proportion of time allocated to hips per day is $p_{hip} = \dfrac{h_{hip}}{24}$. If λ_{true} is the true arrival rate (for hip surgery only) then the steady-state arrival rate required for the queuing equations is given by $\lambda_{ss} = \lambda_{true} \cdot \dfrac{1}{p_{hip}}$. Since λ_{true} is known, λ_{ss} and p_{hip} can be amended in order to assess any impact upon the system. Making this amendment means that existing results from queuing theory can be applied as a non-terminating steady-state system can be assumed. This model essentially assumes that hip patients take priority over all other surgeries and that surgery can be performed at any time.

A similar approach, adjusting the arrival rate such that a non-terminating system can be assumed, has been reported for the M | M | 1 queuing system applied to the operating room (Tucker et al. 1999). This was done to ascertain the likelihood of needing to utilise a back-up staffing team during a night shift in order to make decisions regarding staffing requirements. The arrival rate was altered so that it represented a 24-hour system and then subsequently investigations were made into the probability of two or more patients simultaneously needing the operating room for a range of arrival volumes.

In order for the assumption of $\rho < 1$ to be satisfied (so that the queue does not become infinitely large and the system saturated), it must be true that $\dfrac{\lambda_{ss}}{\mu} < 1$. The mean service time is 99 minutes, so $\mu = \dfrac{1}{99}$. From this, it follows that in order to have a utilisation factor less than 1, it must be true that $p_{hip} > \dfrac{99.0 \text{x} 1.270}{60 \text{x} 24} = 0.0873$. This proportion equates to a theatre availability of 2.096 hours per day. The true numbers of hours used on average for hip patients is 2.3 hours per day, according to the data.

It is an additional requirement that $c_s^2 < 1$; here this value was $0.3606^2 = 0.1300$.

Results for the measures quoted in the previous section are now displayed for various theatre availability times, calculated by varying the arriving rate; 2.1 hours is excluded due to the instability of results when $\rho \approx 1$. Zero turnaround time is assumed. It can be seen that results are particularly sensitive between the values of 2.2 and 2.7 (approximately).

Results relating to number of patients are in solid lines and are recorded using the left-sided *y*-axis; results relating to the time in the system are in dotted lines and recorded using the right-sided *y*-axis. Values obtained were checked to be accurate via a Simul8 model (Simul8 1993-2010©).
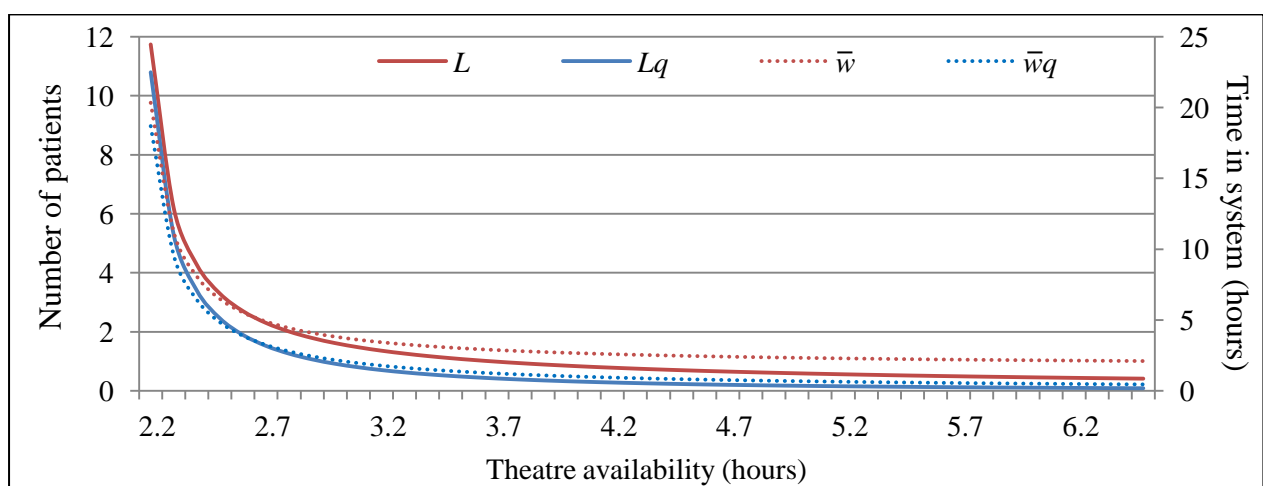

Figure 9.4.1i: Results for key queuing theory measures as the arrival rate is varied

It may be supposed that it is desirable in this scenario to have at least one patient in the queue since an empty operating theatre is a valuable waste of resources and so there should always be somebody waiting to enter service; that is $1 < \dfrac{\rho^2\left(1+c_s^2\right)}{2\left(1-\rho\right)}$, where $\rho = \dfrac{\lambda_{ss}}{\mu}$.

Rearranging for λ_{ss} yields $\left(1+c_s^2\right)\lambda_{ss}^2 + 2\mu\lambda_{ss}^2 - 2\mu^2 < 0.$

Solving this quadratic gives

$$\lambda_{ss} < \frac{\mu\left(-1+\sqrt{3+2c_s^2}\right)}{1+c_s^2},$$

and then by substituting values of μ, c_s^2, λ_{true} and λ_{ss} it can be shown that p_{hip} should be at most 0.122 in order to ensure that $L_q > 1$. This is equivalent to stating that no more than 2.94 hours per day should be allocated to hip patients. Of course this does not guarantee that there is always a queue, but that there will be at least one person waiting on average.

9.4.2 Sensitivity analysis

Results from the previous section implied that the theatre needs to be available for at least 2.1 hours per day in order that $\rho < 1$. Performing a sensitivity analysis on the arrival rate allows information to be gained on theatre availability and system results should the number of arrivals change. Minimum h_{hip} is therefore the number of hours per day which must be allocated to hip patients in order to ensure $\rho < 1$. Results are then given for setting the daily allocation to 2.5 hours for hip fracture patients. This value is chosen as it ensures that $\rho < 1$ while keeping $L_q > 1$. Inspection of Figure 9.4.1i also shows considerable gains are not made by increasing theatre availability beyond this value.

It can be seen that a 10% increase in the arrival rate would in fact lead to more than doubling both the queue length and queue time. The wait times here are evidently less than those observed in practice and will be in part due to assumptions made when formulating this queuing model; however, despite the necessity of making these assumptions, the results nonetheless provide some valuable insight.

Table 9.4.2i: Sensitivity analysis results for altering λ_{true}

Percentage change for λ_{true}	Minimum h_{hip}	Results for $h_{hip} = 2.5$			
		L	L_q	$\bar{w}$ (hours)	$\bar{w}_q$ (hours)
+10%	2.31	7.1	6.2	12.7	11.0
+5%	2.20	4.5	3.7	8.5	6.8
+2%	2.14	3.7	2.8	7.1	5.5
+1%	2.12	3.5	2.6	6.8	5.1
None	2.10	3.3	2.5	6.5	4.8
-1%	2.07	3.1	2.3	6.2	4.5
-2%	2.05	3.0	2.1	5.9	4.3
-5%	1.99	2.6	1.8	5.3	3.6
-10%	1.89	2.1	1.3	4.5	2.9

9.4.3 Distribution of P_n

Let P_n represent the probability that there are n patients in the system at any given time. As there is no explicit formula for P_n for the M | G | 1 queuing system, a method using probability generating functions is instead used to find these probabilities. Also let k_j represent the probability of j arrivals during a service time, so that for a service time distribution of $f(t)$, $k_j = \int_0^\infty \frac{(\lambda t)^j e^{-\lambda t}}{j!} f(t)\, dt$.

Let $G(z)$ and $K(z)$ be the probability generating functions for P_n and k_j respectively, such that $G(z) = \sum_{n=0}^{\infty} z^n P_n = P_0 + zP_1 + z^2 P_2 + ...$ and $K(z) = \sum_{n=0}^{\infty} z^j k_j = k_0 + zk_1 + z^2 k_2 + ...$.

Additionally, $K(z) = \int_0^\infty e^{-\lambda(1-z)t} f(t)\, dt = L_{\lambda(1-z)}\{f(t)\}$.

Then it may be shown that

$$G(z) = \frac{(z-1)(1-\rho)K(z)}{z - K(z)} = \frac{(1-\rho)(1-z)}{1 - \dfrac{z}{K(z)}}.$$

Recall that $L_s\{h_c(t)\} = \dfrac{\alpha_2^4 \alpha_3^5}{(\alpha_2 + s)^4 (\alpha_3 + s)^5}$, where $h_c(t)$ represents the PDF of clinical time.

Then in this case,

$$K(z) = L_{\lambda(1-z)}\{h_c(t)\} = \frac{\alpha_2^4 \alpha_3^5}{\left(\alpha_2 + \lambda(1-z)\right)^4 \left(\alpha_3 + \lambda(1-z)\right)^5}$$

$$\text{and } G(z) = (1-\rho)(1-z)\left[1 - z\frac{\left(\alpha_2 + \lambda(1-z)\right)^4 \left(\alpha_3 + \lambda(1-z)\right)^5}{\alpha_2^4 \alpha_3^5} \right]^{-1}.$$

By expanding this expression of $G(z)$ as an ascending power series in z, P_n is then given as the coefficient of z^n, achieved here using the fact that the power series expansion of the quotient $\dfrac{1}{1-x}$ is equal to $1 + x + x^2 + x^3 + \dots$ when $|x| < 1$.

Letting $x = z\dfrac{\left(\alpha_2 + \lambda(1-z)\right)^4 \left(\alpha_3 + \lambda(1-z)\right)^5}{\alpha_2^4 \alpha_3^5}$, expanding the above formulation of $G(z)$ and collecting terms in z, yields the result on the following page.

Substituting numerical values into this expression and selecting the relevant coefficient returns the probability of n patients in the system, $n = 0, 1, 2, \dots$. Note that if $n \geq 1$ then there is one patient in the operating theatre (the service channel) and the remainder are waiting.

Varying the arrival rate gives different values for P_n as the proportion of time allocated to hip fracture patients per day is altered. The arrival rate and service utilisation used here were calculated according to different restrictions on theatre allocation and cumulative results for $h_{hip} = 2.3, 2.5, 2.7$ and 2.9 are presented in Figure 9.4.3i, where $P(X \leq n) = \sum_{i=0}^{n} P_i$.

$$G(z) = (1-\rho)$$

$$+ z\left[(\rho-1)+(1-\rho)\frac{(\alpha_2+\lambda)^4(\alpha_3+\lambda)^5}{\alpha_2^4\alpha_3^5}\right]$$

$$+ z^2\left[(\rho-1)\frac{(\alpha_2+\lambda)^4(\alpha_3+\lambda)^5}{\alpha_2{}^4\alpha_3{}^5}\right.$$

$$\left.+(1-\rho)\left(\frac{(\alpha_2+\lambda)^3(\alpha_3+\lambda)^4}{\alpha_2{}^4\alpha_3{}^5}\right)\left(-\lambda\left(5(\alpha_2+\lambda)+4(\alpha_3+\lambda)\right)+\frac{(\alpha_2+\lambda)^8(\alpha_3+\lambda)^{10}}{\alpha_2{}^4\alpha_3{}^5}\right)\right]$$

$$+ z^3\left[(\rho-1)\left(\frac{(\alpha_2+\lambda)^3(\alpha_3+\lambda)^4}{\alpha_2{}^4\alpha_3{}^5}\right)\left(-\lambda\left(5(\alpha_2+\lambda)+4(\alpha_3+\lambda)\right)+\frac{(\alpha_2+\lambda)^8(\alpha_3+\lambda)^{10}}{\alpha_2{}^4\alpha_3{}^5}\right)\right.$$

$$+(1-\rho)\left(\frac{(\alpha_2+\lambda)^2(\alpha_3+\lambda)^3}{\alpha_2{}^4\alpha_3{}^5}\right)\left(\lambda^2\left(20(\alpha_2+\lambda)(\alpha_3+\lambda)+10(\alpha_2+\lambda)^2+6(\alpha_3+\lambda)^2\right)\right.$$

$$-\lambda\left(\frac{10(\alpha_2+\lambda)^6(\alpha_3+\lambda)^6+8(\alpha_2+\lambda)^5(\alpha_3+\lambda)^7}{\alpha_2{}^4\alpha_3{}^5}\right)$$

$$\left.\left.+\frac{(\alpha_2+\lambda)^{10}(\alpha_3+\lambda)^{12}}{\alpha_2{}^8\alpha_3{}^{10}}\right)\right]$$

$$+ \dots .$$
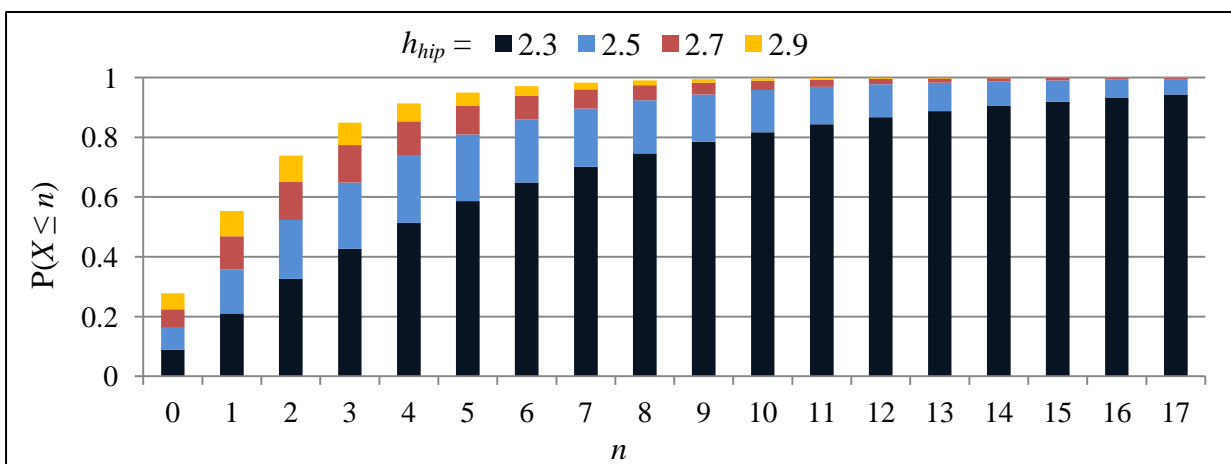
Figure 9.4.3i: P($X \leq n$) for different values of h_{hip}

It appears that 2.5 hours per day is the most appropriate theatre time allocation for this patient group. A shorter allocation than this (2.3 hours) will mean that there are five or more patients in the system approximately half of the time and ten or more approximately one-fifth of the time, which is unacceptable.

For 2.5 hours allocation, the probability that an arriving patient can be operated on immediately is given by the constant term (zero power in z), which is equal to $P_0 = 1 - \rho = 1 - 0.8382 = 0.1618$. It was previously hypothesised that having no patients in the system is not desirable due to wasted resources, and therefore it would be up to the hospital managers to decide whether a system that is empty 16% of the time, on average, is acceptable. Extending the allocation gives more preferable results from a patient perspective; an allocation of 2.7 hours gives a probability of no patients in the system of 22%, increasing to 28% for an allocation of 2.9 hours.

More detailed results for the chosen value of $h_{hip} = 2.5$ are now investigated. This results in a service utilisation of 0.8382 and results for P_n are presented in Figure 9.4.3ii. It can be seen that P_1 returns the highest probability and this is equal to 0.197. P_n then monotonically decreases as n increases (for $n \geq 1$). It is interesting to note that a modal value of one and the shape displayed below were seen for each of the four theatre allocations detailed previously, with the exception of $h_{hip} = 2.9$ where $P_0 > P_i > P_{i+1}$, $i \geq 1$ (but $P_0 \approx P_1$).
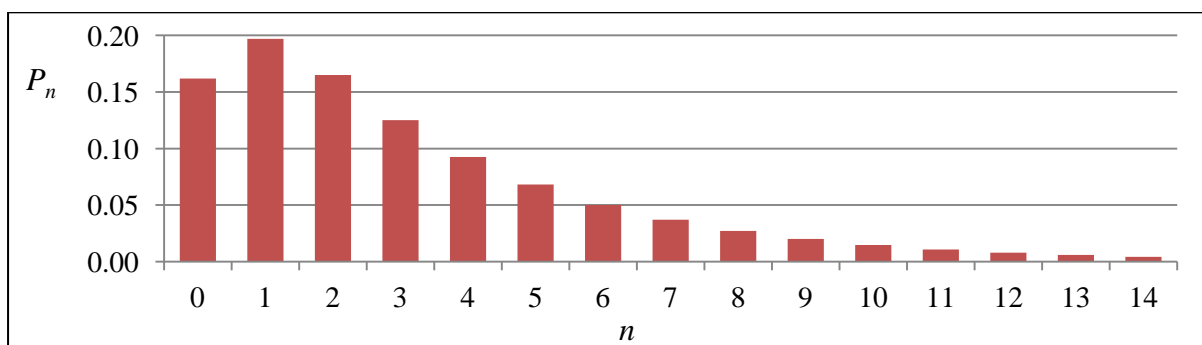

Figure 9.4.3ii: P_n for $h_{hip} = 2.5$

It follows that the probability that a patient arrives to find a system which is not empty, and therefore has to wait, is given by $P_1 + P_2 + P_3 + ... = 1 - P_0 = 1 - (1 - \rho) = 0.8382$.

The expected value of n is found via expansion of $E(n) = \sum_{n=0}^{\infty} n P_n$. The summation was terminated once probabilities diminished below 0.001 (at $n = 17$, as shown; note also that $\sum_{n=0}^{17} P_n = 0.9946$).

$$\sum_{n=0}^{\infty} n P_n = 0 P_0 + 1 P_1 + 2 P_2 + 3 P_3 + 4 P_4 + 5 P_5 + \ldots$$

$$= 0(0.1618) + 1(0.1968) + 2(0.1649) + 3(0.1250) + 4(0.0925) + 5(0.0682) + 6(0.0503)$$

$$+ 7(0.0370) + 8(0.0273) + 9(0.0201) + 10(0.0148) + 11(0.0109) + 12(0.0080)$$

$$+ 13(0.0059) + 14(0.0044) + 15(0.0032) + 16(0.0028) + 17(0.0007) \; [+\ldots]$$

$$= 3.168.$$

Recall also that L, the expected number in the system at any given time, was previously shown to equal 3.3 when $h_{hip} = 2.5$; the discrepancy being due to the termination of the sequence above. Inspection of the results gives the percentiles displayed in Table 9.4.3iii, where the right skew displayed previously is again evident.

Table 9.4.3iii: P_n percentiles for $h_{hip} = 2.5$

Percentile	1st	5th	10th	25th	50th	75th	90th	95th	99th
n	0	0	0	1	2	5	8	10	15

Despite the system being empty 16% of the time, the 75th percentile shows that there are five or more patients in the system during the busiest quartile. The busiest 10% of time will see eight or more patients in the system (seven or more waiting), while 1% of the time will see 15 or more patients in the system. The median value is two. It is concluded that setting $h_{hip} = 2.5$ gives a suitable trade-off between resource utilisation and patient management.

9.4.4 Waiting time

Let Q, C and T be the random variables of queuing time, clinical time and total time in the system, given by PDFs $h_q(t)$, $h_c(t)$ and $h_t(t)$ respectively. The Laplace transforms are given by $L_Q(s) = L_s\{h_q(t)\}$, $L_C(s) = L_s\{h_c(t)\}$ and $L_T(s) = L_s\{h_t(t)\}$. Then it can be shown that

$$L_T(s) = \frac{s(1-\rho)L_C(s)}{s - \lambda + \lambda L_C(s)}.$$

Since $Q + C = T$, where the total time T represents theatre delay (the time between arrival and surgery) plus theatre time (but not any post-operation length of stay), and Q and C can be assumed to be independent, then

$$L_Q(s) = \frac{L_T(s)}{L_C(s)} = \frac{s(1-\rho)}{s - \lambda + \lambda L_C(s)} = \frac{s(1-\rho)}{s - \lambda + \lambda\left(\dfrac{\alpha_2^4\alpha_3^5}{(\alpha_2 + s)^4(\alpha_3 + s)^5}\right)}.$$

Moments of Q may be found by differentiating and evaluating at $s = 0$ as previously shown:

$$L_Q'(s)\big|_{s=0} = \frac{5\lambda(\rho-1)\left(4\alpha_2\alpha_3 + 3\alpha_2^2 + 2\alpha_3^2\right)}{\left(\alpha_2\alpha_3 - 5\lambda\alpha_2 - 4\lambda\alpha_3\right)^2}$$

and $L_Q''(s)\big|_{s=0} = \dfrac{10\lambda(\rho-1)\left(7\alpha_2^4\alpha_3 + 12\alpha_2^3\alpha_3^2 + 10\alpha_2^2\alpha_3^3 + 4\alpha_2\alpha_3^4 + \lambda\theta\right)}{\alpha_2\alpha_3\left(\alpha_2\alpha_3 - 5\lambda\alpha_2 - 4\lambda\alpha_3\right)^3}$

where $\theta = 10\alpha_2^4 + 32\alpha_2^3\alpha_3 + 42\alpha_2^2\alpha_3^2 + 20\alpha_2\alpha_3^3 + 4\alpha_3^4$.

Substituting appropriate parameter values gives the mean and standard deviation of the waiting time for various theatre allocations, see Table 9.4.4i. Using $h_{hip} = 2.5$ as a benchmark, allocating 0.2 less hours per day would lead to a 98% increase in the mean waiting time, while an additional 0.2 hours would lead to 33% reduction in mean waiting time. If this is increased to $h_{hip} = 3.0$, then the mean waiting time is reduced by 55%. The percentage changes in standard deviation are slightly less, leading to an almost linear

relationship between h_{hip} and the coefficient of variation of waiting time; for each extra 0.1 hours allocated to hip surgery, the coefficient of variation increases by approximately 0.03.

Table 9.4.4i: Waiting time results (minutes) for different theatre allocations

h_{hip}	Mean	S.D.	C.V.
2.3	572.9	613.3	1.07
2.4	384.8	424.5	1.10
2.5	289.7	328.8	1.13
2.6	232.3	270.8	1.17
2.7	193.9	231.8	1.20
2.8	166.4	203.8	1.23
2.9	145.7	182.7	1.25
3.0	129.6	166.1	1.28

The relationship between waiting time, h_{hip} and ρ is now considered further. Using $-L'_Q(s)\big|_{s=0} = \bar{w}_q$ and rearranging and solving for λ, the value of h_{hip} and ρ that would achieve this mean waiting time can be calculated. Note that it is still a requirement that $\rho < 1$. As $\bar{w}_q$ increases, $\rho \rightarrow 1$ and $h_{hip} \rightarrow 2.096$ (the value of h_{hip} when $\rho = 1$).
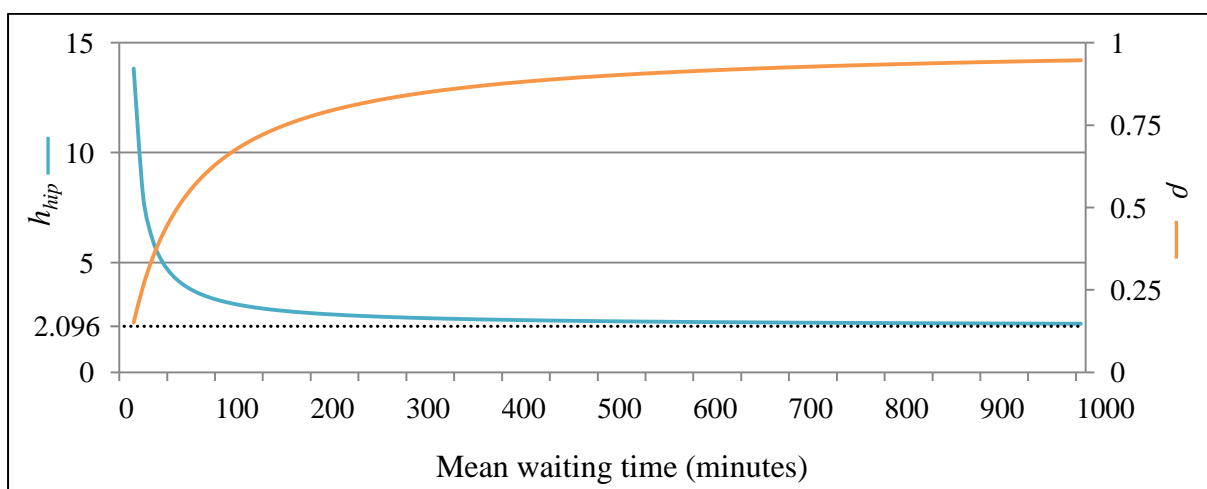
Figure 9.4.4ii: Relationship between waiting time, h_{hip} and ρ

Moments of T may be found directly from its Laplace transform or by using $T = Q + C$. Employing the latter approach and using

$$E[T] = E[Q] + E[C]$$

$$\text{and } \text{Var}[T] = \text{Var}[Q] + \text{Var}[C]$$

(there is no covariance term since Q and C are assumed to be independent) gives the following results for total time in the system.
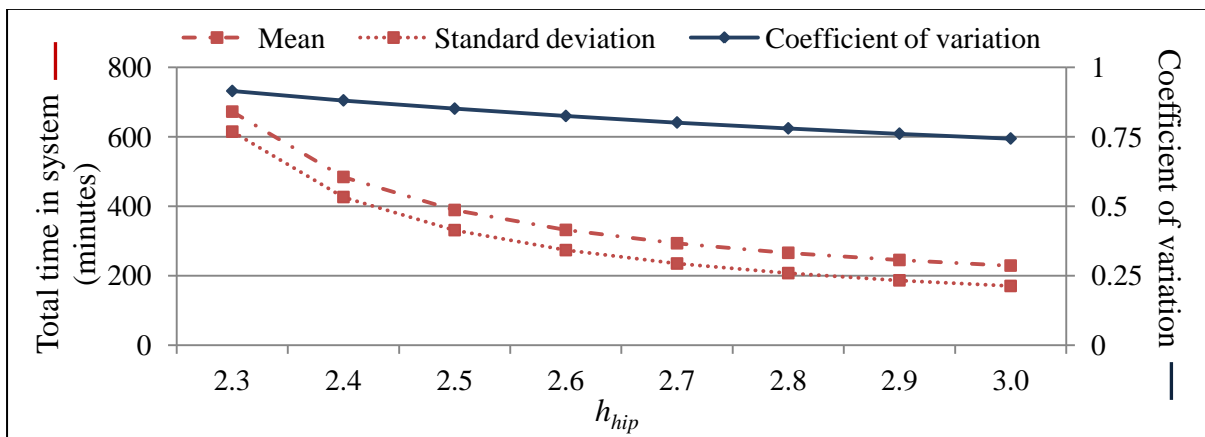

Figure 9.4.4iii: Results for total time in system

It is interesting to note that, while the mean and standard deviation are again relatively similar (as was found for queuing time), the mean in this case exceeds the standard deviation each time. Each coefficient of variation is now less than one and is declining as h_{hip} increases. Inversion of $L_Q(s)$ allows a theoretical waiting time distribution to be found, which may then be compared with simulated results. Due to the structure of $L_Q(s)$, inversion results in the Dirac-Delta function being included in $h_q(t)$.

Consider the function

$$f(t) = \frac{1}{\varepsilon} \quad a \leq t \leq a + \varepsilon$$

$$= 0 \quad \text{elsewhere;}$$

then the Dirac-Delta function, $\delta(t-a)$, is defined as $\delta(t-a) = \lim_{\varepsilon \to 0} f(t)$.

The result of this is an infinitely thin and infinitely tall 'spike' in the graph of $f(t)$ at a, with a value of 0 elsewhere. Inverting $L_Q(s)$ gave a probability distribution for waiting time, $h_q(t)$, as represented in Figure 9.4.4iv for different values of h_{hip}. (The formula is excluded here due to its substantial length.)

The Dirac-Delta part of $h_q(t)$ was found to be $(1-\rho)\delta(t-0)$; that is, there is a probability of $(1-\rho)$ of no wait. This is equivalent to the result shown earlier of a patient arriving to find an empty system, or $P_0 = 1 - \rho = 0.1618$. This term is excluded from the graph for display purposes.
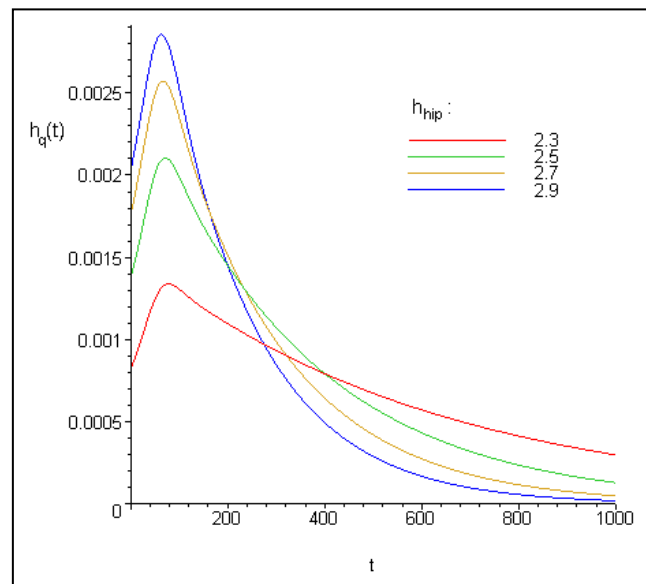
Figure 9.4.4iv: Waiting time distribution for different theatre allocations

A Simul8 model was used to find a number of other results and to verify this waiting time distribution. Results when $h_{hip} = 2.5$ are now presented. The shape initially appears to differ from Figure 9.4.4iv but this is because the (omitted) spike at $t = 0$ and the initial increase in $h_q(t)$ is encompassed into the first interval in Figure 9.4.4v. To compare the distributions more formally, consider the probability that the queuing time is less than 139 minutes (approximately). It can be seen that the Simul8 model returns a probability of 43%, while the theoretical formula gives a probability of 42.3%. A value of 139 minutes was used since it was automatically outputted by Simul8.
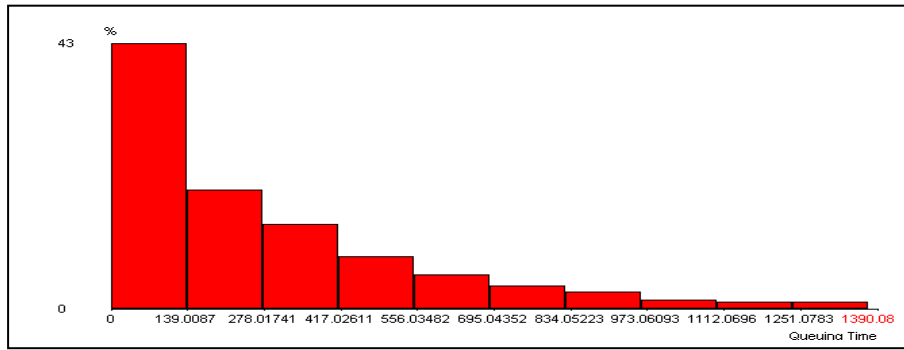
Figure 9.4.4v: Simulated waiting time distribution for $h_{hip} = 2.5$

An average queuing time of 4.8 hours was confirmed by the simulation model, 95% confidence interval (CI) [4.5, 5.1], as was the standard deviation of queuing time of 5.5 hours, 95% CI [5.0, 6.0]. The percentage of patients who have to wait (84%) was also verified. If only non-zero queuing times are considered, the average waiting time increases to 5.8 hours, 95% CI [5.4, 6.1]. The maximum queue length was 25.9 patients on average, 95% CI [21.6, 30.2].

Finally, probabilities of various wait times are now presented. These results are consistent between the theoretical and simulated models and again are displayed for various h_{hip} allocations.
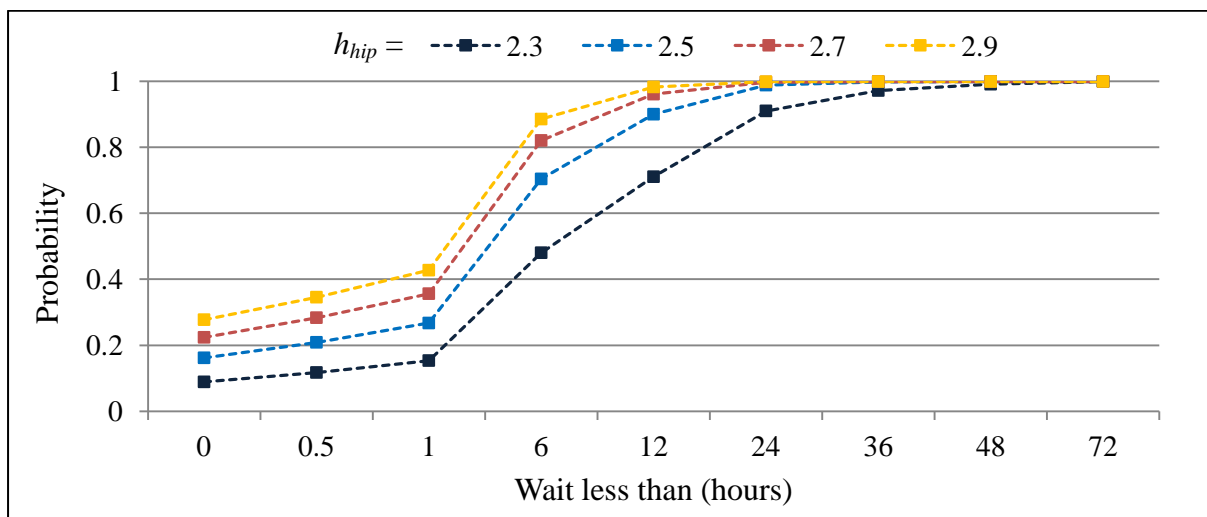
Figure 9.4.4vi: Waiting time results for different values of h_{hip}

Now consider a different way to interpret these results, where instead the wait times experienced by various patient proportions are presented. For example, half of all patients have wait time less than or equal to 3.1 hours (186 minutes) when $h_{hip} = 2.5$, while 90% of patients have a wait not exceeding 12 hours. No value is given for 25%, $h_{hip} = 2.9$ since $1 - \rho = 0.277$ and thus since more than 25% of patients have no wait it would be misleading to enter a zero value here.

Table 9.4.4vii: Waiting time (hours) for various theatre allocations

Proportion	Maximum wait time (hours)			
	$h_{hip} = 2.3$	$h_{hip} = 2.5$	$h_{hip} = 2.7$	$h_{hip} = 2.9$
25%	2.3	0.9	0.2	-
50%	6.4	3.1	2.0	1.4
75%	13.5	6.9	4.7	3.6
90%	22.9	12.0	8.3	6.4

With reference to earlier investigations in this thesis relating to operative delay, it is interesting to note that this system returns an average of 98.8% of patients receiving surgery within one day, rising to 99.9% within two days. The detrimental effects of operative delay have been well-documented throughout this thesis, relating to both mortality and additional time in hospital. This queuing model has demonstrated the vast improvements that could be achieved should an efficient operating theatre suite have 2.5 hours per day allocated to hip fracture patients.

9.5 Chapter summary

A mathematical approach has been used in this chapter in order to model the trauma operating theatre at the UHW. This has focussed on the Erlang distribution and in particular an extension to the traditional Erlang distribution in order to model clinical (service) time. The extension involved formulating a distribution with different rates between the phases. It was not possible to substitute numerical values into this formula directly, thus inversion of the Laplace transform of the distribution was required. The probability distribution function for clinical time formulated by this method was shown to represent the data with a high degree of accuracy.

This service time distribution was then used to apply to results of the M | G | 1 queuing system. In order to do this, a novel approach was taken to modelling arrivals. Since the trauma theatre is not, in reality, a non-terminating system, the arrival rate was amended so that this could in fact be assumed. This meant that recommendations could be made on the number of hours per day that should be assigned to hip fracture surgery.

Results were then formulated for a variety of measures, including the distribution of both the number in the system and waiting time. The maximum wait times for different proportions of patients at different values of assigned hip surgery hours per day were also calculated, which are useful in that they show which assignments will lead to decreases in operative delay. Overall it was recommended that agreeable results can be achieved with an allocation of 2.5 hours for hip surgery per day.

CHAPTER 10: FURTHER THEORETICAL APPROACHES TO MODELLING THE TRAUMA THEATRE

10.1 Introduction

In this chapter, the trauma theatre is modelled via a more traditional queuing theory approach. Several different models are introduced and their relevance to the system is explained.

The trauma theatre suite is the system under consideration. In these models, a patient (customer) joins the queue once they are 'asked for' from the ward, since this is when they become under the care of theatre staff. Service starts once their clinical time starts, which is at the commencement of the anaesthetic procedure. Service continues through the operation time and finishes when the patient exits the theatre, at which point they exit the system.

There are two types of random arrival, hip patients and non-hip patients. Hip arrivals to the UHW requiring surgery have previously been shown follow the Poisson distribution (see Section 9.4) and therefore the demand on the theatre by these patients is also be assumed to be random. Note that results presented previously show the satisfaction of random arrivals to the hospital, and not necessarily the trauma theatre suite. They are used here as a proxy to estimate the demand on the trauma theatre by trauma hip fracture patients. While this is an approximation, this approximation allows for the mathematical formulation to be carried forward; highly compatible results between the model and the data were subsequently achieved. This assumption is further corroborated by a comparison of the daily number of trauma hip surgeries per day against the Poisson distribution with mean 1.47 (see Section 10.2.2), see Figure D10.1a. The same logic can be applied to non-hip patients (Moore 2003).

Service time is dependent upon patient type. There is one server, the trauma theatre suite, and patients are served on a first come first served basis. Reneging is not permitted and zero turnover between operations is assumed; one patient enters the anaesthetic room as another leaves the operating room.

In order to evaluate the accuracy of the models, output from the real world system is required and a summary is given in Table 10.1i. The results quoted as percentages give the proportion of the total time for which the given measure was observed. A total of 1365 days was available for this analysis and counts of the number, location (queue or service) and type (hip

or non-hip) of patients were made at each minute of every day, giving almost two million data points in total. 0.12% of data points were deleted due to erroneous data; more than one patient in theatre, for example.

Table 10.1i: Trauma theatre data summary

Measure	Value	Result
In service	Hip	9.5%
	Non-hip	35.6%
	Nobody	54.9%
Number in system	0	52.1%
	1	38.6%
	2	9.3%
	3	< 0.1%
	≥ 4	0.0%
	Mean	0.572
	S.D.	0.657
Number in queue	0	87.1%
	1	12.9%
	2	< 0.1%
	3	< 0.1%
	≥ 4	0.0%
	Mean	0.130
	S.D.	0.338
Number of hip patients in system	Mean	0.131
	S.D.	0.122
Number of non-hip patients in system	Mean	0.440
	S.D.	0.590
Number of hip patients in queue	Mean	0.035
	S.D.	0.034
Number of non-hip patients in queue	Mean	0.094
	S.D.	0.294

It will be seen later that P_0 is defined as the probability of an empty system. For the data, this can be calculated in two ways; firstly, the proportion of time that there is nobody in the trauma theatre, and secondly, the proportion of time that there is nobody in the system, which were 0.549 and 0.521 respectively.

These differ because in reality there could be a patient in the loading bay, waiting to go into theatre, but no patient currently in theatre (indeed this will always be true for the first scheduled patient of the day). However, in a queuing model, an arrival which finds an empty system will start service immediately.

10.2 Negative Exponential model

There are two sources of Poisson arrivals, hip patients (type 1) and non-hip patients (type 2), which arrive according to rates λ_1 and λ_2 respectively. To begin with, it is assumed that service times follow two Negative Exponential distributions according to rates μ_1 and μ_2. The number of patients allowed in the system at any time is limited to a maximum of l_{sys}, so that if there are already l_{sys} patients in the system, no arrivals are permitted.

Using standard notation, an $M(\lambda_1, \lambda_2) | M(\mu_1, \mu_2) | 1 | l_{sys} |$ FIFO system has been described. For a system with a limit of l_{sys} patients, there are a total of

$$1 + 2\sum_{m=1}^{l_{sys}} m = 1 + 2\left(\frac{l_{sys}(l_{sys}+1)}{2}\right) = l_{sys}^2 + l_{sys} + 1$$

different system states.

10.2.1 Formulation

Let $P_{h,n}(t)$ be the probability of h hip patients and n non-hip patients in the system at time t, with steady-state probability $P_{h,n}$. An asterisk is used to denote which type of patient is in service, where relevant – for example, $P_{1*,1}$ gives the steady-state probability of one hip patient and one non-hip patient in the system, where the hip patient is in service. The data showed that the total number of patients in the system never exceeded three, and thus $l_{sys} = 3$, giving $3^2 + 3 + 1 = 13$ system states in total, with differential-difference equations as follows.

Consider the formulation of the equation for $P_{1*,1}(t + \delta t)$, and in particular the third term on the right-hand side of this equation. At time t, there are two hip patients and one non-hip patient in the system, with one of the hip patients in service. By time $t + \delta t$, this hip patient has been served (term $\mu_1 \delta t$). The next patient to enter service must be a hip patient to achieve a system state of 1*,1 at time $t + \delta t$, so that $\dfrac{\lambda_1}{\lambda_1 + \lambda_2}$, the proportion of hip arrivals in

comparison with all arrivals, is also included in the product. The same logic is applied to the fourth term in this equation and the following equation for $P_{1,1^*}(t+\delta t)$.

$$P_{0,0}(t+\delta t) = P_{0,0}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)$$
$$+ P_{1,0}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(\mu_1\delta t)$$
$$+ P_{0,1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(\mu_2\delta t)$$

$$P_{1,0}(t+\delta t) = P_{1,0}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-\mu_1\delta t)$$
$$+ P_{0,0}(t)(\lambda_1\delta t)(1-\lambda_2\delta t)$$
$$+ P_{2,0}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(\mu_1\delta t)$$
$$+ P_{1,1^*}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(\mu_2\delta t)$$

$$P_{0,1}(t+\delta t) = P_{0,1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-\mu_2\delta t)$$
$$+ P_{0,0}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)$$
$$+ P_{0,2}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(\mu_2\delta t)$$
$$+ P_{1^*,1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(\mu_1\delta t)$$

$$P_{2,0}(t+\delta t) = P_{2,0}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-\mu_1\delta t)$$
$$+ P_{1,0}(t)(\lambda_1\delta t)(1-\lambda_2\delta t)(1-\mu_1\delta t)$$
$$+ P_{3,0}(t)(\mu_1\delta t)$$
$$+ P_{2,1^*}(t)(\mu_2\delta t)$$

$$P_{0,2}(t+\delta t) = P_{0,2}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-\mu_2\delta t)$$
$$+ P_{0,1}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)(1-\mu_2\delta t)$$
$$+ P_{0,3}(t)(\mu_2\delta t)$$
$$+ P_{1^*,2}(t)(\mu_1\delta t)$$

$$P_{1*,1}(t + \delta t) = P_{1*,1}(t)(1 - \lambda_1 \delta t)(1 - \lambda_2 \delta t)(1 - \mu_1 \delta t)$$

$$+ P_{1,0}(t)(1 - \lambda_1 \delta t)(\lambda_2 \delta t)(1 - \mu_1 \delta t)$$

$$+ P_{2*,1}(t)(\mu_1 \delta t)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$$

$$+ P_{1,2*}(t)(\mu_2 \delta t)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$$

$$P_{1,1*}(t + \delta t) = P_{1,1*}(t)(1 - \lambda_1 \delta t)(1 - \lambda_2 \delta t)(1 - \mu_2 \delta t)$$

$$+ P_{0,1}(t)(\lambda_1 \delta t)(1 - \lambda_2 \delta t)(1 - \mu_2 \delta t)$$

$$+ P_{1,2*}(t)(\mu_2 \delta t)\left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)$$

$$+ P_{2*,1}(t)(\mu_1 \delta t)\left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)$$

$$P_{3,0}(t + \delta t) = P_{3,0}(t)(1 - \mu_1 \delta t)$$

$$+ P_{2,0}(t)(\lambda_1 \delta t)(1 - \lambda_2 \delta t)(1 - \mu_1 \delta t)$$

$$P_{0,3}(t + \delta t) = P_{0,3}(t)(1 - \mu_2 \delta t)$$

$$+ P_{0,2}(t)(1 - \lambda_1 \delta t)(\lambda_2 \delta t)(1 - \mu_2 \delta t)$$

$$P_{2*,1}(t + \delta t) = P_{2*,1}(t)(1 - \mu_1 \delta t)$$

$$+ P_{2,0}(t)(1 - \lambda_1 \delta t)(\lambda_2 \delta t)(1 - \mu_1 \delta t)$$

$$+ P_{1*,1}(t)(\lambda_1 \delta t)(1 - \lambda_2 \delta t)(1 - \mu_1 \delta t)$$

$$P_{1,2*}(t + \delta t) = P_{1,2*}(t)(1 - \mu_2 \delta t)$$

$$+ P_{0,2}(t)(\lambda_1 \delta t)(1 - \lambda_2 \delta t)(1 - \mu_2 \delta t)$$

$$+ P_{1,1*}(t)(1 - \lambda_1 \delta t)(\lambda_2 \delta t)(1 - \mu_2 \delta t)$$

$$P_{1*,2}(t+\delta t) = P_{1*,2}(t)(1-\mu_1\delta t)$$
$$+ P_{1*,1}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)(1-\mu_1\delta t)$$

$$P_{2,1*}(t+\delta t) = P_{2,1*}(t)(1-\mu_2\delta t)$$
$$+ P_{1,1*}(t)(\lambda_1\delta t)(1-\lambda_2\delta t)(1-\mu_2\delta t)$$

The steady state equations, including the requirement that all probabilities must sum to one, are as follows:

$$(\lambda_1+\lambda_2)P_{0,0} = \mu_1 P_{1,0} + \mu_2 P_{0,1} \tag{1}$$

$$(\lambda_1+\lambda_2+\mu_1)P_{1,0} = \lambda_1 P_{0,0} + \mu_1 P_{2,0} + \mu_2 P_{1,1*} \tag{2}$$

$$(\lambda_1+\lambda_2+\mu_2)P_{0,1} = \lambda_2 P_{0,0} + \mu_2 P_{0,2} + \mu_1 P_{1*,1} \tag{3}$$

$$(\lambda_1+\lambda_2+\mu_1)P_{2,0} = \lambda_1 P_{1,0} + \mu_1 P_{3,0} + \mu_2 P_{2,1*} \tag{4}$$

$$(\lambda_1+\lambda_2+\mu_2)P_{0,2} = \lambda_2 P_{0,1} + \mu_2 P_{0,3} + \mu_1 P_{1*,2} \tag{5}$$

$$(\lambda_1+\lambda_2+\mu_1)P_{1*,1} = \lambda_2 P_{1,0} + \mu_1\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)P_{2*,1} + \mu_2\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)P_{1,2*} \tag{6}$$

$$(\lambda_1+\lambda_2+\mu_2)P_{1,1*} = \lambda_1 P_{0,1} + \mu_2\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)P_{1,2*} + \mu_1\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)P_{2*,1} \tag{7}$$

$$\mu_1 P_{3,0} = \lambda_1 P_{2,0} \tag{8}$$

$$\mu_2 P_{0,3} = \lambda_2 P_{0,2} \tag{9}$$

$$\mu_1 P_{2*,1} = \lambda_2 P_{2,0} + \lambda_1 P_{1*,1} \tag{10}$$

$$\mu_2 P_{1,2*} = \lambda_1 P_{0,2} + \lambda_2 P_{1,1*} \tag{11}$$

$$\mu_1 P_{1*,2} = \lambda_2 P_{1*,1} \tag{12}$$

$$\mu_2 P_{2,1*} = \lambda_1 P_{1,1*} \tag{13}$$

$$P_{0,0} + P_{1,0} + P_{0,1} + P_{2,0} + P_{0,2} + P_{1*,1} + P_{1,1*} + P_{3,0} + P_{0,3} + P_{2*,1} + P_{1,2*} + P_{1*,2} + P_{2,1*} = 1 \tag{14}$$

10.2.2 Results

These equations were solved using MAPLE but are omitted due to their considerable length. The sums $P_h = P_{1,0} + P_{2,0} + P_{1*,1} + P_{3,0} + P_{2*,1} + P_{1*,2}$ and $P_{nh} = P_{0,1} + P_{0,2} + P_{1,1*} + P_{0,3} + P_{2*,1} + P_{1*,2}$ give the total proportion of time that hip and non-hip patients occupy the theatre respectively, while $P_{0,0}\ (= P_0)$ gives the proportion of time that the theatre is empty. Clearly, $P_0 + P_h + P_{nh} = 1$.

It can be seen from the data (Table 10.1i) that empirical values are $P_{h_E} = 0.095$, $P_{nh_E} = 0.356$ and $P_{0_E} = 0.549$. The value for P_0 includes the time the theatre is closed as well as when it is open but not being used. This empirical value is taken for P_0 rather than the proportion of time that the system was empty since overall theatre usage is of primary interest.

The proportion of time that the theatre is occupied by hip, non-hip or no patients is a key output of this queuing model and Solver was used to find values for λ_2 and μ_2 such that the above proportions are achieved to as high a degree of accuracy as possible. It is not appropriate to acquire the values for λ_2 and μ_2 from the data. The arrival rate that would be calculated would not represent the true arrival rate but an artificial rate of the amount of arrivals that can be accommodated by the system. Unlike with hip fracture patients, whose arrival rate to the theatre cannot be amended, there is some scope to move these patients to other theatres. The empirical mean clinical time for non-hip patients is 85.5 minutes, giving a service rate of 0.702 per hour, but since the 'real' arrival rate is not used, it is also not appropriate to fix μ_2 to this value.

λ_1 and μ_1 were fixed according the data. λ_1 was amended slightly to 1.47 (per day); the previous value of 1.27 for hip arrival rate was calculated from the ward data, while 1.47 was calculated from theatre data and is more appropriate to use here since other theatre data is used to consolidate values obtained from this analysis. μ_1 was determined using the mean service time given by $h_c(t)$ in Section 9.3.4; the mean service time was 99.0 minutes giving a service rate of $\mu_1 = 0.606$ per hour. The aim was to find arrival and service rates for non-hip patients such that the squared discrepancy between the empirical and analytical proportions

were minimised. Values outputted were $\lambda_2 = 5.85$ (per day) and $\mu_2 = 0.629$ (per hour), giving analytical proportions of $P_h = 0.095$, $P_{nh} = 0.363$ and $P_0 = 0.542$.

System probabilities, excluding when the system is empty, are shown in order of likelihood Figure 10.2.2i. The most likely system state (after P_0) is one non-hip patient in theatre with no patients in the queue, while the least likely is three hip patients in the system, one in theatre and two in the queue.
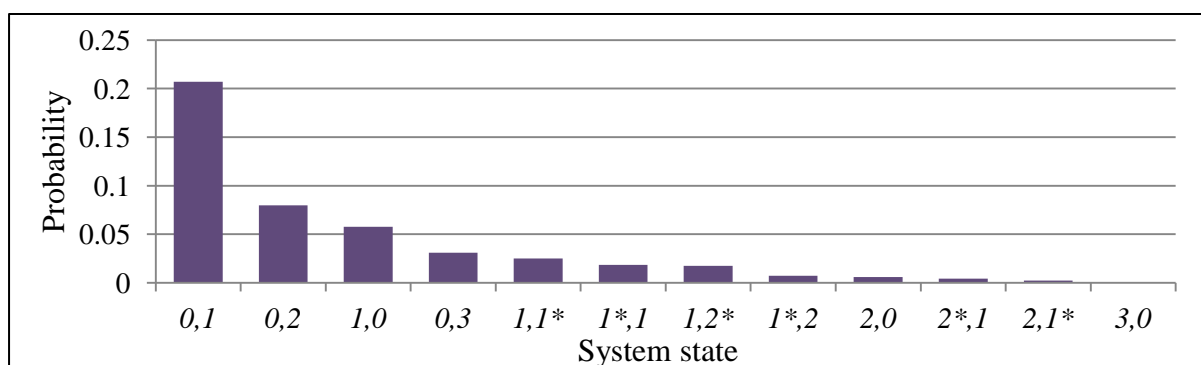

Figure 10.2.2i: System state probabilities for $M(\lambda_1, \lambda_2) \mid M(\mu_1, \mu_2) \mid 1 \mid 3_{sys} \mid$ FIFO model

The analytical probabilities of one, two or three patients in the system are given by

$$P_1 = P_{1,0} + P_{0,1} = 0.265,$$
$$P_2 = P_{2,0} + P_{0,2} + P_{1^*,1} + P_{1,1^*} = 0.129$$
$$\text{and} \quad P_3 = P_{3,0} + P_{0,3} + P_{2^*,1} + P_{1,2^*} + P_{1^*,2} + P_{2,1^*} = 0.063.$$

The theoretical mean number in the system is given by $L = \sum_{n=0}^{3} n\,P_n = 0.713$.

The probabilities of zero, one or two patients in the queue are thus given by $Q_0 = P_0 + P_1 = 0.807$, $Q_1 = P_2 = 0.129$ and $Q_2 = P_3 = 0.063$ respectively. The theoretical mean number in the queue is equal to $L_q = \sum_{n=0}^{2} n\,Q_n = 0.256$. If there is one patient in the queue, it is a hip patient 24.1% of the time $\left(\dfrac{P_{2,0} + P_{1,1^*}}{P_2} \right)$. If there are two patients in the queue, 4.8% of the time they are both hip patients, 60.6% of the time they are both non-hip

patients and 34.5% of the time there is one of each patient type; calculated via $\dfrac{P_{3,0} + P_{2,1^*}}{P_3}$,

$\dfrac{P_{0,3} + P_{1^*,2}}{P_3}$ and $\dfrac{P_{2^*,1} + P_{1,2^*}}{P_3}$ respectively.

Keeping λ_2 and μ_2 fixed, the effect of changing λ_1 and μ_1 on L is shown in Figure 10.2.2ii. λ_1 was varied between the values of one and three per day, translating to 0.000694 to 0.002083 per minute as displayed on the graph. Service time was varied between 50 and 400 minutes, translating to changes in μ_1 from 0.02 to 0.0025, as displayed. If the hip arrival rate is also fixed, then $L > 1$ once hip service time exceeds 261 minutes, while fixing hip service time instead means that $L > 1$ once hip arrival rate exceeds 3.93 per day.
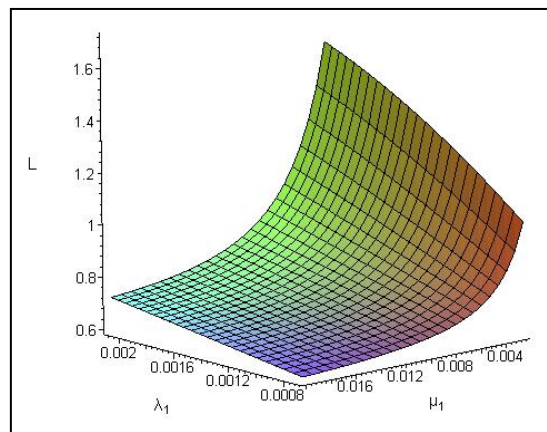

Figure 10.2.2ii: Impact on L as λ_1 and μ_1 are varied

10.2.3 Conclusions

While this model has proved fairly straightforward to solve and investigate, some of the results are not particularly accurate when comparing to the data. The analytical values for mean number in the system and the queue are both overstated. It has been previously shown that clinical time did not follow the monotonically-decreasing shape of the Negative Exponential distribution and the violation of this assumption will inevitably lead to some discrepancies between the model and the true system. Despite this, theoretical results were promising in that they were not wildly different to the data and thus this model gives a sound basis from which to develop.

10.3   Erlang model

It was shown in the previous section that assuming a Negative Exponential service time led to some results which did not accord very well with the data. Previous work (see Sections 8.2.2 and 8.2.3) has additionally shown that a Gamma or Lognormal distribution accommodates operation time more appropriately, while a Gamma distribution can be used to model anaesthetic time. In Section 9.3.4, a general distribution was used to represent clinical time (anaesthetic plus operation time) for hip patients. This was the convolution of Erlang-*4* and Erlang-*5* distributions.

The aim now is to find two distributions to model total clinical time, one for hip patients and one for non-hip patients.

Using Solver, it was found that clinical time for hip patients could in fact be modelled by an Erlang distribution with parameters $k_1 = 9$ and $\mu_1 = 0.604$ per hour so that the mean service rate per phase is $9\mu_1$ and the mean service time is 99.4 minutes, standard deviation 33.1 minutes. These results compare favourably with empirical values of 103.9 and 41.2 minutes respectively. The fit is displayed in Figure 10.3i.
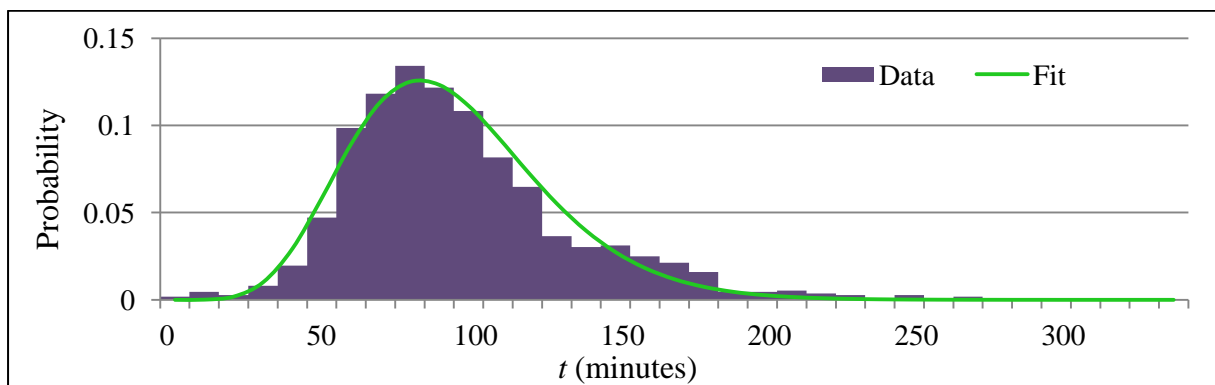

Figure 10.3i: Erlang fit for clinical time for all hip surgeries

Clinical time for all non-hip operations was also found to fit an Erlang distribution, as displayed in Figure 10.3ii, with parameters $k_2 = 3$ and $\mu_2 = 0.702$ (per hour). This results in mean service time of 85.5 minutes, standard deviation 49.4, comparing closely with respective empirical values of 85.5 and 54.7 minutes.
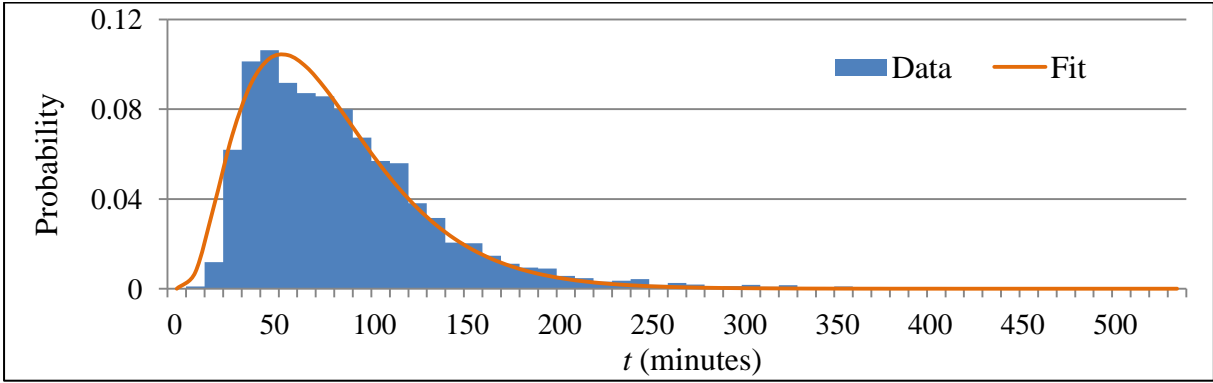
Figure 10.3ii: Erlang fit for clinical time for all non-hip surgeries

Arrival rates are still assumed to be λ_1 for hip patients and λ_2 for non-hip patients. Using standard notation, a general $M(\lambda_1, \lambda_2)|\ E_k(k_1, \mu_1; k_2, \mu_2)|\ 1\ |\ l_{sys}\ |$ FIFO system has been described for a system limit of l_{sys} patients. This system is represented in Figure 10.3iii below for $k_1 = 9$ and $k_2 = 3$.
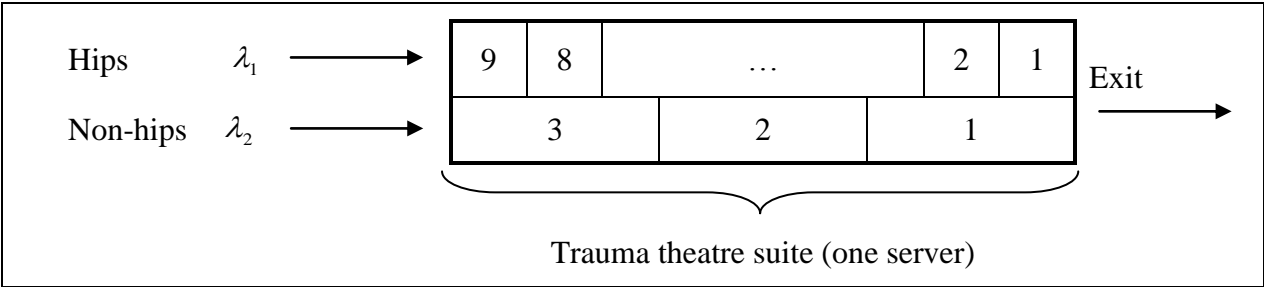

Figure 10.3iii: Pictorial representation of Erlang model

The total number of system states when there are two Erlang distributions with parameters of k_1 and k_2, for a limit of l_{sys} customers in the system, is given by

$$1 + (k_1 + k_2)\sum_{m=1}^{l_{sys}} m = 1 + \frac{1}{2}(k_1 + k_2)(l_{sys}^2 + l_{sys}).$$

10.3.1 Formulation

For this system, if $l_{sys} = 2$ then there are $1 + \frac{1}{2}(9+3)(2^2+2) = 37$ total system states. Setting $l_{sys} = 3$ results in 36 extra system states, 73 in total, but the data showed that the probability of three or more patients in the system was less than 0.1%, so a more stringent limit of $l_{sys} = 2$ used in this case should have negligible effect. This is investigated further in Section 10.3.3.

Let $P_{h,n,k}(t)$ be the probability of h hip patients and n non-hip patients in the system at time t, where the patient in service (where relevant) is in phase k. The steady-state probability of $P_{h,n,k}(t)$ is $P_{h,n,k}$. An asterisk is again used to denote which patient is in service, where relevant.

The differential-difference equations when $l_{sys} = 2$ are as follows:

$$P_{0,0}(t+\delta t) = P_{0,0}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)$$
$$+ P_{1,0,1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(9\mu_1\delta t)$$
$$+ P_{0,1,1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(3\mu_2\delta t)$$

$$i = 1,..8; \quad P_{1,0,i}(t+\delta t) = P_{1,0,i}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-9\mu_1\delta t)$$
$$+ P_{1,0,i+1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(9\mu_1\delta t)$$

$$P_{1,0,9}(t+\delta t) = P_{1,0,9}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-9\mu_1\delta t)$$
$$+ P_{0,0}(t)(\lambda_1\delta t)(1-\lambda_2\delta t)$$
$$+ P_{2,0,1}(t)(9\mu_1\delta t)$$
$$+ P_{1,1^*,1}(t)(3\mu_2\delta t)$$

$$j = 1,2; \quad P_{0,1,j}(t+\delta t) = P_{0,1,j}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-3\mu_2\delta t)$$
$$+ P_{0,1,j+1}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(3\mu_2\delta t)$$

$$P_{0,1,3}(t+\delta t) = P_{0,1,3}(t)(1-\lambda_1\delta t)(1-\lambda_2\delta t)(1-3\mu_2\delta t)$$
$$+ P_{0,0}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)$$
$$+ P_{0,2,1}(t)(3\mu_2\delta t)$$
$$+ P_{1^*,1,1}(t)(9\mu_1\delta t)$$

$i = 1,..8;$
$$P_{2,0,i}(t+\delta t) = P_{2,0,i}(t)(1-9\mu_1\delta t)$$
$$+ P_{2,0,i+1}(t)(9\mu_1\delta t)$$
$$+ P_{1,0,i}(t)(\lambda_1\delta t)(1-\lambda_2\delta t)(1-9\mu_1\delta t)$$

$$P_{2,0,9}(t+\delta t) = P_{2,0,9}(t)(1-9\mu_1\delta t)$$
$$+ P_{1,0,9}(t)(\lambda_1\delta t)(1-\lambda_2\delta t)(1-9\mu_1\delta t)$$

$j = 1, 2;$
$$P_{0,2,j}(t+\delta t) = P_{0,2,j}(t)(1-3\mu_2\delta t)$$
$$+ P_{0,2,j+1}(t)(3\mu_2\delta t)$$
$$+ P_{0,1,j}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)(1-3\mu_2\delta t)$$

$$P_{0,2,3}(t+\delta t) = P_{0,2,3}(t)(1-3\mu_2\delta t)$$
$$+ P_{0,1,3}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)(1-3\mu_2\delta t)$$

$i = 1,..8;$
$$P_{1^*,1,i}(t+\delta t) = P_{1^*,1,i}(t)(1-9\mu_1\delta t)$$
$$+ P_{1^*,1,i+1}(t)(9\mu_1\delta t)$$
$$+ P_{1,0,i}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)(1-9\mu_1\delta t)$$

$$P_{1^*,1,9}(t+\delta t) = P_{1^*,1,9}(t)(1-9\mu_1\delta t)$$
$$+ P_{1,0,9}(t)(1-\lambda_1\delta t)(\lambda_2\delta t)(1-9\mu_1\delta t)$$

$$j = 1, 2; \qquad P_{1,1^*,j}\left(t+\delta t\right) = P_{1,1^*,j}\left(t\right)\left(1-3\mu_2\delta t\right)$$

$$+ P_{1,1^*,j+1}\left(t\right)\left(3\mu_2\delta t\right)$$

$$+ P_{0,1,j}\left(t\right)\left(\lambda_1\delta t\right)\left(1-\lambda_2\delta t\right)\left(1-3\mu_2\delta t\right)$$

$$P_{1,1^*,3}\left(t+\delta t\right) = P_{1,1^*,3}\left(t\right)\left(1-3\mu_2\delta t\right)$$

$$+ P_{0,1,3}\left(t\right)\left(\lambda_1\delta t\right)\left(1-\lambda_2\delta t\right)\left(1-3\mu_2\delta t\right)$$

The steady-state equations are:

$$\left(\lambda_1+\lambda_2\right)P_{0,0} = 9\mu_1 P_{1,0,1} + 3\mu_2 P_{0,1,1} \tag{1}$$

$$\left(\lambda_1+\lambda_2+9\mu_1\right)P_{1,0,i} = 9\mu_1 P_{1,0,i+1} \quad i=1,..,8. \tag{2)-(9}$$

$$\left(\lambda_1+\lambda_2+9\mu_1\right)P_{1,0,9} = \lambda_1 P_{0,0} + 9\mu_1 P_{2,0,1} + 3\mu_2 P_{1,1^*,1} \tag{10}$$

$$\left(\lambda_1+\lambda_2+3\mu_2\right)P_{0,1,j} = 3\mu_2 P_{0,1,j+1} \quad j=1,2. \tag{11)-(12}$$

$$\left(\lambda_1+\lambda_2+3\mu_2\right)P_{0,1,3} = \lambda_2 P_{0,0} + 9\mu_1 P_{1^*,1,1} + 3\mu_2 P_{0,2,1} \tag{13}$$

$$9\mu_1 P_{2,0,i} = 9\mu_1 P_{2,0,i+1} + \lambda_1 P_{1,0,i} \quad i=1,..,8. \tag{14)-(21}$$

$$9\mu_1 P_{2,0,9} = \lambda_1 P_{1,0,9} \tag{22}$$

$$3\mu_2 P_{0,2,j} = 3\mu_2 P_{0,2,j+1} + \lambda_2 P_{0,1,j} \quad j=1,2. \tag{23)-(24}$$

$$3\mu_2 P_{0,2,3} = \lambda_2 P_{0,1,3} \tag{25}$$

$$9\mu_1 P_{1^*,1,i} = 9\mu_1 P_{1^*,1,i+1} + \lambda_2 P_{1,0,i} \quad i=1,..,8. \tag{26)-(33}$$

$$9\mu_1 P_{1*,1,9} = \lambda_2 P_{1,0,9} \tag{34}$$

$$3\mu_2 P_{1,1*,j} = 3\mu_2 P_{1,1*,j+1} + \lambda_1 P_{0,1,j} \quad j = 1, 2. \tag{35)-(36}$$

$$3\mu_2 P_{1,1*,3} = \lambda_1 P_{0,1,3} \tag{37}$$

$$P_{0,0} + \sum_{i=1}^{9}\left(P_{1,0,i} + P_{2,0,i} + P_{1*,1,i}\right) + \sum_{j=1}^{3}\left(P_{0,1,j} + P_{0,2,j} + P_{1,1*,j}\right) = 1 \tag{38}$$

10.3.2 Results

There are therefore 37 unknowns and 38 equations, including the normalising equation. These equations were solved in MAPLE but the solutions again are omitted due to their considerable length.

The sums $P_h = \sum_{i=1}^{9}\left(P_{1,0,i} + P_{2,0,i} + P_{1*,1,i}\right)$ and $P_{nh} = \sum_{j=1}^{3}\left(P_{0,1,j} + P_{0,2,j} + P_{1,1*,j}\right)$ give the total proportion of time that hip and non-hip patients occupy the theatre respectively and were used to solve the equations for λ_2 and μ_2, as previously explained. λ_1 and μ_1 were set equal to 1.47 (per day) and 0.604 (per hour). Values outputted were $\lambda_2 = 6.16$ (per day) and $\mu_2 = 0.629$ (per hour), with $P_h = 0.089$, $P_{nh} = 0.361$ and $P_0 = 0.550$. Results by system state, excluding when the server is idle, are displayed in Figure 10.3.2i.
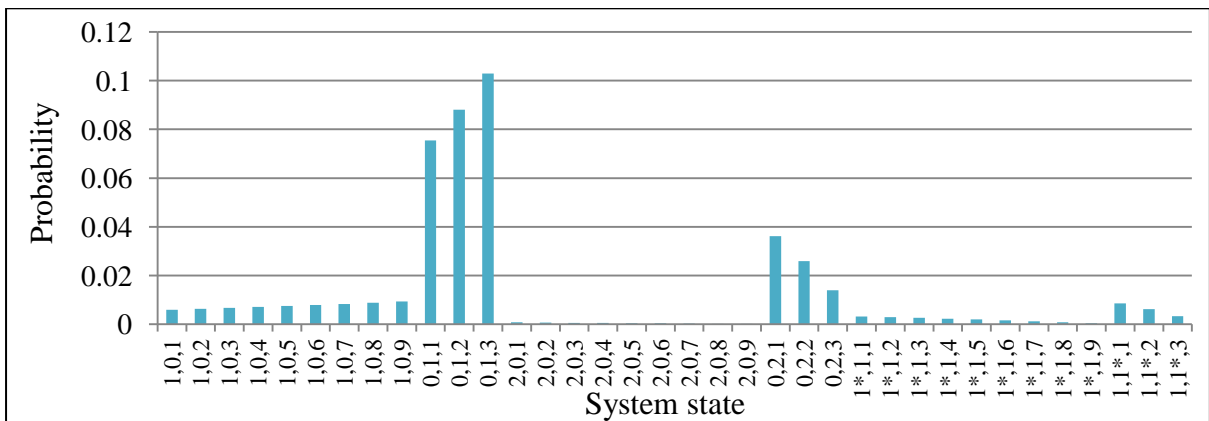

Figure 10.3.2i: Steady-state probabilities for $M\left(\lambda_1, \lambda_2\right) | E_k\left(9, \mu_1; 3, \mu_2\right) | 1 | 2_{sys}$ | FIFO model

The most likely system state, again excluding when the server is idle, is *0,1,3* – one non-hip patient in phase three of service. The least likely system state is *2,0,9* – one hip patient in phase nine of service and another hip patient in the queue.

The probability of one patient in the system (either type) is given by

$$P_1 = \sum_{i=1}^{9} P_{1,0,i} + \sum_{j=1}^{3} P_{0,1,j} = 0.334,$$

while the probability of two patients in the system is given by

$$P_2 = \sum_{i=1}^{9} \left(P_{2,0,i} + P_{1^*,1,i} \right) + \sum_{j=1}^{3} \left(P_{0,2,j} + P_{1,1^*,j} \right) = 0.116.$$

The data showed that there was one patient in the system 38.6% of the time, and two patients 9.3% of the time, so these results compare reasonably favourably. The mean number in the system was 0.572 (S.D. 0.657), compared with $L = \sum_{n=0}^{3} n P_n = 0.566$ (S.D. 0.691), indicating a high overall level of compatibility.

The probability of no patients in the queue is thus given by $Q_0 = P_0 + P_1 = 0.884$, while the probability of one patient in the queue is equal to $Q_1 = P_2 = 0.116$. (Two or more patients in the queue is not possible.) These are very close to the empirical probabilities of 0.871 and 0.129. The analytical mean number in the queue is given by $L_q = \sum_{n=0}^{2} n Q_n = 0.116$ (S.D. 0.109), compared with an empirical mean of 0.130 (S.D. 0.338).

The probability of a hip patient in the queue is given by

$$P_{hq} = \sum_{i=1}^{9} P_{2,0,i} + \sum_{j=1}^{3} P_{1,1^*,j} = 0.022,$$

while the probability of a non-hip patient in the queue is given by

$$P_{nhq} = \sum_{i=1}^{9} P_{1^*,1,i} + \sum_{j=1}^{3} P_{0,2,j} = 0.093.$$

Incidentally, since there is a limit of two patients in the system (and so one in the queue), the values of P_{hq} and P_{nhq} also give the mean number of hip and non-hip patients in the queue

and compare well to data values of 0.035 and 0.094. Given that there is a patient in the queue, the probability of them being a hip patient or a non-hip patient are thus $\dfrac{0.022}{P_2} = 0.193$

and $\dfrac{0.093}{P_2} = 0.807$ respectively.

It is concluded that this queuing model represents the trauma theatre with sufficient level of compatibility and relevance. The model is now used to investigate a number of scenarios in order to assess the impact on a variety of system measures.

(a) What-if scenario: Change in arrival rates

λ_1 and λ_2 were altered both independently and simultaneously in order to assess the impact of changing them on a number of factors. The rates were increased as a percentage as opposed to a crude increase in the number.

The effect of doing this on P_h, P_{nh} and P_0 are presented in Figure 10.3.2ii. Note that some results are almost identical in this graph. A change of up to ±20% in the hip arrival rate has little impact on the system (solid lines), while a much greater effect is seen when non-hip or both arrival rates are changed. A simultaneous increase of 14% or more each to both λ_1 and λ_2 means that $P_0 < P_h + P_{nh}$ and the theatre is more likely to be busy than empty.

The impact upon the number in the system is displayed in Figure 10.3.2iii and Table 10.3.2iv. The largest increase of 20% to both arrival rates sees P_1 increase from 0.334 to 0.365 and P_2 increase from 0.116 to 0.153, resulting in an increase in the mean number in the system from 0.57 to 0.67.
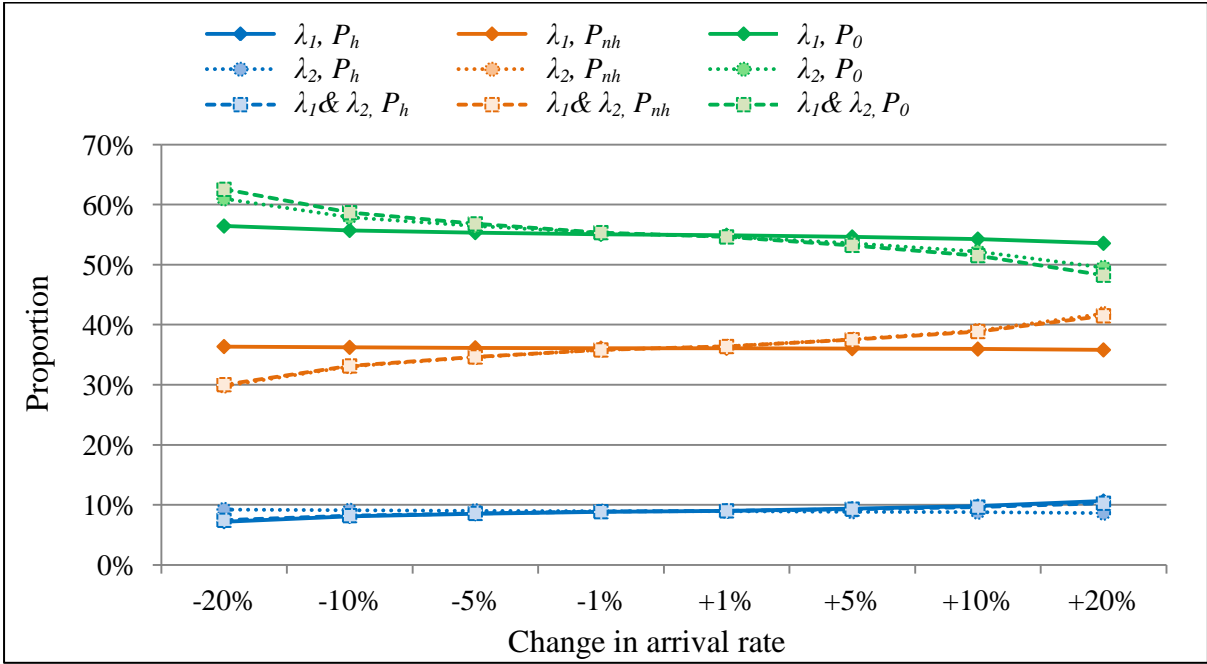
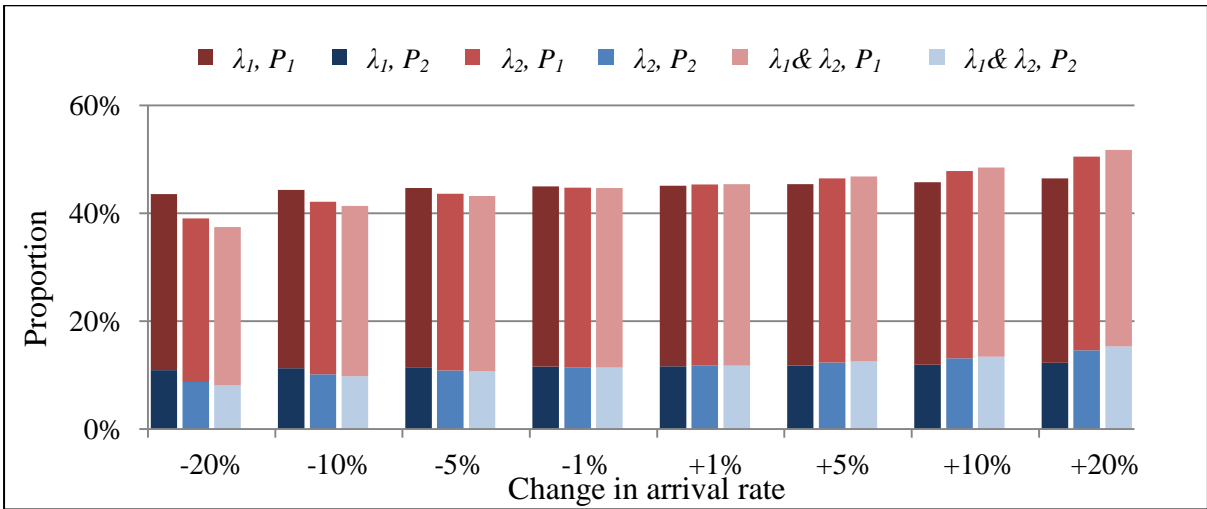**Figure 10.3.2ii:** Effect of changes to the arrival rate on who is in service

Figure 10.3.2iii: Effect of changes to the arrival rate on number in system

Table 10.3.2iv: Effect of changes to the arrival rate on mean number in system

i	Percentage change in λ_i							
	-20%	**-10%**	**-5%**	**-1%**	**+1%**	**+5%**	**+10%**	**+20%**
1	0.54	0.56	0.56	0.56	0.57	0.57	0.58	0.59
2	0.48	0.52	0.54	0.56	0.57	0.59	0.61	0.65
1&2	0.46	0.51	0.54	0.56	0.57	0.59	0.62	0.67

Next the composition of the queue is considered. The maximum number allowed in the queue is one; Table 10.3.2v displays the percentage of time that the patient in the queue is a hip patient, assuming that a queue exists. If simultaneous and equivalent changes are made to the arrival rates, then no change is seen. Increasing the hip arrival rate by 20% means that, when there is patient waiting, 22.3% of the time it is a hip patient, compared with 19.3% of the time when no changes are made.

Table 10.3.2v: Percentage of time that the patient in the queue is a hip patient

i	Percentage change in λ_i							
	-20%	-10%	-5%	-1%	+1%	+5%	+10%	+20%
1	16.0	17.7	18.5	19.1	19.4	20.0	20.8	22.3
2	23.0	21.0	20.1	19.4	19.1	18.5	17.8	16.6
1&2	19.3	19.3	19.3	19.3	19.3	19.3	19.3	19.3

Finally, an increase in λ_1 is considered, with a simultaneous decrease in λ_2, so an increase in the arrival rate of hip patients is compensated for by a decrease in the arrival rate of other patients, by moving them to other trauma theatres, for example. λ_1 is increased by 5%, 10% and 20%, λ_2 is decreased by 5%, 10% and 20%. Figure 10.3.2vi shows the impact upon the mean number in the system, the mean number in the queue, the probability of a hip patient in the queue and the probability of a non-hip patient in the queue, for each of these combinations.

A 20% increase to λ_1 and 5% decrease to λ_2 has the least effect on the mean number in the system/queue. This suggests that to keep the overall theatre utilisation at the current level, an increase of 20% more hip patients would mean that 5% of non-hip surgeries would need to be moved to another theatre. There would be a shift seen towards a greater proportion of time that the patient in the queue is a hip patient. The greatest effect is seen by a 5% increase to hip arrival rate coupled with a 20% decrease in non-hip arrival rate, where L and L_q would reduce by 14.7% and 24% respectively. This is the only scenario considered where there is a decrease in the probability of a hip patient in the queue.
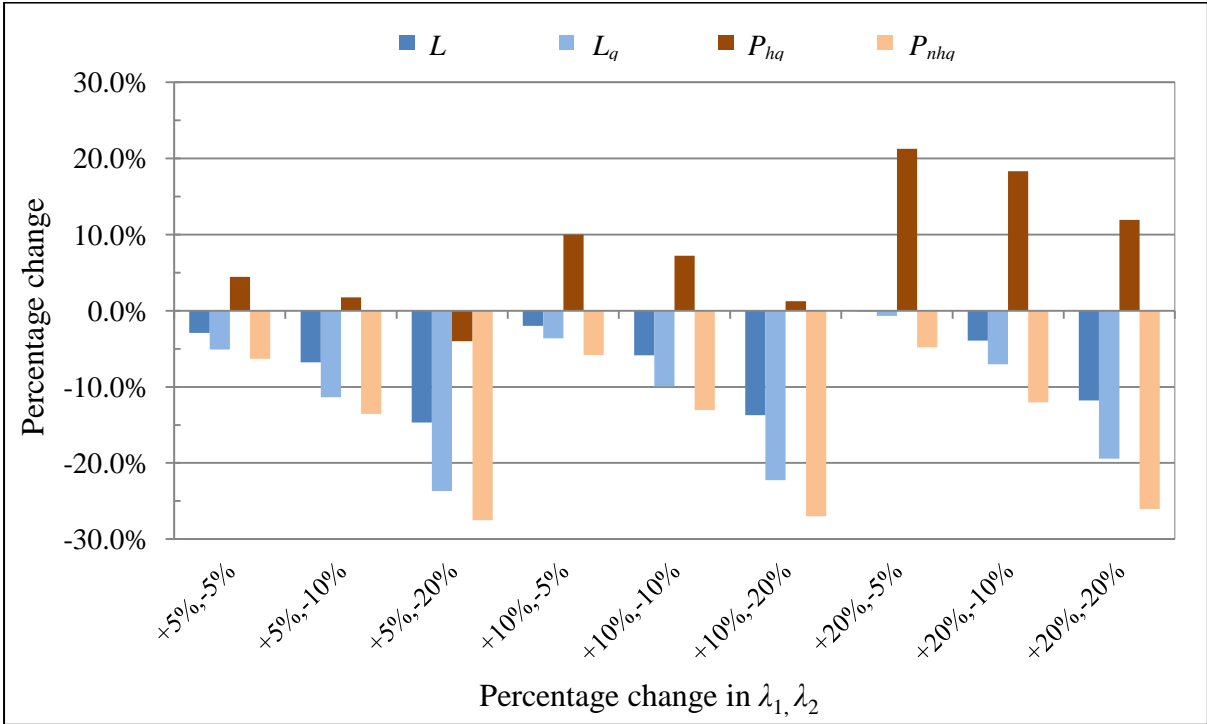
Figure 10.3.2vi: Impact on various measures due to an increase in λ_1 and decrease in λ_2

(b) What-if scenario: Change in hip clinical time

Changes to μ_1 were made to investigate the impact of a change in clinical time for hip patients. Recall that the baseline model had a mean service time for hip patients of 99.4 minutes. The number of hours per day that the theatre is used for each patient type, or when it is empty, is shown in Figure 10.3.2vii.
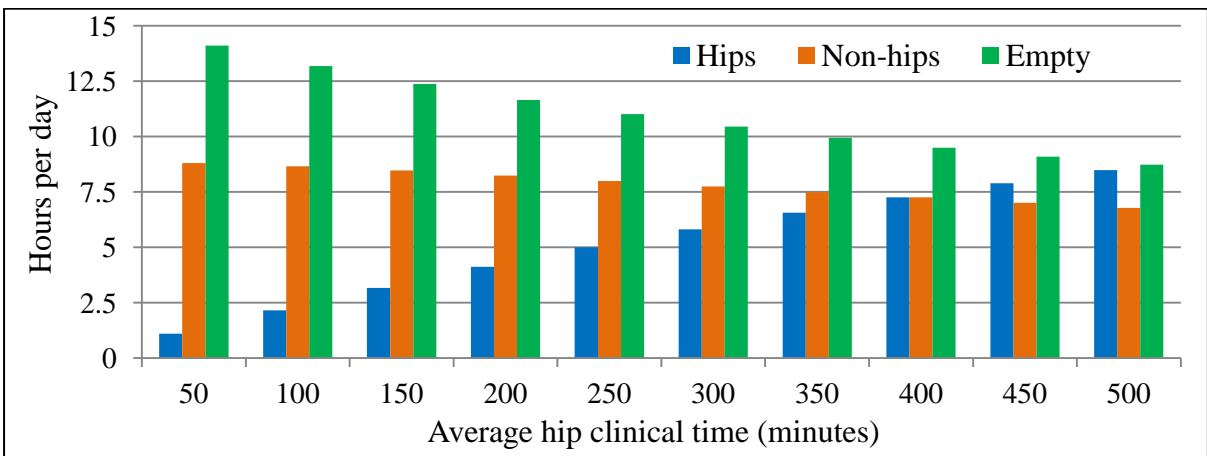

Figure 10.3.2vii: Effect of changes to hip clinical time on theatre usage

Hip patients would spend longer than non-hip patients in theatre per day should their clinical time exceed 399 minutes, while total busy time exceeds empty time once hip clinical time exceeds 174 minutes. The mean number of hip patients in the system does not exceed one until hip clinical time reaches 580 minutes.

(c) What-if scenario: Change in turnover time

Finally, turnover time was incorporated into the model, where it had been previously assumed to be zero. This was done by adding to the service time so that, for example, for a turnover time of t following a hip operation, μ_1 becomes $\dfrac{1}{99.4 + t}$ per minute, since average service time is 99.4 minutes. μ_1 and μ_2 were altered both independently and simultaneously and the impact on theatre usage is displayed in Figure 10.3.2viii. Again a lesser effect is seen when a change is applied only to hip patients due to the fewer number of them entering theatre. The mean anaesthetic room/theatre turnover, as required here, was 23 minutes. If this is incorporated into both service times, then P_0 decreases from 0.550 to 0.470, P_h increases to 0.105 and P_{nh} increases to 0.425. The mean number in the system increases by 22% to 0.690, and exceeds one when $t > 93$ for a simultaneous and equal change in service time to both patient types. Queue composition is unaffected but the mean number in the queue increases by 38% to 0.160.
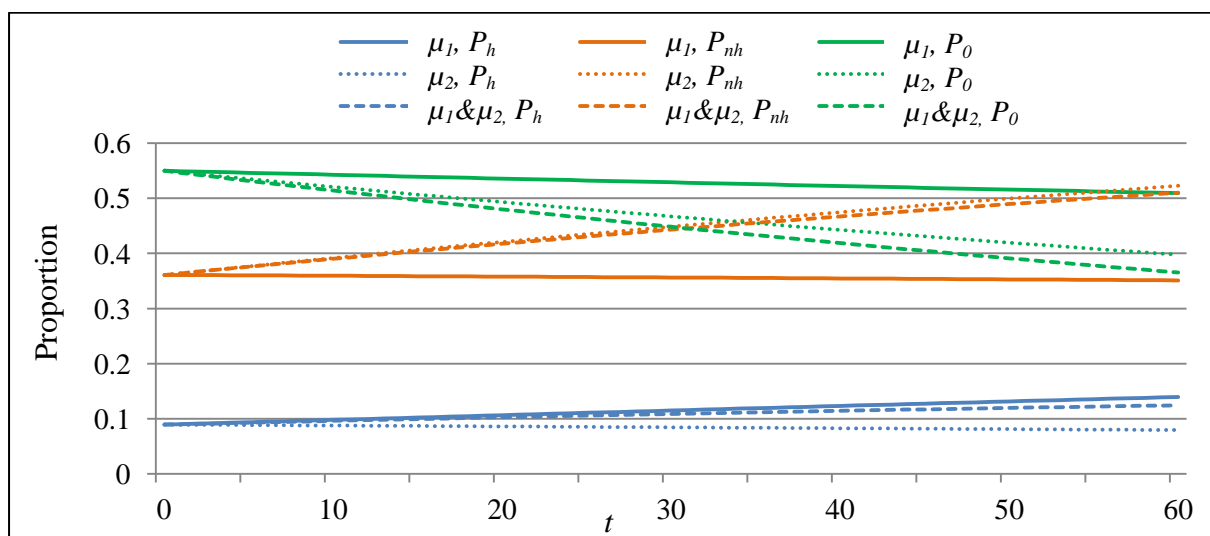

Figure 10.3.2viii: Effect of changes to turnover on theatre usage

10.3.3 Extension to $l_{sys} = 3$

Letting the system threshold $l_{sys} = 3$ means that the system can then be described by the notation $\mathrm{M}(\lambda_1, \lambda_2) \mid \mathrm{E}_k(k_1, \mu_1; k_2, \mu_2) \mid 1 \mid 3_{sys} \mid$ FIFO. Setting $k_1 = 9$, $k_2 = 3$ and $l_{sys} = 3$ results in a total of 73 system states.

Using the same notation as defined in Section 10.3.1, the 73 steady-state probabilities are:

$P_{0,0}$	$P_{1,0,1}$	$P_{0,1,1}$	$P_{2,0,1}$	$P_{0,2,1}$	$P_{1*,1,1}$	$P_{1,1*,1}$
	$P_{1,0,2}$	$P_{0,1,2}$	$P_{2,0,2}$	$P_{0,2,2}$	$P_{1*,1,2}$	$P_{1,1*,2}$
	$\vdots$	$P_{0,1,3}$	$\vdots$	$P_{0,2,3}$	$\vdots$	$P_{1,1*,3}$
	$P_{1,0,9}$		$P_{2,0,9}$		$P_{1*,1,9}$	

$P_{3,0,1}$	$P_{0,3,1}$	$P_{2*,1,1}$	$P_{2,1*,1}$	$P_{1*,2,1}$	$P_{1,2*,1}$
$P_{3,0,2}$	$P_{0,3,2}$	$P_{2*,1,2}$	$P_{2,1*,2}$	$P_{1*,2,2}$	$P_{1,2*,2}$
$\vdots$	$P_{0,3,3}$	$\vdots$	$P_{2,1*,3}$	$\vdots$	$P_{1,2*,3}$
$P_{3,0,9}$		$P_{2*,1,9}$		$P_{1*,2,9}$	

With the additional requirement that all probabilities must sum to one, formulation of the steady-state equations results in 74 equations and 73 unknowns. These equations are omitted here but were entered into MAPLE and subsequently solved. This resulted in very long algebraic expressions for the probabilities with any manipulation proving to be difficult.

However, since each expression involved $P_{0,0}$, by entering the values of λ_1, λ_2, μ_1 and μ_2 that had been previously found, a simple expression of the form $w_i P_{0,0}$ could be found for each of the other 72 steady-state probabilities, where w_i is a weighting such that $0 \le w_i \le 1$, $i = 1,\ldots,72$. $P_{0,0}$ was then found by rearranging $P_{0,0} + \sum_{i=1}^{72} w_i P_{0,0} = 1$ to obtain $P_{0,0} = \dfrac{1}{1 + \sum_{i=1}^{72} w_i}$, resulting in $P_{0,0} = 0.515$.

The equations have therefore not been solved for P_h and P_{nh} as previously, but some insight into the performance of this model is still possible despite this.

A summary of values and results obtained through this method are given in Table 10.3.3i. The probability of more than two patients in the system has been overestimated, as has the mean number in the system and the queue. The queue breakdown provides interesting results, showing that one non-hip patient in the queue is in fact more likely than all other queuing compositions combined (58% of the time a queue had formed). There would be two hip patients waiting only 0.8% of the time that a queue had formed.

Table 10.3.3i: Summary values for $\text{M}\left(\lambda_1,\lambda_2\right)|\,\text{E}_k\left(9,\mu_1;3,\mu_2\right)|\,1\,|\,3_{sys}\,|$ FIFO system

Measure	Value	Result
In service	Hip	10.0%
	Non-hip	38.5%
Number in system	0	51.5%
	1	31.3%
	2	12.1%
	3	5.1%
	Mean	0.707
	S.D.	0.868
Number in queue	0	82.8%
	1	12.1%
	2	5.1%
	Mean	0.223
	S.D.	0.524
Queue composition	1 hip	2.1%
	1 non-hip	10.0%
	2 hips	0.1%
	2 non-hips	4.0%
	1 hip, 1 non-hip	1.0%

Due to the lack of workability of the steady-state equations, coupled with the fact that the data showed that three (or more) patients are present in the system less than 0.1% of the time, this model is not pursued further. It has however provided an interesting academic exercise.

10.3.4 Conclusions

The first model presented in this section proved to represent the trauma theatre suite with a good level of accuracy. A limit of two patients in the system was imposed but on inspection of the data it was deemed that this would have little to no effect on model outcomes. It was found that a queuing model with two arrival sources, each with their own service time as represented by different Erlang distributions, provides a suitable mathematical representation of the theatre. The limit was then increased to three and a second model formulated, inputting results from the first model to obtain some summary results.

10.4 Vacation queuing models

10.4.1 Overview

The field of queuing systems with vacations is a well-researched area and thorough surveys of queues with vacations have been reported by several authors (Doshi 1986, Ke et al. 2010, Takagi 1991, Tian and Zhang 2006).

Consider the trauma theatre as a queuing system with vacations. If the theatre suite is occupied by a hip fracture patient, then the server (the theatre) is working and is not on vacation, but when the theatre is used for other surgery types or is closed, the server can be considered to be on vacation.

Several vacation policies have been developed and a brief overview is now presented. It would be infeasible to include a full literature review on this topic, so only the key themes have been discussed here.

In a single vacation model, the server takes a vacation when it becomes idle at the end of a busy period. On returning from the vacation, service is either immediately resumed (if there is a customer waiting) or the server waits until a customer arrives (if there are no customers waiting). In a multiple vacation model, if a server returns from a vacation to find an empty system then they will keep on taking vacations until, on returning, they find a customer waiting for service. Working vacation models have also been developed where a server works at a different rate, instead of being completely idle, during the vacation period. A key feature of these models is that the server only takes a vacation when the system becomes empty; this is known as exhaustive service. Extensions of the above models to incorporate some control of the vacation period have been comprehensively researched.

The concept of N-policy was first introduced several decades ago (Yadin and Naor 1963). On returning from vacation, N-policy dictates that the server only resumes service if there are at least N (≥ 1) customers in the queue. N-policy was first studied for the M | G | 1 queuing system five years later (Heyman 1968) and has been since developed in several other studies (Artalejo 1998, Wang and Ke 2000). In particular this model has been extended to include two vacation types, long and short (Zhang et al. 1997). Specifying lower and upper thresholds, say L and U, then upon returning from a vacation the server takes a long vacation

if there are less than *L* customers waiting, a short vacation if there are between *L* and *U*-1 customers waiting, or no vacation if there are *U* or more customers waiting.

The concept of *T*-policy was also first introduced for the M | G | 1 system by Heyman (Heyman 1977). Again the server takes a vacation at the end of a busy period, where the length of the vacation is a fixed time of *T* units. Service is resumed if there is at least one customer waiting, otherwise the server takes another vacation of fixed time *T*. Tadj extended earlier results to obtain, amongst other results, the PGF of the number of customers in the system and the optimum value of *T* (Tadj 2003).

Other vacation policies not considered in greater detail include *D*-policy, also first introduced in the 1970s (Balachandran 1973), whereby the server is turned off at the end of a busy period and turned on when the cumulative amount of work which has arrived during the vacation reaches some pre-defined value of *D*. Thus in this model, the service times of the waiting customers are taken into account.

Results for a min(*N*, *T*) policy, or simply *NT*-policy, were first established several years after the individual policies were developed (Gakis et al. 1995), whereby server vacation is terminated if either *N* customers have arrived or *T* time units have elapsed since the end of the last busy period (or the end of the last *T* time units and at least one customer has arrived). Results have since been extended by several authors (Alfa and Li 2000, Hur et al. 2003).

Start-up times have also been considered, firstly by Minh (Minh 1988), where the server needs a warm-up period after a vacation before service can be resumed. This has also been extended to a closing time, firstly by Takagi (Takagi 1991), where the server needs some time for shutting down prior to starting a vacation and is therefore busy but is unable to serve customers.

The concept of unreliable servers was also introduced several decades ago and a good early overview of different approaches was given in 1963 (Avi-Itzhak and Naor 1963), where servers may break down at a random time, service is interrupted, and the server is repaired with repair time following a random variable. Research into vacation queuing models has primarily focussed on reliable servers, but results have been extended over recent years to include unreliable servers (Jain and Jain 2010, Li et al. 1997). Specifically, results including

start-up and closing times have been obtained for the *N*-policy (Ke 2003), the *T*-policy (Ke 2005) and the *NT*-policy (Ke 2006) queuing models.

Lastly, the concept of Bernoulli vacations is introduced. After completion of service, the server either goes on vacation with probability p $(0 \leq p \leq 1)$, or continues to serve the next customer, if there is one waiting, with probability $q = 1 - p$. Recent advances in this field include an unreliable server who is subject to Bernoulli vacations under *N*-policy, where arrivals complete two heterogeneous phases of service (Tadj et al. 2012).

Results from a selection of the aforementioned literature, plus others (Madan 1999, Mehdi 2002, Scholl and Kleinrock 1983), were used to find summary outputs for three different vacation queuing models.

10.4.2 Multiple vacation model

Firstly, a multiple vacation model is considered; if there are no customers (hip patients) in the queue then the server (theatre) has a vacation of period v, which is taken from an arbitrary distribution with first and second moments of $E[v]$ and $E[v^2]$ respectively, and which has a known Laplace transform. The server keeps taking vacations until it returns from a vacation to find a customer waiting. This scenario is considered in detail, as opposed to a single vacation model, as it is more appropriate for the trauma theatre.

The vacation time is taken as the general service time for all other operations, so that if there are no hip fracture patients waiting, the theatre is used for other operations. Since busy time for hip patients is of interest, a 24-hour theatre is assumed; vacations thus may in reality be used for closedown as well as other surgeries. It was shown in Section 10.3 that service time for non-hip operations could be modelled by an Erlang distribution with parameters $k_v = 3$ and $\mu_v = 0.0117$ (per minute) so that $\alpha_v = k_v \mu_v = 0.0351$, $E[v] = 85.5$ and $E[v^2] = 9740$.

Let $h_v(t)$ denote the PDF of vacation time, so that

$$L_s\{h_v(t)\} = \left(\frac{\alpha_v}{\alpha_v + s}\right)^3.$$

This can be extended to more than one additional non-hip surgery during the vacation, so that $h_v^j(t)$ represents the PDF of vacation (clinical) time for j surgeries. Recall that $h_c*(s)$ is the Laplace transform of the PDF of clinical time for hip surgeries, as defined in Section 9.3.4, and let $h_v^j*(s)$ be the Laplace transform of the PDF of vacation time for j non-hip surgeries, so that

$$h_v^j*(s) = L_s\left\{h_v^j(t)\right\} = \left(\frac{\alpha_v}{\alpha_v + s}\right)^{3j}$$

$$\text{and} \quad h_v^j(t) = \frac{\alpha_v(\alpha_v t)^{3j-1} e^{-\alpha_v t}}{(3j-1)!}, \quad t \geq 0.$$

As is usual, ρ is given by $\dfrac{\lambda}{\mu}$ and is the proportion of time that the server is busy serving customers. Note that 'customer' here refers exclusively to hip fracture patients; non-hip patients are no longer considered as customers of the system but instead their service time is incorporated as vacation time.

$E[v]$ thus becomes $E[jv] = jE[v]$ for j surgeries during a vacation and $E[v^2]$ is also amended accordingly; the first and second moments of vacation time for j surgeries are subsequently denoted respectively by $E[v_j]$ and $E[v_j^2]$.

If P_n represents the probability of n customers in the system just after a departure instant, then the PGF of the number of customers in the system just after a departure is given by $Q(z)$, where

$$Q(z) = \sum_{n=0}^{\infty} z^n P_n$$

$$= \frac{h_c*(\lambda - \lambda z)}{\lambda E[v_j]} \cdot \frac{(1-\rho)\left(1 - h_v^j*(\lambda - \lambda z)\right)}{h_c*(\lambda - \lambda z) - z} \quad \text{(Scholl and Kleinrock 1983)}$$

$$= \frac{\alpha_2^4 \alpha_3^5}{\left(\alpha_2 + \lambda(1-z)\right)^4 \left(\alpha_3 + \lambda(1-z)\right)^5 \lambda E[v_j]} \cdot \frac{(1-\rho)\left(1 - \left(\dfrac{\alpha_v}{\alpha_v + \lambda(1-z)}\right)^{3j}\right)}{\left(\dfrac{\alpha_2^4 \alpha_3^5}{\left(\alpha_2 + \lambda(1-z)\right)^4 \left(\alpha_3 + \lambda(1-z)\right)^5} - z\right)} \cdot$$

The probability of no customers waiting just after a departure point is hence given by

$$Q(0) = \frac{(1-\rho)(1-h_v^j*(\lambda))}{\lambda E[v_j]} = \frac{(1-\rho)\left(1-\left(\dfrac{\alpha_v}{\alpha_v+\lambda}\right)^{3j}\right)}{\lambda E[v_j]}.$$

This is equivalent to the probability that the system is idle and is less than the probability that the server is idle; the server may be on vacation when a customer arrives, thus the server is idle but the system is busy. Substituting values into this equation yields the results given in Table 10.4.2i. The probability of an idle server in this case is

$$1-\rho = 1 - \frac{\lambda}{\mu} = 1 - \frac{\left(1.27/24 \times 60\right)}{\left(1/99\right)} = 0.913.$$

Table 10.4.2i: Probability of an idle system for different j

j	1	2	3	4	5	10	15
$Q(0)$	0.868	0.837	0.808	0.779	0.752	0.635	0.542

The remaining probabilities are calculated via $P_n = \dfrac{Q^{(n)}(0)}{n!}$ (Casella and Berger 1990) and

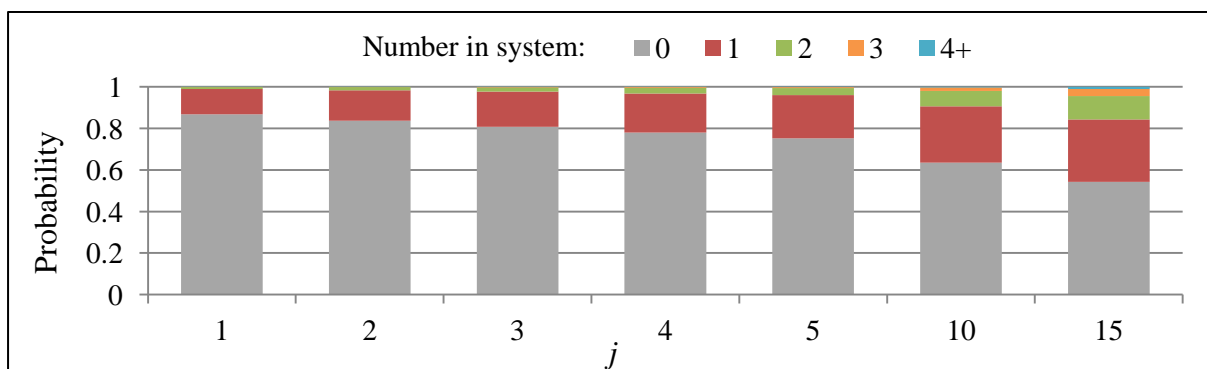are plotted for different j in Figure 10.4.2ii.

Figure 10.4.2ii: P_n for different j

Additionally letting $E[c]$ and $E[c^2]$ be the first and second moments of the PDF of clinical time for hip surgery (so that $E[c] = \frac{1}{\mu}$), then the expected number in the system, $\bar{n}$, is obtained by differentiating $Q(z)$ and evaluating the result at $z = 1$;

$$\bar{n} = \lambda E[c] + \frac{\lambda^2 E[c^2]}{2(1-\rho)} + \frac{\lambda E[v_j^2]}{2E[v_j]}.$$

Note that the first two terms of the sum represent the expected number in the system of a regular M | G | 1 queue, and so the proportion of $\bar{n}$ attributable to the inclusion of vacations in the model can be calculated. This is defined here as $\bar{n}_p$. Results displayed in Figure 10.4.2iii show an increase in $\bar{n}$ and $\bar{n}_p$ as j increases, but that the increase in $\bar{n}_p$ diminishes as j increases.

Note that the expected time in the system, $\bar{t}$, is given by

$$\bar{t} = \frac{\bar{n}}{\lambda} = E[c] + \frac{\lambda E[c^2]}{2(1-\rho)} + \frac{E[v_j^2]}{2E[v_j]},$$
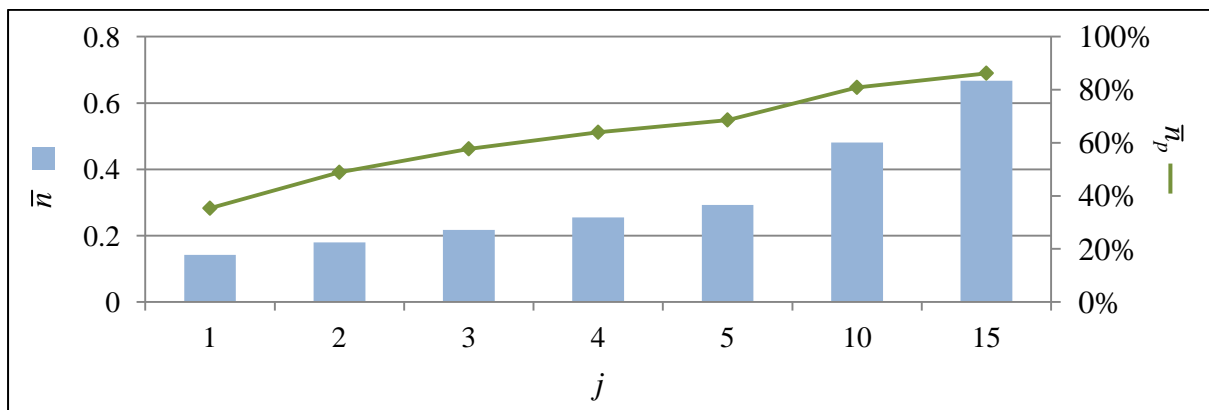
and therefore follows a similar trend to $\bar{n}$.

Figure 10.4.2iii: Expected number in system for different j

Finally, let $h_{t,VM}^{j}(t)$ be the PDF of the total time in the system for the multiple vacation model with j non-hip surgeries per vacation, with Laplace transform $h_{t,VM}^{j}*(s) = L_s\left\{h_{t,VM}^{j}(t)\right\}$.

Then

$$h_{t,VM}^{j}*(s) = \frac{s(1-\rho)h_c*(s)}{s-\lambda+\lambda h_c*(s)} \cdot \frac{1-h_v^{j}*(s)}{sE[v_j]} \quad \text{(Mehdi 2002)}$$

$$= \frac{s(1-\rho)\alpha_2^4\alpha_3^5}{(\alpha_2+s)^4(\alpha_3+s)^5\left(s-\lambda+\dfrac{\lambda\alpha_2^4\alpha_3^5}{(\alpha_2+s)^4(\alpha_3+s)^5}\right)} \cdot \frac{\left(1-\left(\dfrac{\alpha_v}{\alpha_v+s}\right)^{3j}\right)}{sE[v_j]}.$$

Note that the first term of the product represents the result for a regular M | G | 1 queue.

While a single vacation model is less appropriate to this situation, results are also presented if this model were employed. The PDF and associated Laplace transform are given by $h_{t,VS}^{j}(t)$ and $h_{t,VS}^{j}*(s) = L_s\left\{h_{t,VS}^{j}(s)\right\}$, and

$$h_{t,VS}^{j}*(s) = \frac{s(1-\rho)h_c*(s)}{s-\lambda+\lambda h_c*(s)} \cdot \frac{h_v^{j}*(\lambda)+\left(\dfrac{\lambda}{s}\right)\left(1-h_v^{j}*(s)\right)}{\lambda E[v_j]+h_v^{j}*(\lambda)} \quad \text{(Mehdi 2002)}$$

$$= \frac{s(1-\rho)\alpha_2^4\alpha_3^5}{(\alpha_2+s)^4(\alpha_3+s)^5\left(s-\lambda+\dfrac{\lambda\alpha_2^4\alpha_3^5}{(\alpha_2+s)^4(\alpha_3+s)^5}\right)} \cdot \frac{\left(\dfrac{\alpha_v}{\alpha_v+\lambda}\right)^{3j}+\left(\dfrac{\lambda}{s}\right)\left(1-\left(\dfrac{\alpha_v}{\alpha_v+s}\right)^{3j}\right)}{\lambda E[v_j]+\left(\dfrac{\alpha_v}{\alpha_v+\lambda}\right)^{3j}}.$$

A plot of PDFs of the total time in the system for an M | G | 1 queue with one, two and three vacations is displayed in Figure 10.4.2iv, for both single (VS, solid lines) and multiple (VM, dotted lines) vacation models, along with the PDF if there were no vacations. A longer and more varied system time is seen for the multiple vacation model in each case. Employing a single vacation model has far less effect on overall time in the system.
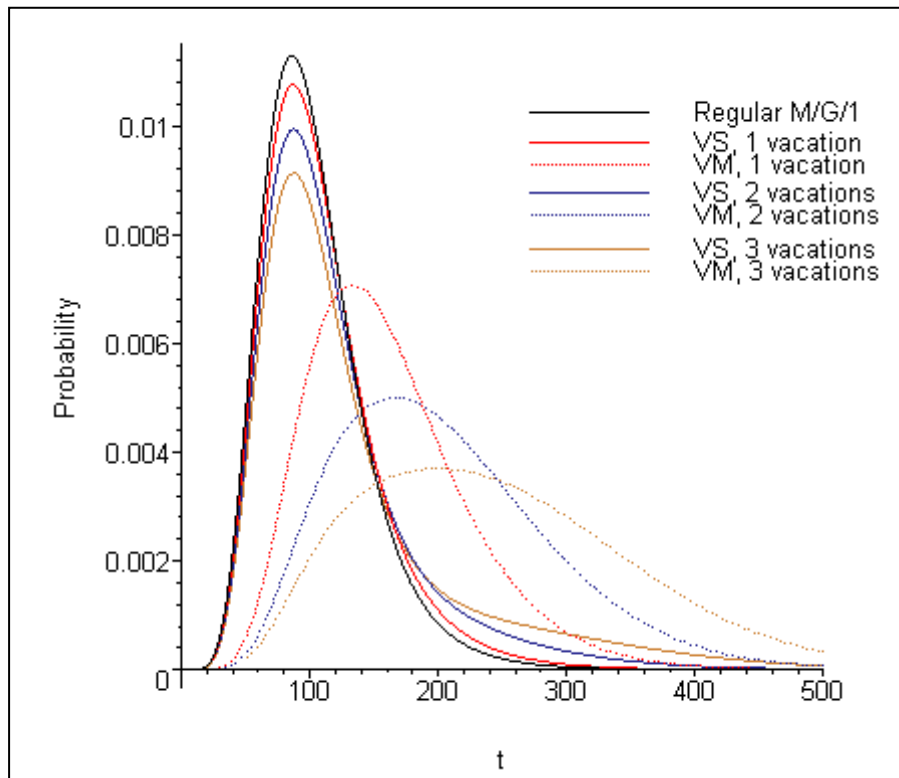
Figure 10.4.2iv: System time for different vacation policies

10.4.3 Vacation model with *N*-policy

Some results are now presented for *N*-policy. Under *N*-policy, the server is turned on whenever there are N (≥ 1) or more customers present. The server is only turned off (that is, it goes on vacation) when there are no customers present. After a vacation, the server does not resume serving customers until there are N customers waiting. Clearly, it would be unlikely that a healthcare provider would not tend to patients unless a certain number of those patients had presented themselves, and so less focus is given to this policy as it is less likely to be employed in the hospital.

The expected number of customers in the system is given by

$$\bar{n} = \frac{N-1}{2} + \rho + \frac{\lambda^2 E\left[c^2\right]}{2(1-\rho)},$$

and thus increases by 0.5 for each additional N (Wang and Ke 2000).

The PGF of the number of customers in the system is given in this case by

$$P(z) = \frac{(1-\rho)(1-z^N)h_c*(\lambda - \lambda z)}{N[h_c*(\lambda - \lambda z) - z]} \quad \text{(Wang and Ke 2000)}$$

$$= \frac{(1-\rho)(1-z^N)\alpha_2^4\alpha_3^5}{(\alpha_2 + \lambda(1-z))^4(\alpha_3 + \lambda(1-z))^5 N\left(\dfrac{\alpha_2^4\alpha_3^5}{(\alpha_2 + \lambda(1-z))^4(\alpha_3 + \lambda(1-z))^5} - z\right)}.$$

Manipulation of these formulae yields the results displayed in Table 10.4.3i and Figure 10.4.3ii. Note that the probability of no customers in the system is given by $P_0 = \dfrac{1-\rho}{N}$. It can be seen that $P_k = \dfrac{1}{N}$ for $0 < k < N$, and $P_k \approx 0$ for $k \geq N$.

Table 10.4.3i: Expected number in system for *N*-policy

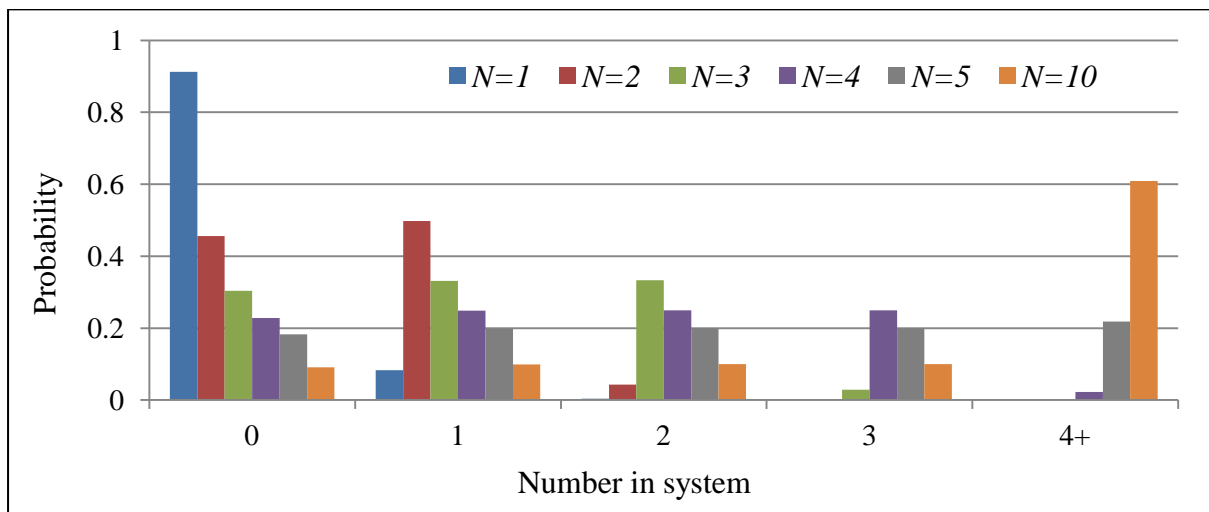| N	1	2	3	4	5	6	7	8	9	10
$\bar{n}$	0.09	0.59	1.09	1.59	2.09	2.59	3.09	3.59	4.09	4.59

Figure 10.4.3ii: P_n results for *N*-policy

10.4.4 Fixed length Bernoulli vacations

An alternative to the previous model is now considered. Instead, after every service, the server may take a vacation with probability p, $0 \le p \le 1$. The length of a vacation is a fixed time interval of d units. Thus, for every hip operation performed, the theatre becomes available afterwards for a fixed time of d units, $p\%$ of the time. In reality, hospital staff can determine not only p and d, but also on which of the $p\%$ of occasions the server takes a vacation.

In this scenario, the utilisation factor is not equal to $\dfrac{\lambda}{\mu}$. Defined in this case as ρ_{Bd}, it is calculated via $\rho_{Bd} = \dfrac{\lambda(1 + p\mu)}{\mu(1 - \lambda pd + p\lambda)}$ (which reduces to $\dfrac{\lambda}{\mu}$ when there are no vacations and $p = 0$). The relationship between p, d and ρ_{Bd} is displayed in Figure 10.4.4i. If $d > 1034.8$ (minutes) then $\rho_{Bd} > 1$ and the model becomes invalid (when p is at its maximum value of 1).
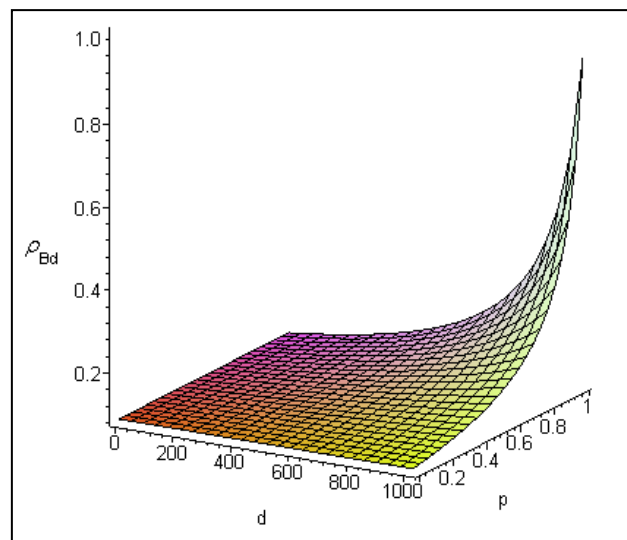

Figure 10.4.4i: Relationship between p, d and ρ_{Bd}

Let $P_q(z)$ and $P(z)$ be the probability generating functions of the number in the queue and the number in the system respectively, which are defined irrespective of whether the server is on vacation, then

$$P_q(z) = \frac{\left(\left(h_c*(\lambda-\lambda z)\right)(\lambda pz-\lambda p+1)-1\right)(1-\rho_{Bd})}{z-h_c*(\lambda-\lambda z)+h_c*(\lambda-\lambda z)p\left(1-e^{-\lambda d(1-z)}\right)} \quad \text{(Madan 1999)}$$

$$= \frac{\left(\dfrac{\alpha_2^4\alpha_3^5(\lambda pz-\lambda p+1)}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}-1\right)(1-\rho_{Bd})}{z-\dfrac{\alpha_2^4\alpha_3^5}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}+\dfrac{\alpha_2^4\alpha_3^5 p\left(1-e^{-\lambda d(1-z)}\right)}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}}$$

and $P(z) = (1-\rho_{Bd}) + zP_q(z)$

$$= (1-\rho_{Bd}) + \frac{z\left(\dfrac{\alpha_2^4\alpha_3^5(\lambda pz-\lambda p+1)}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}-1\right)(1-\rho_{Bd})}{z-\dfrac{\alpha_2^4\alpha_3^5}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}+\dfrac{\alpha_2^4\alpha_3^5 p\left(1-e^{-\lambda d(1-z)}\right)}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}}$$

$$= (1-\rho_{Bd})\cdot\left(\frac{\dfrac{\alpha_2^4\alpha_3^5}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}\left(z(\lambda pz-\lambda p+1)+p\left(1-e^{-\lambda d(1-z)}\right)-1\right)}{z-\dfrac{\alpha_2^4\alpha_3^5}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}+\dfrac{\alpha_2^4\alpha_3^5 p\left(1-e^{-\lambda d(1-z)}\right)}{(\alpha_2+\lambda(1-z))^4(\alpha_3+\lambda(1-z))^5}}\right).$$

General corresponding results for when there are no vacations are thus

$$P_q(z)\Big|_{p=0} = \frac{\left(h_c*(\lambda-\lambda z)-1\right)\left(1-\dfrac{\lambda}{\mu}\right)}{z-h_c*(\lambda-\lambda z)}$$

$$\text{and}\quad P(z)\Big|_{p=0} = \frac{(1-z)h_c*(\lambda-\lambda z)\left(1-\dfrac{\lambda}{\mu}\right)}{h_c*(\lambda-\lambda z)-z},$$

which are well known results for a regular M | G | 1 queuing system. A system with compulsory vacations, also known as a limited service system, is achieved when $p=1$.

Three values of d are considered for further investigation; 300, 600 and 900 minutes (5, 10 and 15 hours), while p is varied at increments of 0.1 across the interval $[0, 1]$. Results for the probability of zero, one and two or more customers in the system are displayed in Figure 10.4.4ii. Results when the vacation is fixed to either 300 or 600 minutes are not particularly dissimilar, but a disparity is seen when d is increased to 900 minutes.
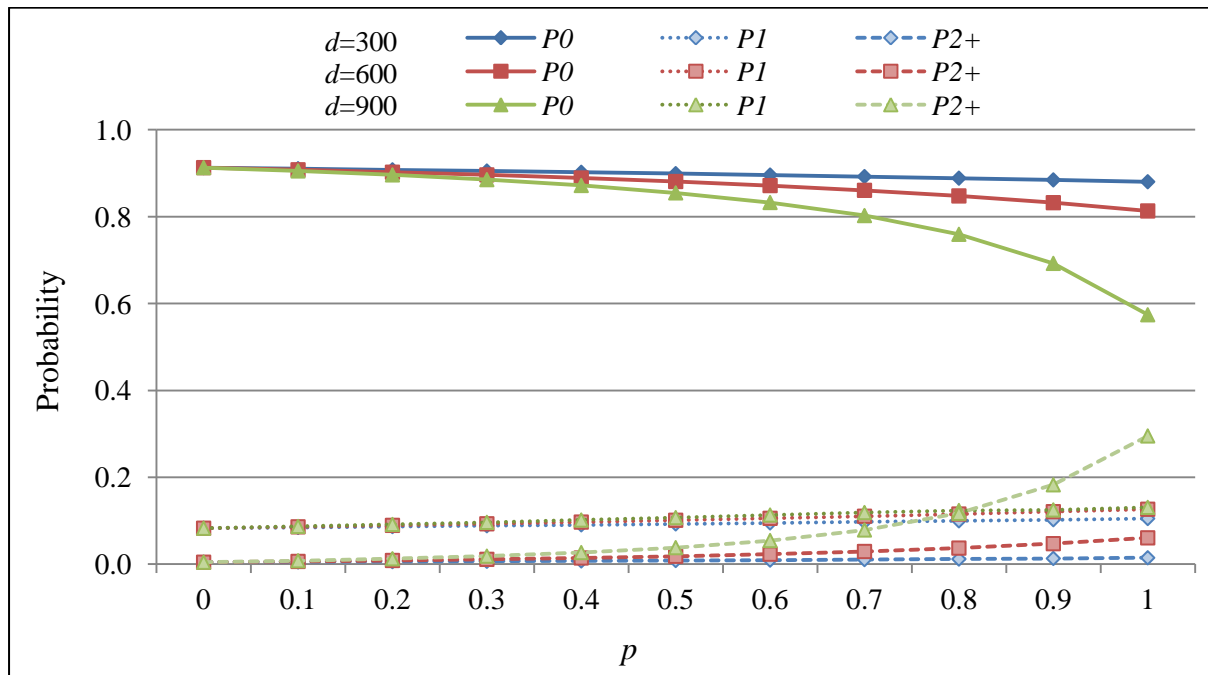

Figure 10.4.4ii: Results for number in system for different p and d

Further investigation shows additional insight into these differences and results are given in Table 10.4.4iii.

L_q, the expected number of customers in the queue, may be found by differentiating $P_q(z)$ at $z = 1$; then after some simplification the following result is achieved:

$$L_q = \frac{\lambda^2 (1 - \rho_{Bd})}{2} \left(\frac{E[c^2](1 + \lambda p(1-d)) + \frac{2p}{\mu}\left(1 - \frac{\lambda}{\mu}\right) + \lambda pd\left(pd + \frac{d}{\mu} + \frac{2}{\mu^2}\right)}{\left(1 - \frac{\lambda}{\mu} - \lambda pd\right)^2} \right).$$

This result may then be used to gain the expected waiting time in the queue, W_q, as well as the expected number in the system, L, and the expected time in the system W:

$$W_q = \frac{L_q}{\lambda}, \quad L = L_q + \rho_{Bd}, \quad W = \frac{L_q + \rho_{Bd}}{\lambda}.$$

The expected number in both the queue and system only exceeds one if vacations are compulsory and have a fixed length of 900 minutes; that is, these values exceed one if the theatre is used for other surgery types or closedown after every hip operation, and that the time allocated to these other tasks is fixed at 900 minutes (15 hours).

It can be seen that at this vacation length, the four performance measures all increase rapidly after $p = 0.7$. This suggests that, should a vacation time of 900 minutes be employed, it should not happen after more than 70% of hip operations so that the system is not too heavily impacted.

Having a set vacation time of ten hours shows very little difference to a vacation of five hours, and so a ten hour vacation is recommended. This gives a longer time to perform other operation types sequentially, or a longer and more realistic closedown time. Altering p also has little impact on results at this vacation length. Therefore, if allocating ten hours to other surgeries or closedown after every hip operation is implemented, little impact is seen on results when changed are made to how often this ten hour period is actually used.

Table 10.4.4iii: Performance measures for different p and d

p	d	L_q	W_q (minutes)	L	W (minutes)
0	300	0.005	5.4	0.092	104.4
	600	0.005	5.4	0.092	104.4
	900	0.005	5.4	0.092	104.4
0.1	300	0.005	6.2	0.095	108.0
	600	0.007	8.0	0.099	112.6
	900	0.010	10.8	0.104	118.5
0.2	300	0.006	7.1	0.099	111.8
	600	0.010	11.2	0.108	122.1
	900	0.016	18.5	0.120	136.4
0.3	300	0.007	8.1	0.102	116.0
	600	0.013	15.3	0.118	133.3
	900	0.026	29.5	0.141	159.8
0.4	300	0.008	9.3	0.106	120.4
	600	0.018	20.4	0.129	146.4
	900	0.040	45.9	0.169	191.4
0.5	300	0.009	10.6	0.110	125.2
	600	0.024	27.0	0.143	162.2
	900	0.063	71.0	0.208	235.8
0.6	300	0.011	12.0	0.115	130.3
	600	0.031	35.7	0.160	181.5
	900	0.099	111.7	0.266	301.7
0.7	300	0.012	13.6	0.120	135.9
	600	0.042	47.1	0.181	205.3
	900	0.161	182.7	0.359	406.7
0.8	300	0.014	15.4	0.125	141.9
	600	0.055	62.7	0.208	235.6
	900	0.282	320.1	0.523	593.0
0.9	300	0.015	17.5	0.131	148.5
	600	0.075	84.5	0.243	275.0
	900	0.559	633.6	0.867	982.4
1.0	300	0.017	19.8	0.137	155.7
	600	0.102	115.9	0.289	327.9
	900	1.414	1602.8	1.840	2085.7

10.5   Perfect world model

Finally, a 'perfect world' model is considered. Consider a system where all admitted hip patients are nominated for theatre the day after arrival. There are no cancellations or interruptions and all nominated patients will receive surgery as intended. This allows some planning to occur, whereby it is not assumed that all patients can enter theatre immediately, adding some feasibility to the model. By employing this model, there would be no unknowns for patient or staff; the number in surgery tomorrow is determined by the number of arrivals today.

The timing of arrivals at the operating theatre suite is fixed at a given time (say 8:30am), and all patients arrive from the ward together. Patients wait in the loading bay until the operating theatre suite becomes available (the first patient will enter the suite immediately). Service time (clinical time) for each patient is represented by the random variable C, which follows the probability distribution $h_c(t)$, as previously defined. Once all hip patients have undergone surgery, the operating theatre becomes available for other surgery types or for closedown (vacation); see Figure 10.5i.
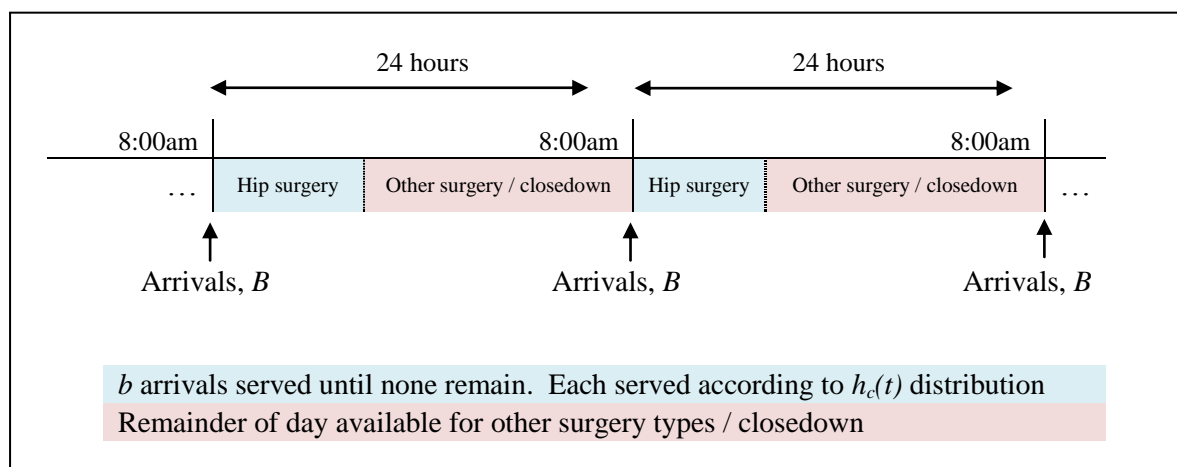

Figure 10.5i: The 'perfect world' model

The number of arrivals (requiring surgery) per day has been shown to fit the Poisson distribution. There are therefore batch arrivals of varying size which are assumed to arrive at a fixed time. Let the batch size be represented by the random variable B with mean $\bar{b}$, so

342

that the number of arrivals, b, on any day is given by the Poisson distribution,

$$P(B=b) = \frac{e^{-\bar{b}}\bar{b}^{b}}{b!},$$ and the expected number of arrivals annually is given by $\bar{b}$ x 365. Thus

when $B=0$, the blue section of Figure 10.5i disappears, but the model becomes invalid if the length of this blue section exceeds 24 hours.

10.5.1 Results

Let H be the total theatre time used for hip fracture patients, so that H is the sum of B instances of C. The number of arrivals, b, is taken from the random variable B. The result is therefore a sum of a random number, B, of independent identically distributed random variables. Using standard results, the mean and variance of H are given by

$$E[H] = E[B]E[C]$$

$$= b\,E[C]$$

$$\text{and } \operatorname{Var}[H] = E[B]\operatorname{Var}[C] + (E[C])^2 \operatorname{Var}[B]$$

$$= b\operatorname{Var}[C] + b(E[C])^2.$$

Using the data, $E[C]$ and $\operatorname{Var}[C]$ are equal to 99.0 and 35.7^2 (minutes) respectively (see Section 9.3.4), while $E[B] = \operatorname{Var}[B] = \bar{b} = 1.27$. Since the number of hip surgeries is determined by the number of arrivals the previous day, each day the time taken to perform all hip surgeries can be estimated more accurately than using the overall averages calculated previously. Let x be the known number of hip surgeries on a given day. Results for the mean and standard deviation of total theatre time for different x are summarised in Table 10.5.1i, calculated respectively via $x\,E[C]$ and $\sqrt{x\operatorname{Var}[C]}$.

Table 10.5.1ii: Mean and standard deviation of H (minutes) for different x

x	0	1	2	3	4	5	6	7
Mean	0	99	198	297	396	495	594	693
S.D.	-	35.7	50.5	61.8	71.4	79.8	87.4	94.5

By altering $\bar{b}$, the number of expected yearly arrivals can be estimated and the mean and standard deviation of H are consequently impacted; this relationship is shown in Figure 10.5.1ii (C is unaltered). The mean and standard deviation of H both increase as $\bar{b}$ increases, as would be expected, but it is interesting to see that they are becoming less alike as the number of arrivals increases, with the mean increasing at a faster rate. Current arrival rates would consume 18% of the daily theatre allocation on average, or 9% of 24 hours. For each additional 100 hip fracture arrivals per year, an extra 27 minutes extra per day of theatre time is required; this translates to 3.9% of the daily theatre allocation, or 1.8% of 24 hours.
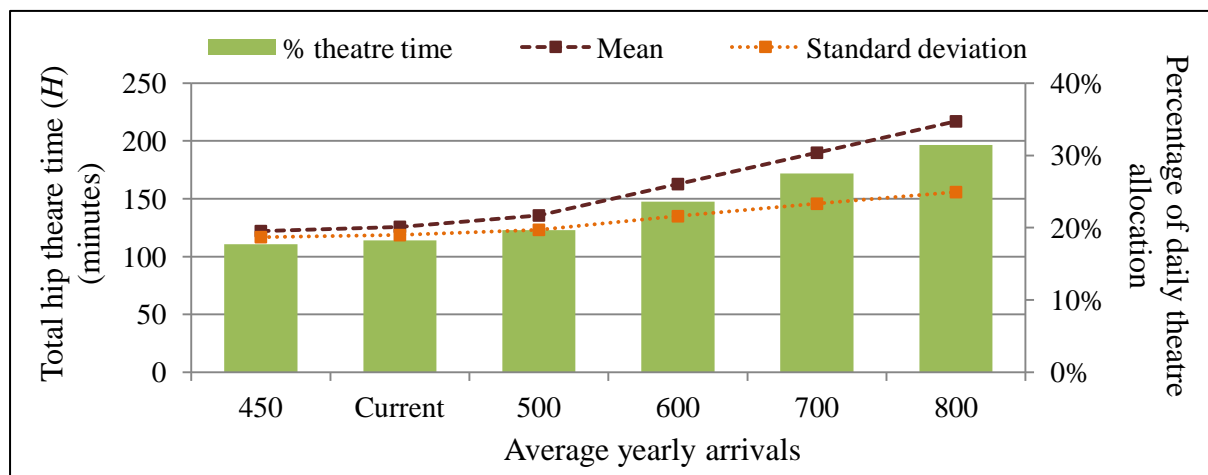

Figure 10.5.1ii: Total hip theatre time results for the 'perfect world' model

Finally the shape of the distribution is considered. Since the Laplace transform for the clinical time taken to complete one operation is known, the Convolution Theorem can be used to find the probability distribution for the clinical time taken for x operations. Denoting this probability distribution function by $h_c^x(t)$, then

$$L_s\left[h_c^x(t)\right] = \left(L_s\left[h_c(t)\right]\right)^x = \left(\frac{\alpha_2^4 \alpha_3^5}{\left(\alpha_2 + s\right)^4 \left(\alpha_3 + s\right)^5}\right)^x.$$

The expressions obtained by inverting this formula are omitted due to their lengths but results are now plotted for $x = 1, .., 4$. The shift of 99 minutes for each additional x, as well as the increase in variation, can be clearly seen in Figure 10.5.1iii.
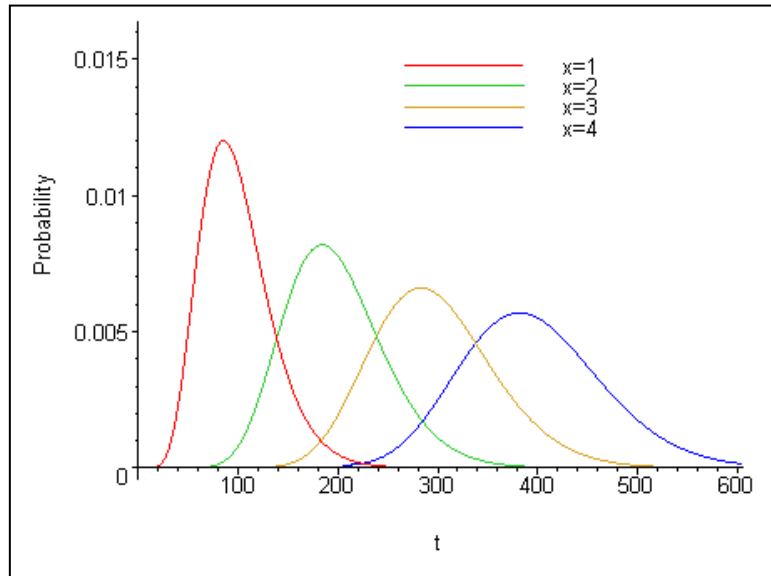
344

Figure 10.5.1iii: Distribution of total clinical time for different x

Having a PDF for total clinical time for all x patients is useful as it allows the probability of the total time to complete all x operations being within a given limit, t_l, to be calculated, simply by integrating $h_c^x(t)$ with respect to t on the interval $[0, t_l]$. Some results are given in Table 10.5.1iv. The median of $h_c^1(t)$ is equal to approximately 94 minutes, which was found by solving $\int_0^m h_c^1(t) = 0.5$ for m. This increases by approximately 99 minutes for each additional x, indicating a slight positive skew in the shape of each distribution.

Table 10.5.1iv: Probabilities of total time within t_l minutes for different x

x	Probability within t_l minutes							
	$t_l = 50$	$t_l = 100$	$t_l = 150$	$t_l = 200$	$t_l = 250$	$t_l = 300$	$t_l = 350$	$t_l = 400$
1	0.054	0.563	0.912	0.989	0.999	>0.999	>0.999	>0.999
2	<0.001	0.009	0.169	0.553	0.850	0.966	0.994	0.999
3	<0.001	0.001	0.003	0.043	0.233	0.551	0.812	0.943
4	<0.001	<0.001	<0.001	0.002	0.011	0.080	0.275	0.552

10.6   Chapter summary

The trauma theatre has been modelled using several mathematical approaches in this chapter. Patients were segregated into two groups, hip patients and non-hip patients. Arrivals were assumed to be random and arrival rates were dependent on source (patient type). Service time was also dependent upon patient type. A limit was placed on the number of patients allowed in the system at any time so that, if the system was already at capacity, then no additional arrivals would be permitted. Empirical data showed that a system limit of two was sensible and any impact realised from an increase beyond this should be negligible.

Initially, equations were formulated and solved under the assumption of a Negative Exponential service time distribution. It was found that while results were promising, the queuing model did not represent the theatre with a high level of compatibility. This was overcome by using the Erlang distribution to represent service time, still dependent upon patient type. System-state probabilities were formulated and solved and this time the model was shown to represent the real system with sufficient compatibility. A number of 'what-if' scenarios were then tested to explore the system further. Arrival and service rates were then inputted into equations formulated for a system limit of three.

Vacation queuing models were also investigated and a number of appropriate policies were considered. A vacation was classified as any time when there was not a hip patient occupying the theatre. Performance measures and summary results were given for a variety of models.

Finally, a novel 'perfect world' model was formulated. In this somewhat idealised situation, all hip patients go to theatre the day after arrival, thus reducing operative delay and removing uncertainty for these patients and their caregivers. Once all hip patients are seen, the theatre becomes free for other surgeries or closedown. Results relating to the total clinical time required for all hip patients, when the number of hip patients on a given day is known, were presented.

CHAPTER 11: CONCLUSIONS AND FURTHER WORK

11.1   Introduction

This thesis is brought to an end in this chapter where overall conclusions are given and some ideas for further work are presented. Four objectives were specified in Chapter 1 and some discussion of the achievement of these objectives is given forthwith.

As with many projects, the conclusion of this one does not necessarily mean that there are no further topics to explore or investigations to be made. Areas for further research can be split into two sections.

Firstly, there are ways in which the work presented here could be directly expanded upon, see Section 11.3.

Secondly, other relevant topics and methodologies can be explored, which have not been investigated in great detail here. Two concepts are introduced as potential possibilities for how this project could be developed further; risk scoring systems and the Fenton-Wilkinson approximation for estimating surgery duration, see Sections 11.4 and 11.5 respectively. Note that the expansion of this field is of course not limited to these two preliminary investigations but that they are included as examples. In each case a review of the literature is given, some initial work completed where deemed feasible and the possibility of additional developments in the area discussed.

11.2 Conclusions

The conclusions of this work on based on the satisfaction of the objectives stated in Section 1.4. More detailed information relating to conclusions has been given in the form of a summary at the end of every chapter.

Briefly, the objectives were: Objective 1 – investigate influencing factors relating to length of stay and mortality, particularly with regard to surgical delay; Objective 2 – build a simulation model of the hip fracture ward; Objective 3 – build a simulation model of the trauma theatre; Objective 4 – model the trauma theatre using queuing theory.

Objective 1 was achieved mainly in Chapters 3 to 5. CART and linear regression were used in Chapter 3 to determine important factors relating to length of stay, and CART and logistic regression were used in Chapter 4 to determine important factors relating to mortality. Variables which consistently indicated a relationship were then scrutinised in greater detail to further quantify any relationship. Surgical (/operative) delay was a key focus of all statistical investigations, not least because it was of primary interest to the clinicians involved in this project. Note that trauma hip fracture data from the UHW had never previously been investigated in such detail.

Mental state was found to be strongly associated with length of stay but this was not used as a marker variable in the simulation models. ASA grade, a measure of medical fitness, was used instead as, in liaison with the clinical team, it was decided that it is a more desirable variable to include due to the reduced ambiguity of assigning a grade to a patient, compared with a mental state score. Detailed linear and logistic regression analysis was performed for each ASA grade grouping and varying results were found. Delay was also shown to be associated with length of stay; the relationship was reduced but persisted once delay was split by pre-/post-operation. Other variables shown to be related to length of stay and mortality, but not included in the simulation model, include age and sex. It was decided that other selected variables provide more flexibility and appropriateness; determining care based on age or gender is clearly less appropriate than basing care options on medical fitness.

In Chapter 5, a principal components analysis (PCA) was used to collapse the dimensionality of the dataset and results inputted into a PCA regression model. While statistically

significant results were found, it was also concluded that perhaps the data is too complex to be reduced in this way.

Objectives 2 and 3 were achieved primarily in Chapters 6 and 8 respectively, drawing on results in each case from preceding chapters and thus highlighting the importance of a thorough statistical investigation prior to building a simulation model. In each case, a host of 'what-if' scenarios were considered in order to display how changes to the system will influence results. A particular consideration was the ageing population and the anticipated increase in demand on hip fracture services.

Many previous studies have focussed on determining whether there exist statistically significant relationships between length of stay / mortality and other variables, commonly with a focus on operative delay. However, no evidence could be found of studies which detail explicitly the implications of changing parameters, such as the proportion of delayed patients, on resources or any other measures. This research expands upon the statistical evaluation, using its output, with the building of two simulation models of the hip fracture ward in Chapter 6. This meant that instead of simply showing a statistical association between delay and length of stay (for example), various parameters, relating to both the distributions representing delay and the proportions representing the prevalence of delay, could be amended. An oversight by many other studies is also to not distinguish between pre- and post-operation length of stay. In Chapter 6, previous analyses were extended to include this distinction. ASA grade was included as a splitting variable in Model I. Results from previous chapters inferred the existence of relationships between ASA grade and length of stay and ASA grade and mortality. By incorporating the variable into the model, the consequences of focussing on a particular grouping could be seen. A similar approach was taken with respect to operation type for Model II.

A detailed examination of trauma theatre data was completed in Chapter 7. This led to a greater knowledge of the workings of the system and, importantly, where advances could be made. For example, tardiness and theatre turnover were shown to be two areas where considerable time savings could be made. Consequently, along with other findings, these were incorporated into the simulation model presented in Chapter 8. A key output of the model was lack of time cancellations and through the thorough scrutiny completed in Chapter 7, an appropriate method of modelling these cancellations could be found. A number of

scenarios were then tested to demonstrate how making changes, often relatively simple, could reduce these cancellations while not compromising theatre usage.

Objective 4 was achieved in Chapters 9 and 10. A variety of theoretical approaches were taken and their relevance and compatibility to modelling the trauma theatre was given. In Chapter 9, results from the M | G | 1 queuing system were used. The Laplace transform of the service time (clinical time) was found as a convolution of two Erlang distributions and arrivals were shown to be random. In order to account for the theatre being a terminating system in reality, results were extrapolated so that a non-terminating system, that had reached steady-state, could be assumed.

A novel and bespoke queuing system based on two types of arrival, hip patients and non-hip patients, each following an Erlang service time with different parameters, was formulated in Chapter 10. This investigation began with the formulation of a model based on random (Negative Exponential) service times and while results were promising, this model was not deemed to appropriately represent the real system. The Erlang model was then presented, with a system limit of two, shown to be valid according to the data, imposed in order to restrict the number of equations to be solved. It was concluded that the model excellently represented the system, with a high level of compatibility. Parameters were then varied in order to investigate system sensitivity to the inputs, and predict system changes based on alterations to these parameters. Results were later used to extend to a system limit of three.

Vacation queuing models were also looked at in detail and a number of existing models were adapted to represent the trauma theatre at the UHW. Server busy time was classified as the time when the theatre was occupied by a trauma hip patient, while vacation time was classified as the time when the theatre was occupied by a non-hip patient or was closed. In particular, the fixed length Bernoulli vacation model provides many useful results. One can choose how often the server goes on vacation and some variability is removed from the system by fixing the length of the vacation to a pre-determined value. It was shown that a ten hour vacation gives desirable results which are only negligibly impacted by changing vacation frequency. Finally, an original and innovative 'perfect world' model was presented, displaying the distributions of performing hip operations sequentially.

11.3   Extensions to this research

It would be unrealistic to cover all investigations relating to hip fracture patients in this piece of research. A number of objectives were shown to be satisfied by this work in the previous section, but there is still scope to extend these results further.

Detailed statistical output was presented primarily in Chapters 3 to 5. While the dataset available for those analyses was fairly large, several of the variables were sparsely-populated. For example, mortality analysis was completed primarily for death on the ward, but also some results were calculated for death within the University Health Board. There was minimal information in the dataset regarding mortality at four months but it was certainly not complete enough to perform further investigations with any level of accuracy. Follow-up information such as this would provide an interesting complement to work already completed, so if data collection improved then there would definitely be scope for further study in this area.

Simulation models have been used in this thesis to represent the trauma hip fracture ward and the trauma operating theatre at the UHW. These models, once validated, were then used to explore a variety of 'what-if' scenarios in order to discover the effect of making changes to the system. While several scenarios were tried in each case, not all possibilities are covered. However, it is considered that the most relevant and practicable scenarios have been investigated. One extension which could be made is to combine the models into one. This would be pursued in particular should additional data become available, so that all emergency trauma admissions could be modelled for the entirety of their hospital stay. It would be particularly interesting, for example, to determine the impact of surgical delay on other patient types, which would be one other way in which the statistical aspects of this thesis may be extended.

There is additional scope to extend the theoretical work presented in Chapters 9 and 10. Patients were split into two types, hip and non-hip, and arrival and service rates were dependent upon patient type. In Chapters 7 and 8, patients were split into three types, hip, spinal and other. Making the same split was not deemed necessary for the mathematical modelling work since the focus was on hip patients. However, if there was a shift in this focus so that spinal patients were also of interest, for example, then the model could be

extended. If the service time for the ith patient type $(i = 1, 2, 3)$ could be represented by an Erlang distribution with k_i phases, then there would be a total of

$$1 + (k_1 + k_2 + k_3) \sum_{n=1}^{l_{sys}} \sum_{m=1}^{n} m = 1 + \frac{1}{2}(k_1 + k_2 + k_3) \sum_{n=1}^{l_{sys}} n(n+1)$$ system states for a system limit of

l_{sys}.

Consider the case where hip patients are represented by an Erlang-9 distribution, as previously seen, and spinal and other patients are represented by two different Erlang-3 distributions. Setting $l_{sys} = 2$ would result in 61 system states, increasing to 151 if $l_{sys} = 3$, each of which would need to be formulated in order to analytically solve the queuing model, a considerable increase in the numbers seen previously.

Consider also the vacation queuing models approach from Section 10.4. Due to the wealth of literature on this topic, only the most relevant models were considered. A potential extension could thus be to apply some of the other vacation queuing models to the trauma theatre, even if it were only as an academic exercise. There could also be some worth of formulating a novel model, specifically designed around the trauma theatre at the UHW. This was beyond the scope of this thesis, since existing models could be used, but should requirements change then it could be a possible avenue to explore in the future.

11.4   Risk scoring systems

There has been some discussion of surgical outcome earlier in this thesis and the feasibility of predicting surgical outcome is now investigated further.   ASA grade, also previously discussed, is a measure of operative risk and clearly provides some useful information to predict surgical outcome. Other scoring systems have also been developed and are discussed forthwith, with the inclusion of surgical risk. It is suggested that the proper function of surgical risk scoring systems is the comparison of outcomes between surgeons and hospitals in a large number of patients (Treasure et al. 2002).

11.4.1 Review of the literature

The method of POSSUM scoring, Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity, was designed to assess outcome after surgery and was developed by multivariate discriminant analysis of 48 physiological and 18 operative factors. The final score consists of 12 physiological and six operative variables, each of which is scored on a scale of one to eight (Copeland et al. 1991). POSSUM was found to be the most appropriate method of risk scoring for general surgical practice at the time of a review of methods (Jones and de Cossart 1999).

A review was performed 12 years after the original methodology was devised and POSSUM was evaluated extensively in both general and specialist surgery and it was concluded that, when used correctly, POSSUM can be usefully applied in order to make comparisons between surgeons and between hospitals (Neary et al. 2003).   This was endorsed by a separate paper which stated that the sorting of patients into risk categories by the POSSUM system is useful for comparing hip fracture mortality between hospitals, but also advised the unfortunate actuality that this scoring system cannot be used for individual patients pre-operatively as a predictor of post-operation outcomes (Theis 2006).

However, another study found less positive results and in particular that POSSUM over-predicts the risk of death by more than twofold and the risk of death for low-risk patients by more than sevenfold. A modified p-POSSUM (Portsmouth POSSUM) predictor equation was thus formulated and found to give better results (Prytherch et al. 1998).

Using a modified operation classification, the original POSSUM equation has been validated for use in orthopaedic surgery (Mohamed et al. 2002). However, when POSSUM was first formally evaluated for fractured neck of femur surgery, it was found to be a poor predictor of outcome after operation (Ramanathan et al. 2005). It was found that POSSUM appeared to overestimate mortality in hip fracture patients, particularly in those patients with a higher predicted risk of dying. It was concluded that POSSUM should not be used to audit outcome after fractured neck of femur surgery and found that its role as a preoperative assessment tool is also limited. However, another investigation found that the orthopaedic POSSUM equations did agree well with observed mortality and morbidity data and it is suggested that if used as an audit tool it would allow an unbiased interpretation of results (Wright et al. 2008). The value of orthopaedic POSSUM in assessing mortality and morbidity following hip fractures over a period of six months was looked at in another study with positive results; the observed data showed a higher number of complications in patients allocated into a higher risk groups (Young et al. 2006).

The Surgical Risk Scale (SRS) scoring system incorporates clinical data using three classifications familiar to clinicians: the Confidential Enquiry into Perioperative Deaths (CEPOD) grade of operative urgency (NCEPOD 2004), ASA grade and the British United Provident Association (BUPA) schedule of operative procedures (BUPA 1990). Analysis showed that the SRS score was significantly predictive of death and did not over-predict mortality for low-risk procedures (Sutton et al. 2002).

One study compared the POSSUM, p-POSSUM and Surgical Risk Scale methodologies for a cohort of higher-risk patients, finding equal accuracy of prediction across all three methods (Brooks et al. 2005). Specifically for hip fractures, it was found that POSSUM and SRS over-predict operative mortality but that they are useful tools in prioritising time of surgery (Ahluwalia et al. 2009).

Another model was developed which predicts mortality based on variables which were found to be significantly correlated with death: ASA status, age, type of surgery (elective, urgent or emergency) and degree of surgery (minor, moderate or major) (Donati et al. 2004). Hip replacement was specified to be of grade two (moderate) surgery. The authors state that the advantage of their model is that "it can be applied preoperatively and does not require the use of intraoperative data"; so that the preoperative risk can be calculated. This differs from

354

POSSUM in that it can be used before an operation to assess risk of mortality instead of simply as an audit tool and while it was concluded that this model can be used to predict operative risk for both elective and emergency surgery in the operating room, some flaws have been highlighted (Ramanathan et al. 2005). Since the Donati score is based only on the four factors given previously, there will be very little variation obtained for hip fractures. Nearly all patients will fit into the older age group (above 70 years) and the operation will be classed as emergency and of a moderate degree for all patients. There will also be little variation between patients in terms of ASA grade and so here the use of this score is limited.

One of the earliest systems of risk stratification examined risk factors contributing to cardiac risk in non-cardiac surgery (Goldman et al. 1978), and was later updated to a more developed version which classifies risk as a score out of six, and is known as the RGCRI (Revised Goldman Cardiac Risk Index) (Lee et al. 1999). Another study formulated the BHOM (Biochemistry and Haematology Outcome Models) system was developed to address the problem of the large number of variables used by POSSUM to model outcome with the aim of excluding the least important variables and also factors deemed to be subjective (Prytherch et al. 2003). A review of different methods found that the p-POSSUM, SRS and BHOM systems were the most capable methods of predicting outcome after surgery but that the RGCBI did not discriminate accurately within the mortality groups; additionally it was suggested that the SRS has the advantage of ease of calculation (Neary et al. 2007).

Other studies have focussed particularly on a specific patient or operative procedure group, some of which are relevant to the patient cohort under study in this thesis. For example, one group assessed a risk-adjusted scoring tool used to predict outcome in patients aged 80 or over and found that their risk-adjusted mortality prediction compared favourably with observed outcomes (Nichols et al. 2008). Another study focussed on the same age group and found a 30-minute increment in duration of operation increased the odds of mortality by 17% and that post-operative mortality and morbidity increased progressively with increasing age (Turrentine et al. 2006). It was also found that risk scores may aid the assessment of sick elderly patients (undergoing abdominal surgery) but that, crucially, the opinion of an experienced clinician is still essential (Rix and Bates 2007).

11.4.2 Hip fracture specific scoring systems

Equations to predict postoperative risk were established specifically for patients with a hip fracture using the so-called Estimation of Physiologic Ability and Surgical Stress (E-PASS) scoring system, again with the aim of predicting morbidity and mortality, which it was claimed to achieve successfully (Hirose et al. 2009). While this may therefore seem a promising development in this field, the validity of the scoring system is under debate. For example, it has been suggested that the population sample used to create the E-PASS system was not representative of the patient cohort that the system represents (Moppett 2010), while the lack of proper statistical assessment also gives cause for concern (Zhou and Fan 2010).

The Nottingham Hip Fracture Score (NHFS) was developed to predict 30 day mortality for hip fracture patients by first determining key prognostic factors of mortality at 30 days and then incorporating these into a risk scoring system which can be used on an individual patient level to predict, at admission, the probability of 30 day mortality. Predictor variables were selected via univariate logistic regression and then entered into a multivariate logistic regression model to construct and validate the scoring system. Surgical and anaesthetic data was deliberately excluded. The area under the ROC curve was 0.719, indicating a reasonable fit, while similar fits were found by applying both the Donati score and a simple model based on ASA grade only to the same dataset (Maxwell et al. 2008).

Six different outcome risk scores were assessed for their predictive value with respect to three variables for elderly patients undergoing hip fracture surgery; incidence of serious complications, ambulation status after three months and survival at 90 days (Burgos et al. 2008). None of the scales investigated were able to predict risk of mortality at 90 days, while the Barthel Index (Mahoney and Barthel 1965) and the Visual Analogue Scale for Risk (RISK-VAS) (Arvidsson et al. 1996) were the most useful for predicting ability to walk three months after fracture. Half of the scoring systems investigated, namely the RISK-VAS score, the Charlson Index (Charlson et al. 1987) and the POSSUM score were found to reach sufficient predictive value of serious complications post-surgery.

11.4.3 Feasibility of further work in this area

It is stated that for a risk scoring system to be of clinical use, it needs to (a) use readily available and verifiable clinical information, (b) have been developed and validated on the population in whom it will be used, and (c) be free from any confounding factors (Maxwell et al. 2008). The main problem here is the satisfaction of (a); while a fair amount of data is collected on trauma hip fracture patients at the UHW, there was no information available on comorbidities. The presence of comorbidities is common in elderly patients and is likely to impact upon risk for these patients, whether this is for functional outcome, development of complications, mortality or another measure. It is regretful that comorbidities could not be included in the mortality analysis completed in Chapter 4 but this is an unfortunate eventuality of data restrictions.

It is therefore concluded that it would be impractical to develop an official scoring system at the current time. If further data could be collected in future, then this would be a potential prospect of further research in this area. The advantages of having an accurate scoring system are well-documented and it is recommended that the team at the UHW would benefit should such a system be created for their group of patients.

11.5 The Fenton-Wilkinson approximation

11.5.1 Theatre scheduling

There have been numerous proposed approaches, using a diverse range of methods, which tackle the issue of scheduling surgeries in operating theatres (Blake and Carter 1997, Cardoen et al. 2010), including making these decisions on the day of surgery (Dexter et al. 2004). This is particularly relevant here due to the nature of a trauma theatre. Methods reported to be used and studied previously include the surgeons' estimation of the duration of each case, although this technique has been shown to produce a high number of cancellations due to underestimations of the time required for a surgical procedure (Schofield et al. 2005). Another commonly-used method allocates procedures to surgical block times using average surgery durations calculated from historical data (Dexter 1996). One difficulty faced here is the time variability of the several processes involved; surgical times can be random by nature. Inaccurate predictions have been shown to decrease utilisation of operating rooms (Goldman et al. 1970), while above-average surgery times can increase net staffing costs (Abouleish et al. 2004). Where surgical time can be accurately estimated, it is advised to schedule the shortest operation first in order to reduce patient waiting time and staff overtime while offering greater predictability to the start time of the rest of the schedule (Lebowitz 2003).

Surgical times have often been shown to follow a Lognormal distribution (Spangler et al. 2004, Stepaniak et al. 2009a); this distribution is suitable due to the left-sided truncation (no negative times) and a long right-sided tail. It has also been suggested that the Lognormal distribution may approximate post-anaesthesia duration in the recovery room (Dexter and Tinker 1995).

Although the consideration of the variability given by these distributions can improve scheduling single cases, it does not tackle the issue of scheduling a block of surgery time, for which the sum of these times would be required. Before the topic of summing multiple Lognormal random variables is considered further, it must first be assessed whether this is an appropriate avenue to explore.

The purpose of this investigation is to assess the feasibility of predicting the total time that the emergency trauma theatre would be in use, given a set of planned operations. This information may then aid planning the theatre schedule for that day, thus reducing

cancellations. Note that due to the nature of emergency admissions, this not a typical operating room scheduling problem, where surgeries can be planned weeks or even months in advance.

11.5.2 Theatre times as Lognormal random variables

The aim here is to estimate the overall time for which the operating theatre is busy, given a planned set of operations. The turnover time must therefore also be considered in addition to surgical completion times. These have both previously been discussed in this thesis (see Chapters 7 and 8), but are considered again here with regard to the Lognormal distribution.

• Surgery duration

The classification of surgery types performed in the emergency trauma theatre is used again here, where there are three types of surgeries performed; hip operations, spinal operations and other operations. Previously the number of procedures performed per theatre episode was also considered but this extra classification is not used here due to loss of generality and data restrictions. The minimum threshold parameter for the Lognormal distribution was taken to be the minimum observed time in each group, while estimates for μ and σ were calculated using Solver and checked via Stat::Fit. Parameter estimates are given in Table 11.5.2i. The median was found to be accurate within one minute when comparing empirical and theoretical values for hip and other surgeries, and was within six minutes for spinal operations. While the first two moments may suggest a poor fit in some cases, a significant statistical fit was given by both the Anderson-Darling and Kolmogorov-Smirnov statistics in all cases, while close graphical fits were also found.

Table 11.5.2i: Parameter estimates and some goodness-of-fit comparisons for the Lognormal distribution fitted to surgery time (minutes)

Operation type	μ	σ	Min	Theoretical		Empirical	
				Mean	**S.D.**	**Mean**	**S.D.**
Hip	4.05	0.49	0	64.9	33.9	65.4	36.3
Spinal	4.46	0.57	5	106.9	63.6	96.1	58.2
Other	3.63	0.95	0	59.6	71.8	49.3	42.3

• **Turnover times**

Turnover times were modelled in Chapter 8 using the Lognormal distribution and the same parameters could therefore be used here. However, on closer inspection of these results it was found that turnover times preceding a hip operation followed a very similar distribution for each of the three operation types that the turnover time followed. These distributions are displayed in Figure 11.5.2ii.
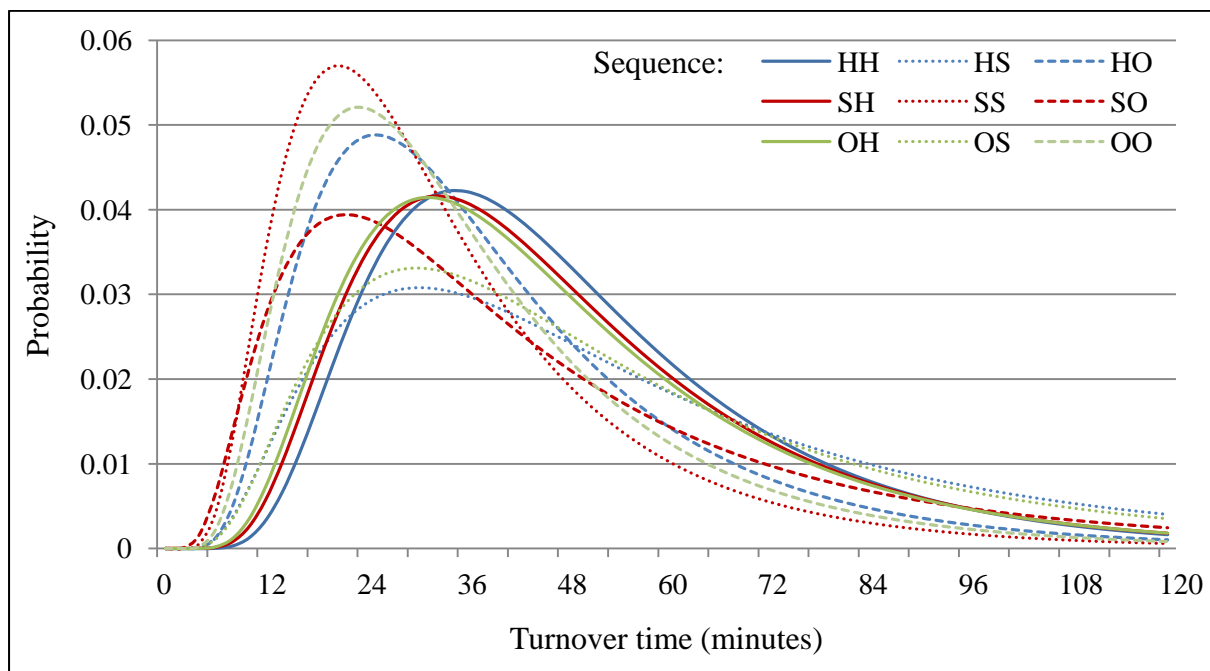

Figure 11.5.2ii: The Lognormal distribution fitted to turnover times for each sequence of operations (surgery types: H – hip, S – spinal, O – other)

All HH, SH and OH turnover times were thus combined and a new set of Lognormal parameters found; these were $\mu = 3.742$, $\sigma = 0.512$ and a minimum of one. Theoretical values are thus 42.9, 20.5 and 38.7 for the mean, standard deviation and median respectively, compared with empirical values of 47.3, 35.8 and 39 minutes.

11.5.3 Methodology

There are various different methods which approximate the sum of Lognormal random variables, as is required here. The chosen method used here is based on the Fenton-Wilkinson approximation of the sum of multiple Lognormal times (Fenton 1960) and is explained in more detail in due course. This approximation has been widely used across many fields including telecommunications (Stüber 2000), bioscience (López-Fidalgo and Sanchez 2005) and the financial sector (Finnerty 2003).

Some of the other methodologies available are rather more complex (Beaulieu and Xie 2004, Schwartz and Yeh 1982, Szyszkowicz and Yanikomeroglu 2009) and thus, should this investigation prove worthwhile, would be difficult to implement in practice. Note that the list of references given here is not exhaustive. Another advantage of the Fenton-Wilkinson approximation is that the differences between predicted and real times are smaller in the tail of the distribution (Wu et al. 2005) and it is important to accurately capture the longer cases as well as the 'normal' schedules.

Consider p independent random variables, X_i $(i = 1, ..., p)$, each having a different Lognormal distribution according to parameters μ_i and σ_i; $i = 1, ..., p$. There is no closed-form expression for the random variable P, where $P = \sum_{i=1}^{p} X_i$, but it can be approximated by another Lognormal random variable, $\hat{P}$.

The Fenton-Wilkinson approximation is obtained by matching the mean and variance parameters of the original p random variables. While the resultant expected value (mean) of

$\hat{P}$ is equal to summing the expected values of the X_i, the variance for this sum is not equal to the sum of the variances of the original Lognormal random variables.

The shape $\left(\sigma_{\hat{P}}\right)$ and scale $\left(\mu_{\hat{P}}\right)$ parameters of $\hat{P}$ are calculated by:

$$\sigma_{\hat{P}} = \sqrt{\ln\left(\frac{\sum_{i=1}^{p} e^{2\mu_i + \sigma_i^2}\left(e^{\sigma_i^2} - 1\right)}{\left(\sum_{i=1}^{p} e^{\mu_i + \sigma_i^2/2}\right)^2} + 1\right)}$$

$$\text{and} \quad \mu_{\hat{P}} = \ln\left(\sum_{i=1}^{p} e^{\mu_i + \sigma_i^2/2}\right) - \frac{\sigma_{\hat{P}}^2}{2}.$$

The minimum threshold parameter, $\min_{\hat{P}}$, was taken for this exercise to be the sum of the minimum values for each of the X_i, $i = 1, ..., p$.

11.5.4 Results

Using historical data, the parameters $\mu_{\hat{P}}$, $\sigma_{\hat{P}}$ and $\min_{\hat{P}}$ were estimated for each surgery schedule using the described methodology, allowing the comparison between the actual time for which the theatre was in use and the time predicted by the new Lognormal random variable. However, this of course poses the problem of how to sample a predicted value from the new random variable, $\hat{P}$, or rather which sampled value to take. There were 959 days suitable for this analysis.

The expected value (mean) is given by

$$E\left[\hat{P}\right] = e^{\mu_{\hat{P}} + \frac{\sigma_{\hat{P}}^2}{2}} + \min_{\hat{P}}.$$

The αth percentile point is calculated using

$$\hat{P}_{\alpha} = e^{\mu_{\hat{P}} + \sigma_{\hat{P}} \Phi^{-1}(\alpha)} + \min_{\hat{P}} \; ,$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative standard Normal distribution.

Using these formulae, a number of estimators of the schedule duration are available. One study used the second tertile cut-off point as the predictor for total theatre time (Alvarez et al. 2010) on the basis of an economic assessment which determined that if more than one third of operating rooms overrun on their schedule, then theatre allocations are not being planned appropriately (McIntosh et al. 2006).

Clearly there is a trade-off to be determined here; choosing a low percentile as the prediction time, underestimations and thus overutilisation is more likely, while sampling from a larger percentile within the distribution will be more prone to overestimations, resulting in underutilisation of the theatre.

First consider simply whether or not the predictor under- or overestimated the actual time the theatre was in use. The mean and a number of percentiles are given, see results in Table 11.5.4i (pc – percentile), where it is interesting to see that the second tertile has more of an even split between under- and overestimations than any other measure, including the median. The magnitude of these discrepancies is also of interest. Only the mean and the 50th, $66.\dot{6}^{\text{th}}$ and 75th percentiles are considered further, as these gave the better results. No predictor will ever be exact since empirical data is recorded in minutes while these calculations come from a continuous distribution.

Table 11.5.4i: Results of estimation accuracy of theatre usage for a number of predictors

Predictor	Percentage	
	Underestimated	Overestimated
Mean	64.8	35.2
$33.\dot{3}$ pc	84.5	15.5
50 pc	70.9	29.1
$66.\dot{6}$ pc	51.2	48.8
75 pc	39.9	60.1
95 pc	6.9	93.1
99 pc	0.7	99.3

On average, the mean of the distribution given by the Fenton-Wilkinson approximation underestimated the actual theatre busy time by 38 minutes, while the median underestimated the true time by an average of one hour. The second tertile was more accurate, overestimating the schedule by nine minutes on average, while the third quartile overestimated by an average of 53 minutes.

Another way to measure the quality of these predictions is to look at the correlation coefficient between actual and estimated values, although results must be taken with caution since they measure the extent to which one variable increases (or decreases) as another variable does the same, and not how similar they are to each other. The Spearman's rank coefficient is used here since data is non-Normal. The correlation coefficient between the actual theatre usage and the value predicted by the mean was $\rho = 0.2890$ ($p < 0.0001$ under the null hypothesis that $\rho = 0$), while it was found to be $\rho = 0.2902$ ($p < 0.0001$) between the actual value and the second tertile, showing a weak to moderate positive correlation in each case.

Results are given by means of a plot of the overall differences in Figure 11.5.4ii and the proportion of predictions within certain ranges in Figure 11.5.4iii, where each of the original measures are once again considered. Note that results in Figure 11.5.4iii are cumulative; for example, the mean predicts the value to be correct within 30 minutes for 18% of all cases, while 32% are within 60 minutes.

The shape of each of the distributions given by Figure 11.5.4ii suggest that the differences could follow a Normal distribution. This was tested for the differences when the mean is used as the predictor and indeed these differences were found to follow a Normal distribution with maximum likelihood parameter estimates found to be $\mu = 37.88$ and $\sigma = 136.9$. The graphical fits for this distribution can be found in Figure D11.5.4a of Appendix D where an excellent fit can be seen. This could be a useful result should this investigation be taken further at a later stage.
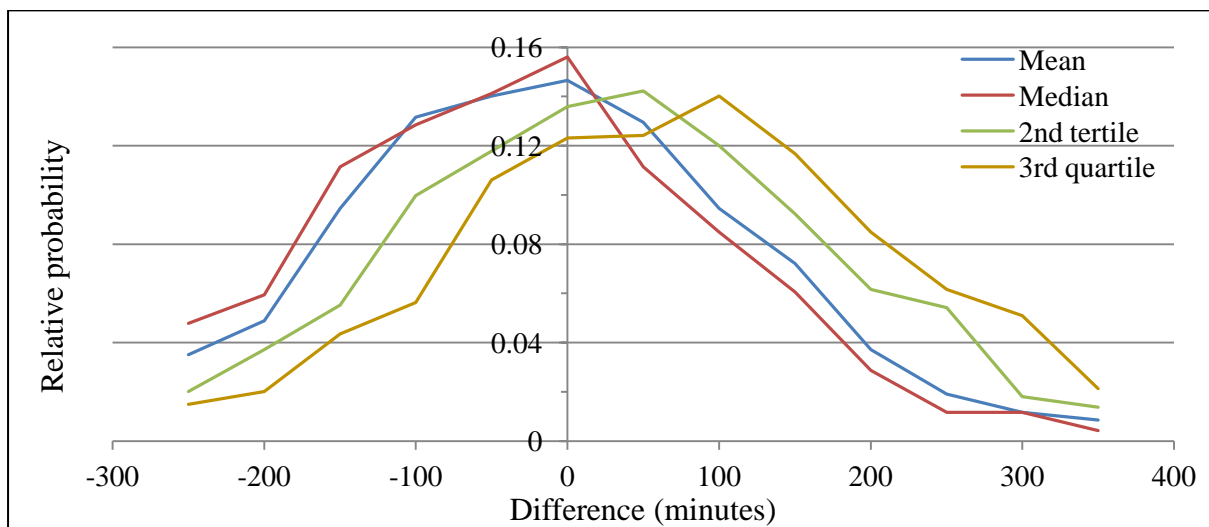
Figure 11.5.4ii: Distribution of differences between predicted and empirical theatre usage time
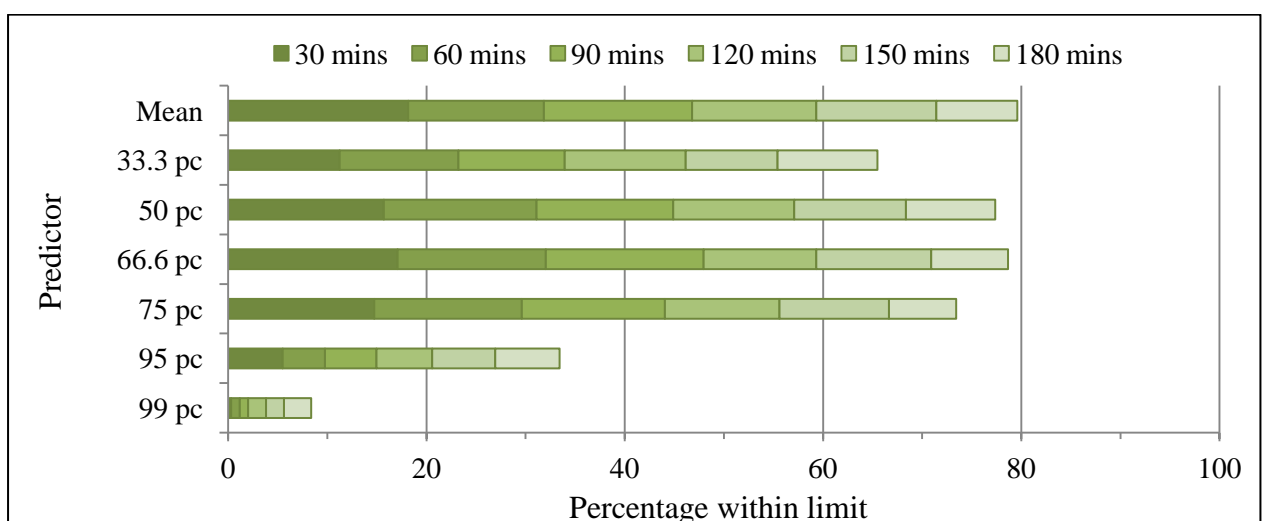

Figure 11.5.4iii: Percentage of predicted cases within various limits

There is a lot of variation in accuracy here and results have shown that the approximation used has not given particularly accurate estimations of theatre time in many cases. When using any of the predictor measures, the theatre time was not found to be correct within 3 hours in at least 20% of cases, which demonstrates once more the unpredictable behaviour of this system. It was already known that the theatre under study is subject to high variability in both demand and use and so these results are not particularly surprising.

It is clear that the success of this method depends on how variable the system is, as well as the choice taken for the measure used to make the prediction. It would be useful therefore to decide on the most suitable predictor measure. Let t_j be the actual length of time for which the theatre was busy on day j; and μ_j, σ_j and $\min_j$ are the parameter estimates for the Lognormal distribution for day j, based on the schedule for that day.

The percentile α_j which would correctly predict the real duration of t_j is calculated by

$$\Phi^{-1}\left(\alpha_j\right) = \frac{\ln\left(t_j - \min_j\right) - \mu_j}{\sigma_j}.$$

Using this formula, α_j was calculated for each day and results are now presented.
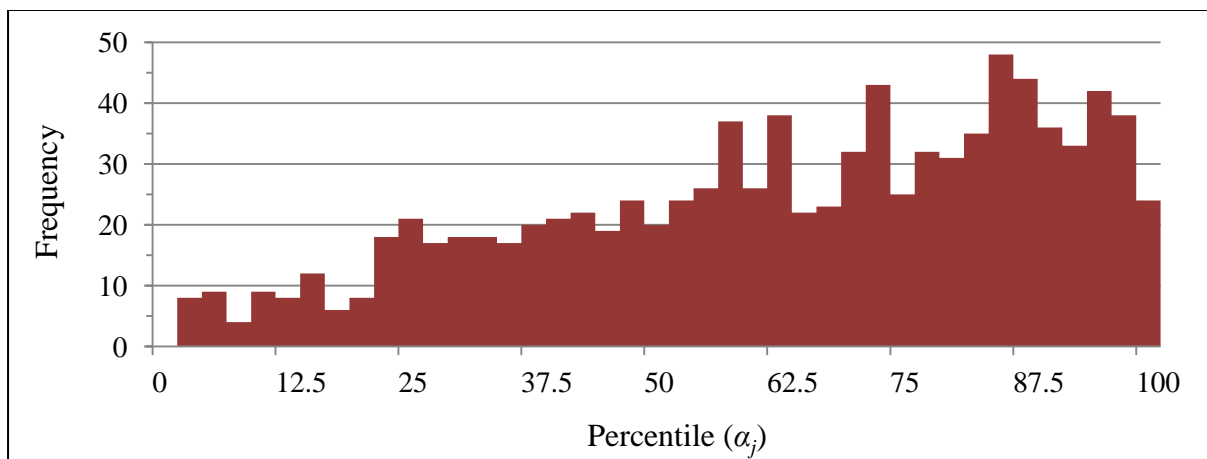

Figure 11.5.4iv: Frequency of percentile values at which the Fenton-Wilkinson approximation would correctly predict theatre busy time

There does not appear to be any clear pattern shown by Figure 11.5.4iv, except that the distribution is negatively-skewed. The mean percentile was 62, but with a relatively high standard deviation of 25.5 percentile points. A distributional fitting exercise showed that the percentile points (in decimal format) statistically fit the Beta distribution with maximum likelihood parameters calculated to be min = 0, max = 1, $p = 1.618$ and $q = 1.029$. This is shown graphically in Figure D11.5.4b of Appendix D.

Since this approximation has not shown to be particularly accurate in some cases, another approach is considered here. The busy time planned for the theatre is 11.5 hours (690 minutes), but this is measured as the time from when the first patient starts anaesthetic to when the last patient leaves the theatre; here only operation times are considered, so the time from when surgery starts for the *first* patient and finishes for the *last* patient is what is being measured (the turnover time accounts for the time taken to leave the theatre for every other patient). Summing the averages of these two additional times (the time taken between starting anaesthesia and starting surgery for the first patient and the time taken between finishing surgery and leaving the theatre for the last patient) gives 33 minutes. This is rounded so that the total time that the theatre is planned to be busy for is set to 11 hours (660 minutes). Using these thresholds, the percentage of cases for which the Fenton-Wilkinson approximation would correctly predict an over-run or an under-run of the scheduled theatre time can be calculated. Therefore while it may not be possible to accurately predict the running time of the theatre, this method would at least permit for planning whether or not the theatre will go over or under schedule. In the former case, cancellations could be made at the beginning of the day thus having a lesser impact on patients, while in the latter case an under-run of the schedule could allow for extra cases to be scheduled and thus maximising the potential utilisation the theatre. Results for the percentage of over-runs (in red) and under-runs (in green) correctly identified for the two discussed thresholds are given in Figure 11.5.4v, for a variety of predictors.
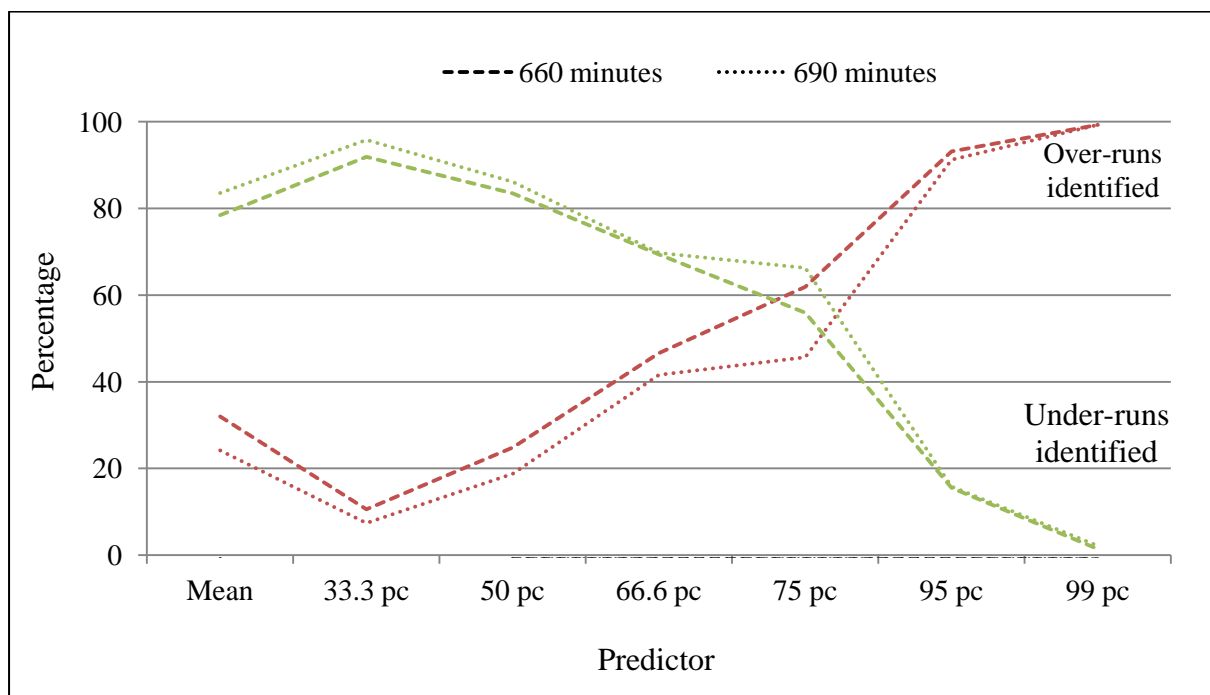
Figure 11.5.4v: Percentage of over- and under-runs of the schedule for two threshold values for theatre availability

There is clearly a trade-off to be found here, where the closest balances between correctly predictions either way are given by the second tertile and third quartile; the second tertile correctly identifies 47% of all over-runs and 70% of all under-runs, while the third quartile correctly identifies 62% and 55% of the same measures (when the threshold is set to 660 minutes). The overall number of correct predictions is in fact greatest overall for the first tertile in both cases, but this is influenced somewhat by there being considerably more under-runs than over-runs in the data set.

11.5.5 Challenges and conclusion

As an example of the difficulty in this area, results from a set of comparable schedules are presented. A sample was taken whereby the number of surgeries taking place on each day was six, which included one hip operation and five other operations (not including spinal surgery) in each case. Another criterion was that the hip operation was not the first or last operation of the day as this would influence turnover times. These decisions were made since

those days selected constitute a 'typical' schedule while providing a large enough sample on which to perform further analysis; in total 43 daily schedules suited these constraints. The mean duration of theatre time used was just over 10 hours (601 minutes), with a standard deviation of 1.21 hours (73 minutes). The minimum usage time was 7.15 hours (429 minutes) and the maximum was 12.1 hours (726 minutes). Underutilisation was seen in 37 of the 43 cases (86%). Some considerable deviation has thus been shown in this relatively small, but homogenous, set of results. The Lognormal distribution given by the Fenton-Wilkinson approximation for this schedule has a mean of 561 minutes, while the median is 536 minutes, the second tertile is 610 minutes and the value at the third quartile is 657 minutes. This led to an average underestimate of 40 minutes and 65 minutes by the mean and median values respectively, with respective overestimates of 9 minutes and 56 minutes by the second tertile and third quartile. On average the second tertile is thus estimating the overall theatre use accurately, but still the variation in actual completed schedules is problematic.

Another obstacle faced here is that there was only a limited amount of data available, meaning that operations were only compartmentalised into three types. Clearly the operation type of *Other* will have an enormous range of surgeries in it, while the complexities of these different operations will undoubtedly influence the time taken to complete the procedure. Furthermore, some patients undergo more than one operation during a single theatre episode. As an example of the diversity in this theatre, consider two patients who underwent surgery under the care of the same lead surgeon and were classified into the *Other* operation type category. They both had one procedure. The first patient's operation lasted 289 minutes (4.82 hours) and was for an open reduction internal fixation of the olecranon, while the second patient spent just 19 minutes (0.32 hours) in surgery for debridement of a wound. While the Lognormal distribution will capture this variation to some extent, the extremity of these differences will make any prediction exercise difficult. It would be desirable to split this group further (by complexity of operation, for example) and this could be an aspect to explore in the future.

Finally, trauma surgery is stochastic by nature; again while this can be captured by the distributions used to some extent, it may not be possible to completely overcome this issue. Another study found closer approximations than were found here (Alvarez et al. 2010), but

this was for cardiac elective surgery which is a more homogenous group of patients than in this case.

In conclusion, the Fenton-Wilkinson approximation has been found to be a reasonable method to use to predict busy time in the trauma theatre. While it has not been found to be specifically accurate in predicting the actual total time the theatre would be busy for, it was found to give reasonable results when predicting overall over- or under-running of the theatre schedule.

There are various prospective avenues to explore should this work be continued later, including further segregation of operation types should more data become available, as well as the investigation into other approximation methods. These were initially not considered here as it was required to find a simple solution which the theatre staff could potentially use on a daily basis. Alternatively it could be explored whether other distributions could be used and summed in a similar fashion using existing results; for example, the Gamma distribution (Moschopoulos 1985).

11.6   Chapter summary

As an end to this thesis, this chapter has included some discussion of how the thesis objectives have been satisfied, followed by some suggestions for future research, considered from two perspectives.

Firstly, a number of extensions to how the work presented in this thesis could be expanded upon were given. This included additional statistical analyses, extensions of the simulation models and expansion to the mathematical modelling.

Secondly, two potential fields of research were discussed which are relevant to this area but were not visited in detail throughout the main part of the thesis. These two areas, risk scoring systems and the Fenton-Wilkinson approximation for predicting total theatre time, were looked at in considerable detail.

It was ultimately concluded that pursuing risk scoring systems would not be appropriate unless additional data was made available. The Fenton-Wilkinson approximation provided some promising results but it is not considered that it would be appropriate to implement this approach to predicting total theatre time in a real life situation due to some inaccuracies which still remain. If these could be dealt with, then this could provide a more successful avenue for future exploration.

REFERENCES

Abouleish A, Dexter F, Whitten C, Zavaleta J and Prough D (2004) Quantifying net staffing costs due to longer-than-average surgical case durations. *Anesthesiology* 100 (2) 403-412.

Adunsky A, Arad M, Levi R, Blankstein A, Zeilig G and Mizrachi E (2005) Five-year experience with the 'Sheba' model of comprehensive orthogeriatric care for elderly hip fracture patients. *Disability & Rehabilitation* 27 (18-19) 1123-1127.

Aharonoff G, Koval K, Skovron M and Zuckerman J (1997) Hip fractures in the elderly: predictors of one year mortality. *Journal of Orthopaedic Trauma* 11 (3) 162-165.

Ahluwalia R, Mueller M, Kempraju N and Fonseca J (2009) The use of POSSUM and Surgical Risk Score (SRS) in predicting hip fractures needing early surgery to improve outcomes. Can we do better? *Journal of Bone and Joint Surgery, British Volume* 91-B 167-a.

Aitkenhead A (2005) Injuries associated with anaesthesia. A global perspective. *British Journal of Anaesthesia* 95 (1) 95-109.

Alegre-López J, Cordero-Guevara J, Alonso-Valdivielso J and Fernández-Melón J (2005) Factors associated with mortality and functional disability after hip fracture: an inception cohort study. *Osteoporosis International* 16 (7) 729-736.

Alfa A and Li W (2000) Optimal (*N*,*T*)-policy for M/G/1 system with cost structures. *Performance Evaluation* 42 (4) 265-277.

Allison P (1999) *Logistic Regression Using the SAS System: Theory and Application*. Cary, NC, USA: SAS Institute Inc.

Alvarez R, Bowry R and Carter M (2010) Prediction of the time to complete a series of surgical cases to avoid cardiac operating room overutilization. *Canadian Journal of Anesthesia* 57 (11) 973-979.

American Society of Anesthesiologists (2010) *ASA Physical Classification System.* http://www.asahq.org/For-Members/Clinical-Information/ASA-Physical-Status-Classification-System.aspx Accessed: May 2010

Anderson K, Koval K and Spratt K (2009) Hip fracture outcome: is there a "July effect"? *American Journal of Orthopedics* 38 (12) 606-611.

Arinzon Z, Shabat S, Peisakh A, Gepstein R and Berner Y (2010) Gender differences influence the outcome of geriatric rehabilitation following hip fracture. *Archives of Gerontology and Geriatrics* 50 86-91.

Arsuaga J and Carretero J (1994) Multivariate analysis of the sexual dimorphism of the hip bone in a modern human population and in early hominids. *American Journal of Physical Anthroplogy* 93 (2) 241-257.

Artalejo J (1998) Some results on the M/G/1 queue with *N*-policy. *Asia-Pacific Journal of Operational Research* 15 (2) 147-157.

Arvidsson S, Ouchterlony J, Sjöstedt L and Svårdsudd K (1996) Predicting postoperative adverse events. Clinical efficiency of four general classification systems. The project perioperative risk. *Acta Anaesthesiologica Scandinavica* 40 (7) 783-791.

Ashouri F, Al-Jundi W, Patel A and Mangwani J (2011) Management of warfarin anticoagulation in patients with fractured neck of femur. *ISRN Hematology* 294628.

Audit Commission (2002) Operating Theatres - A Bulletin for Health Bodies. London, UK: The Audit Commission.

Audit Commission (2003) Operating Theatres - Review of National Findings. London: The Audit Commission.

Avi-Itzhak B and Naor P (1963) Some queuing problems with the service station subject to breakdown. *Operations Research* 11 (3) 303-320.

Aylett V, Rosenberg K and Topp K (2007) Orthogeriatric Models of Care. London, UK: British Geriatrics Society.

Azhar A, Lim C, Kelly E, O'Rourke K, Dudeney S, Hurson B and Quinlan W (2008) Cost induced by hip fracture. *Irish Medical Journal* 101 (7) 213-215.

Baboolal K, Griffiths J, Knight V, Nelson A, Voake C and Williams J (2012) How efficient can an emergency unit be? A perfect world model. *Emergency Medicine Journal* (Published Online First February 2012).

Bailey N (1952) A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. *Journal of the Royal Statistical Society: Series B (Methodological)* 14 (2) 185-199.

Baker R, Squires B, Gargan M and Bannister G (2006) Total hip arthroplasty and hemiarthroplasty in mobile, independent patients with a displaced intracapsular fracture of the femoral neck. A randomized, controlled trial. *Journal of Bone and Joint Surgery* 88 (12) 2583-2589.

Balachandran K (1973) Control policies for a single server system. *Management Science* 19 (9) 1013-1018.

Barnes S, Golden B and Wasil E (2010) MRSA transmission reduction using agent-based modeling and simulation. *INFORMS Journal on Computing* 22 (4) 635-646.

BBC News (2010) *Falls amongst the elderly cost the NHS millions daily.* http://www.bbc.co.uk/news/10353642 Accessed: December 2010

Beaulieu N and Xie Q (2004) An optimal lognormal approximation to lognormal sum distributions. *IEEE Transactions on Vehicular Technology* 53 (2) 479-489.

Beaupre L, Cinats J, Senthilselvan A, Lier D, Jones C, Scharfenberger A, Johnston D and Saunders L (2006) Reduced morbidity for elderly patients with a hip fracture after implementation of a perioperative evidence-based clinical pathway. *Quality and Safety in Health Care* 15 (5) 375-379.

Beaupre L, Jones C, Saunders L, Johnston D, Buckingham J and Majumdar S (2005) Best practices for elderly hip fracture patients. A systematic review of the evidence. *Journal of General Internal Medicine* 20 (11) 1019-1025.

Beech R, Withey C and Morris R (1995) Understanding variations in lengths of stay between hospitals for fractured neck of femur patients and the potential consequences of reduced stay targets. *Journal of Public Health Medicine* 17 (1) 77-84.

Bergeron E, Lavoie A, Moore L, Bamvita J, Ratte S, Gravel C and Clas D (2006) Is the delay to surgery for isolated hip fracture predictive of outcome in efficient systems? *Journal of Trauma* 60 (4) 475-476.

Bhandari M, Devereaux P, Swiontkowski M, Tornetta 3rd P, Obremskey W, Koval K, Nork S, Sprague S, Schemitsch E and Guyatt G (2003) Internal fixation compared with arthroplasty for displaced fractures of the femoral neck. A meta-analysis. *Journal of Bone and Joint Surgery* 85-A (9) 1673-1681.

Bhattacharyya T, Vrahas M, Morrison S, Kim E, Wiklund R, Smith R and Rubash H (2006) The value of the dedicated orthopaedic trauma operating room. *Journal of Trauma* 60 (6) 1336-1340.

Birks Y, Porthouse J, Addie C, Loughney K, Saxon L, Baverstock M, Francis R, Reid D, Watt I and Torgerson D (2004) Randomized controlled trial of hip protectors among women living in the community. *Osteoporosis International* 15 (9) 701-706.

Black D, Bouxsein M, Marshall L, Cummings S, Lang T, Cauley J, Ensrud K, Nielson C, Orwoll E and Osteoporotic Fractures in Men (MrOS) Research Group (2008) Proximal femoral structure and the prediction of hip fracture in men: a large prospective study using QCT. *Journal of Bone and Mineral Research* 23 (8) 1326-1333.

Blake J and Carter M (1997) Surgical process scheduling: a structured review. *Journal of the Society of Health Systems* 5 (3) 17-30.

Bland J and Altman D (2000) The odds ratio. *British Medical Journal* 320 (7247) 1486.

Bleibler F, Konnopka A, Benzinger P, Rapp K and König H (2012) The health burden and costs of incident fractures attributable to osteoporosis from 2010 to 2050 in Germany - a demographic simulation model. *Osteoporosis* (Epub).

Blomfeldt R, Törnkvist H, Eriksson K, Söderqvist A, Ponzer S and Tidermark J (2007) A randomised controlled trial comparing bipolar hemiarthroplasty with total hip replacement for displaced intracapsular fractures of the femoral neck in elderly patients. *Journal of Bone and Joint Surgery, British Volume* 89 (2) 160-165.

BOA (2002) *Education and training for SHOs: a snapshot of the moment and recommendations for the future*. London, UK: British Orthopaedic Association.

BOA (2007) *The care of patients with fragility fracture*. London, UK: British Orthopaedic Association.

Bone P and Hooker L (2007) Just three letters: A case for reform for the NHS. The Cornerstone Group.

BoneMatters® (2011) *What is osteoporosis?* www.bonematters.co.uk Accessed: January 2011

Boonen S, Autier P, Barette M, Vanderschueren D, Lips P and Haentjens P (2004) Functional outcome and quality of life following hip fracture in elderly women: a prospective controlled study. *Osteoporosis International* 15 (2) 87-94.

Bosma E, de Jongh M and Verhofstad M (2010) Operative treatment of patients with pertrochanteric femoral fractures outside working hours is not associated with a higher incidence of complications or higher mortality. *Journal of Bone and Joint Surgery, British Volume* 92-B (1) 110-115.

Botella-Carretero J, Iglesias B, Balsa J, Zamarrón I, Arrieta F and Vázquez C (2008) Effects of oral nutritional supplements in normally nourished or mildly undernourished geriatric patients after surgery for hip fracture: a randomized clinical trial. *Journal of Parenteral and Enteral Nutrition* 32 (2) 120-128.

Bottle A and Aylin P (2006) Mortality associated with delay in operation after hip fracture: an observational study. *British Medical Journal* 332 (7547) 947-951.

Bowerman B and O'Connell R (1990) *Linear Statistical Models: An Applied Approach*. Second ed. Belmont, CA, USA: Duxbury Press.

Bowers J and Mould G (2002) The deferrable elective patient: a means of reducing waiting-lists in orthopaedics. *Journal of Management in Medicine* 16 (2-3) 150-158.

Bowers J and Mould G (2004) Managing uncertainty in orthopaedic trauma theatres. *European Journal of Operational Research* 154 (3) 599-608.

Brailsford S, Harper P, Patel B and Pitt M (2009) An analysis of the academic literature on simulation and modelling in health care. *Journal of Simulation* 3 (3) 130-140.

Braithwaite R, Col N and Wong J (2003) Estimating hip fracture morbidity, mortality and costs. *Journal of the American Geriatrics Society* 51 (3) 364-370.

Brennan née Saunders J, Johansen A, Butler J, Stone M, Richmond P, Jones S and Lyons R (2003) Place of residence and risk of fracture in older people: a population-based study of over 65-year-olds in Cardiff. *Osteoporosis International* 14 (6) 515-519.

Brooks M, Sutton R and Sarin S (2005) Comparison of Surgical Risk Score, POSSUM and p-POSSUM in higher-risk surgical patients. *British Journal of Surgery* 92 (10) 1288-1292.

Browne J, Pietrobon R and Olson S (2009) Hip fracture outcomes: does surgeon or hospital volume really matter? *Journal of Trauma* 66 (3) 809-814.

Bruyere O, Brandi M, Burlet N, Harvey N, Lyritis G, Minne H, Boonen S, Reginster J, Rizzoli R and Akesson K (2008) Post-fracture management of patients with hip fracture: a perspective. *Current Medical Research & Opinion* 24 (10) 2841-2851.

Bryan R, Nair P and Taylor M (2009) Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture risk. *Journal of Biomechanics* 42 (13) 2171-2176.

Buck N, Devlin H and Lunn J (1987) *The report of a confidential enquiry into perioperative deaths*. London, UK: Nuffield Provincial Hospitals Trust.

Bulajic-Kopjar M (2000) Seasonal variations in incidence of fractures among elderly people. *Injury Prevention* 6 (1) 16-19.

BUPA (1990) Schedule of fees. British United Provident Association Ltd. p. 3

Burgos E, Gómez-Arnau J, Díez R, Muñoz L, Fernández-Guisasola J and Garcia del Valle S (2008) Predictive value of six risk scores for outcome after surgical repair of hip fracture in elderly patients. *Acta Anaesthesiologica Scandinavica* 52 (1) 125-131.

Burns R, Moskowitz M, Ash A, Kane R, Finch M and McCarthy E (1999) Do hip replacements improve outcomes for hip fracture patients? *Medical Care* 37 (3) 285-294.

Butler M, Forte M, Joglekar S, Swiontkowski M and Kane R (2011) Evidence summary: systematic review of surgical treatments for geriatric hip fractures. *Journal of Bone and Joint Surgery* 93 (12) 1104-1115.

Byrne A and Jones J (1997) Responses to simulated anaesthetic emergencies by anaesthetists with different durations of clinical experience. *British Journal of Anaesthesia* 78 553-556.

Calder F, Jadhav V and Hale J (1998) The effect of a dedicated emergency theatre facility on emergency operating patterns. *Journal of the Royal College of Surgeons of Edinburgh* 43 (1) 17-19.

Calmes S and Shusterich K (1992) Operating room management: what goes wrong and how to fix it. *Physician Executive* 18 (6) 43-48.

Cardoen B, Demeulemeester E and Beliën J (2010) Operating room planning and scheduling: a literature review. *European Journal of Operational Research* 201 921-932.

Carson J (1986) Convincing users of model's validity is challenging aspect of modeler's job. *Industrial Engineering* 18 (6) 74-85.

Casella G and Berger R (1990) *Statistical Inference*. Belmont, CA, USA: Wadsworth Publishing Company.

Castro I, Barroso L and Sinnecker P (2005) Functional foods for coronary heart disease risk reduction: a meta-analysis using a multivariate approach. *American Journal of Clinical Nutrition* 82 (1) 32-40.

Cayirli T and Veral E (2003) Outpatient scheduling in health care: A review of the literature. *Production and Operations Management* 12 (4) 519-549.

Cayirli T, Veral E and Rosen H (2006) Designing appointment scheduling systems for ambulatory care services. *Health Care Management Science* 9 (1) 47-58.

Chacko A, Ramirez M, Ramappa A, Richardson L, Appleton P and Rodriguez E (2011) Does late night hip surgery affect outcome? *Journal of Trauma* 71 (2) 447-453.

Chakladar A and White S (2010) Cost estimates of spinal versus general anaesthesia for fractured neck of femur surgery. *Anaesthesia* 65 (8) 810-814.

Charalambous C, Yarwood S, Paschalides C, Siddique I, Hirst P and Paul A (2003) Factors delaying surgical treatment of hip fractures in elderly patients. *Annals of the Royal College of Surgeons of England* 85 (2) 117-119.

Charlson M, Pompei P, Ales K and MacKenzie C (1987) A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases* 40 (5) 373-383.

Chia W, Chen Y, Lin C, Liu H and Lin B (2008) Difference in time to x ray with similar time to treatment decision leads to differences in waiting times of patients with hip fracture in a crowded emergency department. *Emergency Medicine Journal* 25 (5) 274-275.

Chikwe J, de Souza A and Pepper J (2004) No time to train the surgeons. *British Medical Journal* 328 (7437) 418-419.

Choong P, Langford A, Dowsey M and Santamaria N (2000) Clinical pathway for fractured neck of femur: a prospective, controlled study. *Medical Journal of Australia* 172 (9) 423-426.

Chow Y and Robbins H (1965) On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *The Annals of Mathematical Statistics* 36 457-462.

Clague J, Craddock E, Andrew G, Horan M and Pendleton N (2002) Predictors of outcome following hip fracture. Admission time predicts length of stay and in-hospital mortality. *Injury* 33 (1) 1-6.

Clement N, Aitken S, Duckworth A, McQueen M and Court-Brown C (2011) The outcome of fractures in very elderly patients. *Journal of Bone and Joint Surgery, British Volume* 93 (6) 806-810.

Cluett J (2010) *Hip fracture.* http://orthopedics.about.com/ Accessed: November 2010

Cochran J and Bharti A (2006) A multi-stage stochastic methodology for whole hospital bed planning under peak loading. *International Journal of Industrial and Systems Engineering* 1 (1-2) 8-36.

Collantes E, Mauffrey C, Lewis C and Brewster M (2008) A review of 1241 trauma cases: A study of efficiency in trauma theatres. *Injury* 39 (7) 742-747.

Commission for Health Improvement (2003) *Cancelled operations.* http://www.chi.nhs.uk/Ratings/Trust/Indicator/indicatorDescriptionShort.asp?indicatorId=1003 Accessed: December 2009

Conway R (1963) Some tactical problems in digital simulation. *Management Science* 10 47-61.

Cook T (2008) The concept of surgical operating list 'efficiency': a formula to describe the term. *Anaesthesia* 63 (1) 98.

Copeland G, Jones D and Walters M (1991) POSSUM: a scoring system for surgical audit. *British Journal of Surgery* 78 (3) 355-360.

Costa A, Ridley S, Shahani A, Harper P, De Senna V and Nielson M (2003) Mathematical modelling and simulation for planning critical care capacity. *Anaesthesia* 58 (4) 320-327.

Cowie H, Lloyd M and Soutar C (1985) Study of lung function data by principal components analysis. *Thorax* 40 (6) 438-443.

Creemers S, Lambrecht M and Vandaele N (2007) Queueing models in healthcare. *Open access publications from Katholieke Universiteit Leuven*.

Crofts T, Griffiths J, Sharma S, Wygrala J and Aitkin R (1997) Surgical training: an objective assessment of recent changes for a single health board. *British Medical Journal* 314 (7084) 891-895.

Cummings S and Melton III L (2002) Epidemiology and outcomes of osteoporotic fractures. *Lancet* 359 (9319) 1761-1767.

Davies H, Crombie I and Tavaloki M (1998) When can odds ratios mislead? *British Medical Journal* 316 (7136) 989-991.

Davies R (1985) An assessment of models of a health system. *Journal of the Operational Research Society* 36 (8) 679-687.

Davis P (1992) *Generalizing concepts of verification, validation and accreditation (VV&A) for military simulation*. Santa Monica, CA, USA: The RAND Corporation.

Davis T, Sher J, Porter B and Checketts R (1988) The timing of surgery for intertrochanteric femoral fractures. *Injury* 19 (4) 244-246.

de Luise C, Brimacombe M, Pedersen L and Sørensen H (2008) Chronic obstructive pulmonary disease and mortality following hip fracture: a population-based cohort study. *European Journal of Epidemiology* 23 (2) 115-122.

DeAnda A and Gaba D (1991) Role of experience in the response to simulated critical incidents. *Anesthesia & Analgesia* 72 (3) 308-315.

Delaney C, Davis N and Tamblyn P (2010) Audit of the utilization of time in an orthopaedic trauma theatre. *ANZ Journal of Surgery* 80 (4) 217-222.

Der G and Everitt B (2002) *A Handbook of Statistical Analyses using SAS*. Second ed. Boca Raton, Florida, USA: Chapman & Hall / CRC.

Dexter F (1996) Application of prediction levels to OR scheduling - managing operating room usage time. *Association of periOperative Registered Nurses (AORN) Journal* 63 (3) 1-8.

Dexter F (2000) A strategy to decide whether to move the last case of the day in an operating room to another empty operating room to decrease overtime labor costs. *Anesthesia & Analgesia* 91 (4) 925-928.

Dexter F, Epstein R, Traub R and Xiao Y (2004) Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. *Anesthesiology* 101 (6) 1444-1453.

Dexter F and Tinker J (1995) Analysis of strategies to decrease postanesthesia care unit costs. *Anesthesiology* 82 (1) 94-101.

Di Monaco M, Di Monaco R, Manca M and Cavanna A (2002) Functional recovery and length of stay after recurrent hip fracture. *American Journal of Physical Medicine & Rehabilitation* 81 (2) 86-89.

Di Monaco M, Vallero F, Di Monaco R, Mautino F and Cavanna A (2003) Functional recovery and length of stay after hip fracture in patients with neurologic impairment. *American Journal of Physical Medicine & Rehabilitation* 82 (2) 143-148.

Dien J, Spencer K and Donchin E (2003) Localization of the event-related potential novelty response as defined by principal components analysis. *Cognitive Brain Research* 17 (3) 637-650.

Divecha H, Smith R, Cairns C and Bayer J (2011) Improving patient flow: the impact of consultant work pattern on trauma ward efficiency. *Surgeon* 9 (4) 175-178.

DoH (2001) National Service Framework for Older People. London, UK: Department of Health.

DoH (2004) Hospital episode statistics - 2003-2004 (England). London, UK: Department of Health.

Dolan P and Torgerson D (1998) The cost of treating osteoporotic fractures in the United Kingdom female population. *Osteoporosis International* 8 (6) 611-617.

Dolk T (1990) Operation in hip fracture patients - analysis of the time factor. *Injury* 21 (6) 369-372.

Donati A, Ruzzi M, Adrario E, Pelaia P, Coluzzi F, Gabbanelli V and Pietropaoli P (2004) A new and feasible model for predicting operative risk. *British Journal of Anaesthesia* 93 (3) 393-399.

Dorotka R, Schoechtner H and Buchinger W (2003a) Influence of nocturnal surgery on mortality and complications in patients with hip fracture. *Unfallchirurg* 106 (4) 287-293.

Dorotka R, Schoechtner H and Buchinger W (2003b) The influence of immediate surgical treatment of proximal femoral fractures on mortality and quality of life. *Journal of Bone and Joint Surgery, British Volume* 85-B (8) 1107-1113.

Doshi B (1986) Queueing systems with vacations - A survey. *Queueing Systems* 1 (1) 29-66.

Downs S and Black N (1998) The feasibility of creating a checklist for the assessment of the methodological quality of both randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health* 52 (6) 377-384.

Dudkiewicz I, Burg A, Salai M and Hershkovitz A (2011) Gender differences among patients with proximal femur fractures during rehabilitation. *Gender Medicine* 8 (4) 231-238.

Duguay C and Chetouane F (2007) Modeling and improving emergency department systems using discrete event simulation. *Simulation* 83 (4) 311-320.

Duncan D, Beck S, Hood K and A J (2006) Using dietetic assistants to improve the outcome of hip fracture: a randomised controlled trial of nutritional support in an acute trauma ward. *Age and Ageing* 35 (2) 148-153.

Durani P, Seagrave M and Neumann L (2005) The use of theatre time in elective orthopaedic surgery. *Annals of The Royal College of Surgeons of England* 87 (5) 170-172.

Dy C, McCollister K, Lubarsky D and Lane J (2011) An economic evaluation of a systems-based strategy to expedite surgical treatment of hip fractures. *Journal of Bone and Joint Surgery* 93 (14) 1326-1334.

Edelstein D, Aharonoff G, Karp A, Capla E, Zuckerman J and Koval K (2004) Effect of postoperative delirium on outcome after hip fracture. *Clinical Orthopaedics and Related Research* 422 195-200.

Edwards C, Counsell A, Boulton C and Moran C (2008) Early infection after hip fracture surgery: risk factors, costs and outcome. *Journal of Bone and Joint Surgery, British Volume* 90 (6) 770-777.

Eijkemans M, van Houdenhoven M, Nguyen T, Boersma E, Steyerberg E and Kazemier G (2010) Predicting the unpredictable: A new prediction model for operating room times using individual characteristics and the surgeon's estimate. *Anesthesiology* 112 (1) 41-49.

El-Darzi E, Vasilakis C, Chaussalet T and Millard P (1998) A simulation modelling approach to evaluating length of stay, occupancy, emptiness and bed blocking in a hospital geriatric department. *Health Care Management Science* 1 (2) 143-149.

Elder G, Harvey E, Vaidya R, Guy P, Meek R and Aebi M (2005) The effectiveness of orthopaedic trauma theatres in decreasing morbidity and mortality: a study of 701 displaced subcapital hip fractures in two trauma centres. *Injury* 36 (9) 1060-1066.

Elliott J, Beringer T, Kee F, Marsh D, Willis C and Stevenson M (2003) Predicting survival after treatment for fracture of the proximal femur and the effect of delays to surgery. *Journal of Clinical Epidemiology* 56 788-795.

Endo Y, Aharonoff G, Zuckerman J, Egol K and Koval K (2005) Gender differences in patients with hip fracture: a greater risk of morbidity and mortality in men. *Journal of Orthopaedic Trauma* 19 (1) 29-35.

Erlang A (1920) Telefon-Ventetider. Et Stykke Sandsynlighedsregning. *Matematisk Tidsskrift* B 25-42.

Eva K (2002) The aging physician: changes in cognitive processing and their impact on medical practice. *Academic Medicine* 77 (10) S1-S6.

Ezzati M, Vander Hoorn S, Rodgers A, Lopez A, Mathers C, Murray C and The Comparative Risk Assessment Collaborating Group (2003) Estimates of global and regional potential health gains from reducing multiple major risk factors. *Lancet* 362 (9380) 271-280.

Fackrell M (2009) Modelling healthcare systems with phase-type distributions. *Health Care Management Science* 12 (1) 11-26.

Faddy M and McClean S (1999) Analysing data on lengths of stay of hospital patients using phase-type distributions. *Applied Stochastic Models in Business and Industry* 15 (4) 311-317.

Fairley J (1991) Computers in medicine. *Journal of the Royal Society of Medicine* 84 566-567.

Fairley R (1976) Dynamic testing of simulation software. In *Proceedings of the 1976 Summer Computer Simulation Conference* (pp. 40-46). Washington D.C., USA.

Family Practice Notebook LLC (2012) *Lymphocyte count.* http://www.fpnotebook.com/hemeonc/lab/lymphcytcnt.htm Accessed: January 2012

Fantini M, Fabbri G, Laus M, Carretta E, Mimmi S, Franchino G, Favero L and Rucci P (2011) Determinants of surgical delay for hip fracture. *Surgeon* 9 (3) 130-134.

Fenton L (1960) The sum of log-normal probability distributions in scatter transmission systems. *IRE Transactions on Communication Systems* 8 (1) 57-67.

Field A (2009) *Discovering Statistics Using SPSS*. Third ed. Sage Publications Ltd.

Finnerty J (2003) The impact of transfer restrictions on stock prices. Washington DC, USA: AFA Meetings.

Fischer J and Johnson M (1990) Low body weight and weight loss in the aged. *Journal of the American Dietetic Association* 90 (12) 1697-1706.

Fishman G (1971) Estimating sample size in computing simulation experiments. *Management Science* 18 (1) 21-38.

Fitzgerald J and Dittus R (1990) Institutionalized patients with hip fractures: characteristics associated with returning to community dwelling *Journal of General Internal Medicine* 5 (4) 298-303.

Fletcher A, Halsall D, Huxham D and Worthington D (2007) The DH accident and emergency department model: A national generic model used locally. *Journal of the Operational Research Society* 58 (12) 1554-1562.

Fomundam S and Herrmann J (2007) A survey of queuing theory applications in healthcare. *The Institute for Systems Research* Technical Report (2007-24).

Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, Roberts K, Coyle E, Bevan G and 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 Medicine* 25 (4) 325-335.

Forster M C and Calthorpe D (2000) Mortality following surgery for proximal femoral fractures in centenarians. *Injury* 31 (7) 537-539.

Forte M, Virnig B, Swiontkowski M, Bhandari M, Feldman R, Eberly L and Kane R (2010) Ninety-day mortality after intertrochanteric hip fracture: does provider volume matter? *Journal of Bone and Joint Surgery* 92 (4) 799-806.

Foss N, Kristensen M and Kehlet H (2006) Prediction of postoperative morbidity, mortality and rehabilitation in hip fracture patients: the cumulated ambulation score. *Clinical Rehabilitation* 20 701-708.

Franklin W and White Jnr K (2008) Stationarity tests and MSER-5: exploring the intuition behind mean-squared-error-reduction in detecting and correcting initialization bias. In Mason, S, Hill, R, Mönch, L, Rose, O, Jefferson, T and Fowler, J (Eds.), *Proceedings of the Winter Simulation Conference* (pp. 541-546). Miami, USA.

Frey R (2010) *Hip revision surgery.* http://www.surgeryencyclopedia.com/Fi-La/Hip-Revision-Surgery.html Accessed: February 2010

Gafarian A, Ancker C and Morisaku T (1978) Evaluation of commonly used rules for detecting "Steady State" in computer simulation. *Naval Research Logistics Quarterly* 25 511-529.

Gakis K, Rhee H and Sivazlian B (1995) Distributions and first moments of the busy and idle periods in controllable M/G/1 queueing models with simple and dyadic policies. *Stochastic Analysis and Applications* 13 (1) 47-81.

Garcia A, Bonnaig J, Yoneda Z, Richards J, Ehrenfeld J, Sethi M, Jahangir A and Obremskey W (2011) Patient variables that may predict length of stay and incurred hospital costs in elderly patients with low-energy hip fracture. *Proceedings of The 2011 Annual Meeting of the Orthopaedic Trauma Association* (#74).

Gawande A, Zinner M, Studdert D and Brennan T (2003) Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 133 (6) 614-621.

Geiger F, Zimmermann-Stenzel M, Heisel C, Lehner B and Daecke W (2007) Trochanteric fractures in the elderly: the influence of primary hip arthroplasty on 1-year mortality. *Archives of Orthopaedic and Trauma Surgery* 127 (10) 959-966.

Gholve P, Kosygan K, Sturdee S and Faraj A (2005) Multidisciplinary integrated care pathway for fractured neck of femur. A prospective trial with improved outcome. *Injury* 36 (1) 93-98.

Gillespie L, Gillespie W, Robertson M, Lamb S, Cumming R and Rowe B (2003) Interventions for preventing falls in elderly people. *Cochrane Database of Systematic Reviews* 4 CD000340.

Gillespie W, Gillespie L and Parker M (2010) Hip protectors for preventing hip fractures in older people. *Cochrane Database of Systematic Reviews* 10 CD001255.

Girotto J, Koltz P and Drugas G (2010) Optimizing your operating room: Or, why large, traditional hospitals don't work. *International Journal of Surgery* 8 (5) 359-367.

Glenville M (2005) *Osteoporosis: The silent epidemic*. London, UK: Kyle Cathie Ltd.

Goldman J, Knappenberger H and Shearon W (1970) A study of the variability of surgical estimates. *Hospital Management* 110 (3) 46.

Goldman L, Caldera D, Southwick F, Nussbaum S, Murray B, O'Malley T, Goroll A, Caplan C, Nolan J, Burke D, Krogstad D, Carabello B and Slater E (1978) Cardiac risk factors and complications in non-cardiac surgery. *Medicine (Baltimore)* 57 (4) 357-370.

Goldsman D, Schruben L and Swain J (1994) Tests for transient means in simulated time series. *Naval Research Logistics* 41 171-187.

González-Montalvo J, Alarcón T, Mauleón G, Gil-Garay E, Gotor P and Martin-Vega A (2010) The orthogeriatric unit for acute patients: a new model of care that improves efficiency in the management of patients with hip fracture. *Hip International* 20 (2) 229-235.

Gorman T (1996) Osteoporosis. Hansard. pp. 314-315

Gorunescu F, McClean S and Millard P (2002) A queueing model for bed-occupancy management and planning of hospitals. *Journal of the Operational Research Society* 53 (1) 19-24.

Graham J, Chang P, Bergés I, Granger C and Ottenbacher K (2008) Race/ethnicity and outcomes following inpatient rehabilitation for hip fracture. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 63 (8) 860-866.

Gray R, Li X, Wawer M, Gange S, Serwadda D, Sewankambo N, Moore R, Wabwire-Mangen F, Lutalo T and Quinn T (2003) Stochastic simulation of the impact of antiretroviral therapy and HIV vaccines on HIV transmission; Rakai, Uganda. *AIDS* 17 (13) 1941-1951.

Green L (2006) Queueing analysis in healthcare. In: Hall, R (ed.) *Patient flow: Reducing delay in healthcare delivery*. Vol. 91. New York, USA: Springer-Verlag.

Green L, Soares J, Giglio J and Green R (2006) Using queueing theory to increase the effectiveness of emergency department provider staffing. *Academic Emergency Medicine* 13 (1) 61-68.

Gregory J and Aspden R (2008) Femoral geometry as a risk factor for osteoporotic hip fracture in men and women. *Medical Engineering & Physics* 30 (10) 1275-1286.

Gregory J, Stewart A, Undrill P, Reid D and Aspden R (2004) Identification of hip fracture patients from radiographs using Fourier analysis of the trabecular structure: a cross-sectional study. *BioMed Central Medical Imaging* 4 (1) 4.

Griffiths J, Jones M, Read M and Williams J (2010) A simulation model of bed-occupancy in a critical care unit. *Journal of Simulation* 4 (1) 52-59.

Griffiths J, Price-Lloyd N, Smithies M and Williams J (2006) A queueing model of activities in an intensive care unit. *IMA Journal of Management Mathematics* 17 (3) 277-288.

Grimes J, Gregory P, Noveck H, Butler M and Carson J (2002) The effects of time-to-surgery on mortality and morbidity in patients following hip fracture. *American Journal of Medicine* 112 (9) 702-709.

Gruson K, Aharonoff G, Egol K, Zuckerman J and Koval K (2002) The relationship between admission hemoglobin level and outcome after hip fracture. *Journal of Orthopaedic Trauma* 16 (1) 39-44.

Gulihar A, Nixon M, Jenkins D and Taylor G (2009) *Clostridium difficile* in hip fracture patients: Prevention, treatment and associated mortality. *Injury* 40 (7) 746-751.

Gullberg B, Johnell O and Kanis J (1997) World-wide projections for hip fracture. *Osteoporosis International* 7 (5) 407-413.

Gupta I, Zoreda J and Kramer N (1971) Hospital manpower planning by use of queueing theory. *Health Services Research* 6 (1) 76-82.

Günal M and Pidd M (2010) Discrete event simulation for performance modelling in health care: a review of the literature. *Journal of Simulation* 4 (1) 42-51.

Haentjens P, Lamraski G and Boonen S (2005) Costs and consequences of hip fracture occurence in old age: an economic perspective. *Disability and Rehabilitation* 27 (18-19) 1129-1141.

Haentjens P, Magaziner J, Colón-Emeric C, Vanderschueren D, Milisen K, Velkeniers B and Boonen S (2010) Meta-analysis: excess mortality after hip fracture among older men and women. *Annals of Internal Medicine* 152 (6) 380-390.

Haleem S, Lutchman L, Mayahi R, Grice J and Parker M (2008) Mortality following hip fracture: trends and geographical variations over the last 40 years. *Injury* 39 (10) 1157-1163.

Hamlet W, Lieberman J, Freedman E, Dorey F, Fletcher A and Johnson E (1997) Influence of health status and the timing of surgery on mortality in hip fracture patients. *American Journal of Orthopedics* 26 (9) 621-627.

Handoll H, Cameron I, Mak J and Finnegan T (2009) Multidisciplinary rehabilitation for older people with hip fractures. *Cochrane Database of Systematic Reviews* 4 CD007125.

Handoll H and Parker M (2008) Conservative versus operative treatment for hip fractures in adults. *Cochrane Database of Systematic Reviews* 3 CD000337.

Hanley J and McNeil B (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143 29-36.

Hannan M, Felson D, Dawson-Hughes B, Tucker K, Cupples L, Wilson P and Kiel D (2000) Risk factors for longitudinal bone loss in elderly men and women: the Framingham Osteoporosis Study. *Journal of Bone and Mineral Research* 15 (4) 710-720.

Hans D, Dargent-Molina P, Schott A, Sebert J, Cormier C, Kotzki P, Delmas P, Pouilles J, Breart G and Meunier P (1996) Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. *Lancet* 348 (9026) 511-514.

Harper P and Leite Jr E (2008) TreeWorks: Advances in scalable decision trees. *International Journal of Healthcare Information Systems and Informatics* 3 (4) 53-68.

Harper P and Shahani A (2002) Modelling for the planning and management of bed capacities in hospitals. *Journal of the Operational Research Society* 53 (1) 11-18.

Harries D and Eastwood H (1991) Proximal femoral fractures in the elderly: does operative delay for medical reasons affect short-term outcome? *Age and Ageing* 20 (1) 41-44.

Harris W (1969) Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *Journal of Bone and Joint Surgery* 51 (4) 737-755.

Hartz A, Kuhn E and Pulido J (1999) Prestige of training programs and experience of bypass surgeons as factors in adjusted patient mortality rates. *Medical Care* 37 93-103.

Haussmann R (1970) Waiting time as an index of quality of nursing care. *Health Services Research* 5 (2) 92-105.

Health Workforce Solutions LLC (2008) *Hospital at Home.* http://www.innovativecaremodels.com/care_models/20 Accessed: July 2011

Healthcare Commission (2005) Day surgery. London: Commission for Healthcare Audit and Inspection.

Hershkovitz A, Kalandariov Z, Hermush V, Weiss R and Brill S (2007) Factors affecting short-term rehabilitation outcomes of disabled elderly patients with proximal hip fracture. *Archives of Physical Medicine and Rehabilitation* 88 (7) 916-921.

Heyman D (1968) Optimal operating policies for M/G/1 queuing system. *Operations Research* 16 (2) 362-368.

Heyman D (1977) The *T* policy for the M/G/1 queue. *Management Science* 23 (7) 775-778.

Hirose J, Ide J, Irie H, Kikukawa K and Mizuta H (2009) New equations for predicting postoperative risk in patients with hip fracture. *Clinical Orthopaedics and Related Research* 467 (12) 3327-3333.

Ho V, Hamilton B and Roos L (2000) Multiple approaches to assessing the effects of delays for hip fracture patients in the United States and Canada. *Health Services Research* 34 (7) 1499-1518.

Hoad K, Robinson S and Davies R (2008a) *Automating DES output analysis: How many replications to run.* http://www2.warwick.ac.uk/fac/soc/wbs/projects/autosimoa/replications/ Accessed: June 2009

Hoad K, Robinson S and Davies R (2008b) Automating warm-up length estimation. In Mason, S, Hill, R, Mönch, L, Rose, O, Jefferson, T and Fowler, J (Eds.), *Proceedings of the Winter Simulation Conference*. Miami, USA.

Hoenig H, Rubenstein L, Sloane R, Horner R and Kahn K (1997) What is the role of timing in the surgical and rehabilitative care of community-dwelling older persons with acute hip fracture? *Archives of Internal Medicine* 157 (5) 513-520.

Hollingworth W, Todd C and Parker M (1995) The cost of treating hip fractures in the twenty-first century. *Journal of Public Health Medicine* 17 (3) 269-276.

Hollingworth W, Todd C, Parker M, Roberts J and Williams R (1993) Cost analysis of early discharge after hip fracture. *British Medical Journal* 307 (6909) 903-906.

Holmes J and House A (2000) Psychiatric illness predicts poor outcome after surgery for hip fracture: a prospective cohort study. *Psychological Medicine* 30 (4) 921-929.

Holt G, Smith R, Duncan K, Hutchison J and Gregori A (2008) Outcome after surgery for the treatment of hip fracture in the extremely elderly. *Journal of Bone and Joint Surgery* 90 (9) 1899-1905.

Holt G, Smith R, Duncan K, Hutchison J and Reid D (2009) Changes in population demographics and the future incidence of hip fracture. *Injury* 40 (7) 722-726.

Holt G, Smith R, Duncan K and McKeown D (2010) Does delay to theatre for medical reasons affect the peri-operative mortality in patients with a fracture of the hip? *Journal of Bone and Joint Surgery, British Volume* 92-B (6) 835-841.

Hommel A, Bjorkelund K, Thorngren K-G and Ulander K (2008a) Differences in complications and length of stay between patients with a hip fracture treated in an orthopaedic department and patients treated in other hospital departments. *Journal of Orthopaedic Nursing* 12 (1) 13-25.

Hommel A, Ulander K, Bjorkelund K B, Norrman P-O, Wingstrand H and Thorngren K-G (2008b) Influence of optimised treatment of people with hip fracture on time to operation, length of hospital stay, reoperations and mortality within 1 year. *Injury* 39 (10) 1164-1174.

Hosmer D and Lemeshow S (1989) *Applied Logistic Regression*. New York, USA: John Wiley & Sons, Inc.

Hosmer D and Lemeshow S (2000) *Applied Logistic Regression*. Second ed. New York, USA: John Wiley & Sons, Inc.

Hossain M, Neelapala V and Andrew J G (2009) Results of non-operative treatment following hip fracture compared to surgical intervention. *Injury* 40 (4) 418-421.

Huang X-M, Taylor C and Stewart I (1995) Computer simulation of A&E department. *OR Insight* 8 18-23.

Huddleston J and Whitford K (2001) Medical care of elderly patients with hip fractures. *Mayo Clinic Proceedings* 76 (3) 295-298.

Hur S, Kim J and Kang C (2003) An analysis of the M/G/1 system with *N* and *T* policy. *Applied Mathematical Modelling* 27 (8) 665-675.

Hutchings L, Fox R and Chesser T (2011) Proximal femoral fractures in the elderly: How are we measuring outcome? *Injury* 42 (11) 1205-1213.

Isenberg D, Maddison P, Woo P, Glass D and Breedveld F (2004) *Oxford Textbook of Rheumatology*. Third ed. Oxford, UK: Oxford University Press.

Jacobsen S, Sargent D, Atkinson E, O'Fallon W and Melton III L (1995) Population-based study of the contribution of weather to hip fracture seasonality. *American Journal of Epidemiology* 141 (1) 79-83.

Jain M and Jain A (2010) Working vacations queueing model with multiple types of server breakdown. *Applied Mathematical Modelling* 34 (1) 1-13.

Jain R, Koo M, Kreder H, Schemitsch E, Davey J and Mahomed N (2002) Comparison of early and delayed fixation of subcapital hip fractures in patients sixty years of age or less. *Journal of Bone and Joint Surgery* 84-A (9) 1605-1612.

Jan F, Tabish S, Qazi S and Atif M (2003) Time utilization of operating rooms at a large teaching hospital. *Journal of the Academy of Hospital Administration* 15 1-6.

Jensen J (1984) Determining factors for the mortality following hip fractures. *Injury* 15 (6) 411-414.

Johansen A, Evans R, Stone M, Richmond P, Lo S and Woodhouse K (1997) Fracture incidence in England and Wales: a study based on the population of Cardiff. *Injury* 28 (9-10) 655-660.

Johnell O (1997) The socioeconmic burden of fractures: today and in the 21st century. *American Journal of Medicine* 103 (2A) S20-S25.

Johnell O and Kanis J (2005) Epidemiology of osteoporotic fractures. *Osteoporosis International* 16 (Supp. 2) S3-S7.

Johnell O, Kanis J, Oden A, Johansson H, De Laet C, Delmas P, Eisman J, Fujiwara S, Kroger H, Mellstrom D, Meunier P, Melton III L, O'Neill T, Pols H, Reeve J, Silman A and Tenenhouse A (2005) Predictive value of BMD for hip and other fractures. *Journal of Bone and Mineral Research* 20 (7) 1185-1194.

Johnston A, Barnsdale L, Smith R, Duncan K and Hutchison J (2010) Change in long-term mortality associated with fractures of the hip: Evidence from the Scottish Hip Fracture Audit. *Journal of Bone and Joint Surgery, British Volume* 92-B (7) 989-993.

Jones D, Woodnutt D and Leyshon R (2003) Operative data inaccuracies: What operations do we perform? - An audit of 'Theatreman' system. *Journal of Bone and Joint Surgery, British Volume* 85-B (Supp. II) 152.

Jones H and de Cossart L (1999) Risk scoring in surgical patients. *British Journal of Surgery* 86 (2) 149-157.

Jorm A and Jolley D (1998) The incidence of dementia: A meta-analysis. *Neurology* 51 728-733.

Juliebø V, Krogseth M, Skovlund E, Engedal K and Wyller T (2010) Medical treatment predicts mortality after hip fracture. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 65A (4) 442-449.

Jun J, Jacobson S and Swisher J (1999) Application of discrete-event simulation in healthcare clinics: A survey. *Journal of the Operational Research Society* 50 (2) 109-123.

Kaiser H (1959) Computer program for varimax rotation in factor analysis. *Educational and Psychological Measurement* 19 413-420.

Kamel H, Iqbal M, Mogallapu R, Maas D and Hoffmann R (2003) Time to ambulation after hip fracture surgery: relation to hospitalization outcomes. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 58 (11) 1042-1045.

Kannegaard P, van der Mark S, Eiken P and Abrahamsen B (2010) Excess mortality in men compared with women following a hip fracture. National analysis of comedications, comorbidity and survival. *Age and Ageing* 39 (2) 203-209.

Kannus P, Parkkari J, Sievänen H, Heinonen A, Vuori I and Järvinen M (1996) Epidemiology of hip fractures. *Bone* 18 (Supp. 1) S57-S63.

Kaptoge S, Beck T, Reeve J, Stone K, Hillier T, Cauley J and Cummings S (2008) Prediction of incident hip fracture risk by femur geometry variables measured by hip structural analysis in the study of osteoporotic fractures. *Journal of Bone and Mineral Research* 23 (12) 1982-1904.

Karaca S, Ayhan E, Kesmezacar H and Uysal O (2012) Hip fracture mortality: is it affected by anesthesia techniques? *Anaesthesiology Research and Practice* 708754.

Karasik D, Cupples L, Hannan M and Kiel D (2004) Genome screen for a combined bone phenotype using principal component analysis: the Framingham study. *Bone* 34 (3) 547-556.

Katz S and Stroud M (1989) Functional assessment in geriatrics. A review of progress and directions. *Journal of the American Geriatrics Society* 37 (3) 267-271.

Ke J-C (2003) The optimal control of an M/G/1 queueing system with server vacations, startup and breakdowns. *Computers & Industrial Engineering* 44 (4) 567-579.

Ke J-C (2005) Modified *T* vacation policy for an M/G/1 queueing system with an unreliable server and startup. *Mathematical and Computer Modelling* 41 (11-12) 1267-1277.

Ke J-C (2006) On M/G/1 system under *NT* policies with breakdowns, startup and closedown. *Applied Mathematical Modelling* 30 (1) 49-66.

Ke J-C, Wu C-H and Zhang Z (2010) Recent developments in vacation queueing models: A short survey. *International Journal of Operations Research* 7 (4) 3-8.

Kelton W and Law A (1983) A new approach for dealing with the startup problem in discrete event simulation. *Naval Research Logistics Quarterly* 30 641-658.

Kelz R, Freeman K, Hosokawa P, Asch D, Spitz F, Moskowitz M, Henderson W, Mitchell M and Itani K (2008) Time of day is associated with postoperative morbidity: An analysis of the national surgical quality improvement program data. *Annals of Surgery* 247 (3) 544-552.

Kelz R, Tran T, Hosokawa P, Henderson W, Paulson E, Spitz F, Hamilton B and Hall B (2009) Time-of-day effects on surgical outcomes in the private sector: A retrospective cohort study. *Journal of the American College of Surgeons* 209 (4) 434-445.

Kendall D (1953) Stochastic processes occurring in the theory of queues and their analysis by the method of imbedded Markov chains. *Annals of Mathematical Statistics* 24 338-354.

Kenzora J, McCarthy R, Lowell J and Sledge C (1984) Hip fracture mortality. Relation to age, treatment, preoperative illness, time of surgery, and complications. *Clinical Orthopaedics and Related Research* 186 45-56.

Kesmezacar H, Ayhan E, Unlu M, Seker A and Karaca S (2010) Predictors of mortality in elderly patients with an intertrochanteric or a femoral neck fracture. *Journal of Trauma* 68 (1) 153-158.

Khan R, Fernandez C, Kashif F, Shedden R and Diggory P (2002) Combined orthogeriatric care in the management of hip fractures: a prospective study. *Annals of The Royal College of Surgeons of England* 84 (2) 122-124.

Khan S, Kalra S, Khanna A, Thiruvengada M and Parker M (2009) Timing of surgery for hip fractures: A systematic review of 52 published studies involving 291,413 patients. *Injury* 40 (7) 692-697.

Khan S, Khanna A, Al-Salahi M and Parker M (2011) Does the place of fall influence the time to specialist treatment in patients sustaining hip fractures? A study of 4917 patients falling in four different settings. *Acta Orthopaedica Belgica* 77 (2) 197-202.

Khasraghi F, Lee E, Christmas C and Wenz J (2003) The economic impact of medical complications in geriatric patients with hip fracture. *Orthopedics* 26 (1) 49-53.

Khaw K, Reeve J, Luben R, Bingham S, Welch A, Wareham N, Oakes S and Day N (2004) Prediction of total and hip fracture risk in men and women by quantitative ultrasound of the calcaneus: EPIC-Norfolk prospective population study. *Lancet* 363 (9404) 197-202.

Khaw K, Wareham N, Bingham S, Welch A, Luben R and Day N (2008) Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *Public Library of Science* 5 (1) 39-47.

Khintchine A (1932) Mathematical theory of a stationary queue. *Matematicheskii Sbornik* 39 (4) 73-84.

Kiel D, Magaziner J, Zimmerman S, Ball L, Barton B, Brown K, Stone J, Dewkett D and Birge S (2007) Efficacy of a hip protector to prevent hip fracture in nursing home residents. *Journal of the American Medical Association* 298 413-422.

Klimek M, van Houdenhoven M and Ottens T (2008) Operating Theatres: Organization, Costs and Audit. In: Gullo, A (ed.) *Anaesthesia, Pain, Intensive Care and Emergency A.P.I.C.E.* Springer Milan.

Knight V and Harper P (2012) Modelling emergency medical services with phase-type distributions. *Health Systems* 1 (1) 58-68.

Komarasamy B, Forster M, Esler C, Harper W and Hall A (2007) Mortality following hip fracture surgery in patients with recent myocardial infarction. *Annals of the Royal College of Surgeons of England* 89 (5) 521-525.

Kondo A, Zieler B, Isokawa Y, Hagino H, Ito Y and Richerson M (2010a) Comparison of lengths of hospital stay after surgery and mortality in elderly hip fracture patients between Japan and the United States - the relationship between the lengths of hospital stay after surgery and mortality. *Disability and Rehabilitation* 32 (10) 826-835.

Kondo A, Zierler B and Hagino H (2010b) Relationship between the length of hospital stay after hip fracture surgery and ambulatory ability or mortality after discharge in Japan. *Japan Journal of Nursing Science* 7 (1) 96-107.

Kopp L, Edelmann K, Obruba P, Procházka B, Blstáková K and Dzupa V (2009) Mortality risk factors in the elderly with proximal femoral fracture treated surgically. *Acta Chirurgiae Orthopaedicae et Traumatologiae Cechoslovaca* 76 (1) 41-46.

Koval K, Maurer S, Su E, Aharonoff G and Zuckerman J (1999) The effects of nutritional status on outcome after hip fracture. *Journal of Orthopaedic Trauma* 13 (3) 164-169.

Koval K, Rust C and Spratt K (2011) The effect of hospital setting and teaching status on outcomes after hip fracture. *American Journal of Orthopedics* 40 (1) 19-28.

Krupka D and Sandberg W (2006) Operating room design and its impact on operating room economics. *Current Opinion in Anaesthesiology* 19 (2) 185-191.

Kutner M, Nachtsheim C and Neter J (2004) *Applied Linear Regression Models*. Fourth ed. New York, USA: McGraw-Hill / Irwin.

LaCroix A, Beck T, Cauley J, Lewis C, Bassford T, Jackson R, Wu G and Chen Z (2010) Hip structural geometry and incidence of hip fracture in postmenopausal women: what does it add to conventional bone mineral density? *Osteoporosis International* 21 (6) 919-929.

Lagergren M (1994) Disability development and the structure of care: some results from simulation of an area-based ssytem of long-term care for elderly people. *Health Policy* 29 (3) 229-246.

Lapierre S, Batson C and McCaskey S (1999) Improving on-time performance in health care organizations: a case study. *Health Care Management Science* 2 (1) 27-34.

Law A (2007) *Simulation Modelling and Analysis*. Fourth ed. New York, USA: McGraw-Hill.

Law A and Kelton W (2000) *Simulation Modelling and Analysis*. Third ed. New York, USA: McGraw-Hill.

Lawrence T, White C, Wenn R and Moran C (2005) The current hospital costs of treating hip fractures. *Injury* 36 (1) 88-91.

Lebowitz P (2003) Schedule the short procedure first to improve OR efficiency. *Association of periOperative Registered Nurses Journal* 78 (4) 651-654, 657-659.

Lee R, Donaldson C and Cook L (2003) The need for evolution in healthcare decision modelling. *Medical Care* 41 (9) 1024-1033.

Lee T, Marcantonio E, Mangione C, Thomas E, Polanczyk C, Cook E, Sugarbaker D, Donaldson M, Poss R, Ho K, Ludwig L, Pedan A and Goldman L (1999) Derivation and propective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation* 100 (10) 1043-1049.

Lefaivre K, Macadam S, Davidson D, Gandhi R, Chan H and Broekhuyse H (2009) Length of stay, mortality, morbidity and delay to surgery in hip fractures. *Journal of Bone and Joint Surgery, British Volume* 91-B (7) 922-927.

Lemos D, Nilssen E, Khatiwada B, Elder G, Reindl R, Berry G and Harvey E (2009) Dedicated orthopedic trauma theatres: effect on morbidity and mortality in a single trauma centre. *Canadian Journal of Surgery* 52 (2) 87-91.

Lenze E, Skidmore E, Dew M, Butters M, Rogers J, Begley A, Reynolds 3rd C and Munin M (2007) Does depression, apathy or cognitive impairment reduce the benefit of inpatient rehabilitation facilities for elderly hip fracture patients? *General Hospital Psychiatry* 29 (2) 141-146.

Leong G, Wilson J and Charlett A (2006) Duration of operation as a risk factor for surgical site infection: comparison of English and US data. *Journal of Hospital Infection* 63 (3) 255-262.

Leung F, Blauth M and Bavonratanavech S (2010a) Surgery for fragility hip fracture-streamlining the process. *Osteoporosis International* 21 (S4) S519-S521.

Leung F, Lau T, Kwan K, Chow S and Kung A (2010b) Does timing of surgery matter in fragility hip fractures? *Osteoporosis International* 21 (S4) S529-S534.

Li W, Shi D and Chao X (1997) Reliability analysis of M/G/1 queueing systems with server breakdowns and vacations. *Journal of Applied Probability* 34 (2) 546-555.

Liebetrau A (1983) Measures of association.*Quantitative Applications in the Social Sciences*. Newbury Park, CA, USA: Sage Publications.

Lin H and Xiraxagar S (2006) Seasonality of hip fractures and estimates of season-attributable effects: a multivariate ARIMA analysis of population-based data. *Osteoporosis International* 17 (6) 795-806.

Lin W, Wen C, Jiang C, Hou S, Chen C and Lin J (2011) Risk factors for hip fracture sites and mortality in older adults. *Journal of Trauma* 71 (1) 191-197.

Lindsay R (1995) The burden of osteoporosis: cost. *American Journal of Medicine* 98 9S-11S.

Little J (1961) A proof for the queuing formula: $L = \lambda W$. *Operations Research* 9 (3) 383-387.

Lucas C, Buechter K, Coscia R, Hurst J, Meredith J, Middleton J, Rinker C, Tuggle D, Viahos A and Wilberger J (2001) Mathematical modeling to define optimum operating room staffing needs for trauma centers. *Journal of the American College of Surgeons* 192 (5) 559-565.

Luger T, Kammerlander C, Gosch M, Luger M, Kammerlander-Knauer U, Roth T and Kreutziger J (2010) Neuroaxial versus general anaesthesia in geriatric patients for hip fracture surgery: does it matter? *Osteoporosis International* 21 (S4) S555-S572.

López-Fidalgo J and Sanchez G (2005) Statistical criteria to establish bioassay programs. *Health Physics* 89 (4) 333-338.

Madan K (1999) An M/G/1 queue with optional deterministic server vacations. *Metron - International Journal of Statistics* LVII (3-4) 83-95.

Maggi S, Siviero P, Wetle T, Besdine R, Saugo M, Crepaldi G and Hip Fracture Study Group (2010) A multicenter survey on profile of care for hip fracture: predictors of mortality and disability. *Osteoporosis International* 21 (2) 223-231.

Mahoney F and Barthel D (1965) Functional evaluation: The Barthel Index. *Maryland State Medical Journal* 14 61-65.

Majumdar S, Beaupre L, Johnston D, Dick D, Cinats J and Jiang H (2006) Lack of association between mortality and timing of surgical fixation in elderly patients with hip fracture: results of a retrospective population-based cohort study. *Medical Care* 44 (6) 552-559.

Malik M, Harwood P, Diggle P and Khan S (2004) Factors affecting rates of infection and nonunion in intramedullary nailing. *Journal of Bone and Joint Surgery, British Volume* 86-B (4) 556-560.

MAPLE (1981-2010©) 14.00. Waterloo, Ontario, Canada: Waterloo Maple Inc.

Marazzi A, Paccaud F, Ruffieux C and Beguin C (1998) Fitting the distributions of length of stay by parametric models. *Medical Care* 36 (6) 915-927.

March L, Cameron I, Cumming R, Chamberlain A, Schwarz J, Brnabic A, O'Meara P, Taylor T, Riley S and Sambrook P (2000) Mortality and morbidity after hip fracture: can evidence based clinical pathways make a difference? *Journal of Rheumatology* 27 (9) 2227-2231.

Marinella M and Markert R (2009) Clinical predictors of prolonged hospitalization in patients with hip fracture. *Journal of Clinical Outcomes Management* 16 (10) 453-458.

Marjamaa R, Torkki P, Hirvensalo E and Kirvelä O (2009) What is the best workflow for an operating room? A simulation study of five scenarios. *Health Care Management Science* 12 (2) 142-146.

Marks R (2010) Hip fracture epidemiological trends, outcomes, and risk factors, 1970-2009. *International Journal of General Medicine* 3 1-17.

Marshall A, McClean S, Shapcott C and Millard P (2002) Modelling patient duration of stay to facilitate resource management of geriatric hospitals. *Health Care Management Science* 5 (4) 313-319.

Marshall A and Shaw B (2008) Modeling the survival of hip fracture patients using a conditional phase-type distribution. *Proceedings of the 21st IEEE International Symposium on Computer-Based Medical Systems* 518-523.

Maxwell M, Moran C and Moppett I (2008) Development and validation of a preoperative scoring system to predict 30 day mortality in patients undergoing hip fracture surgery. *British Journal of Anaesthesia* 101 (4) 511-517.

McClean S and Millard P (1993) Patterns of length of stay after admission in geriatric medicine: An event history approach. *Journal of the Royal Statistical Society, Series D (The Statistician)* 42 (3) 263-274.

McDonald S, Hetrick S and Green S (2004) Pre-operative education for hip or knee replacement. *Cochrane Database of Systematic Reviews* 1 CD003526.

McIntosh C, Dexter F and Epstein R (2006) The impact of service-specific staffing, case scheduling, turnovers, and first-case starts on anesthesia group and operating room productivity: A tutorial using data from an Australian hospital. *Anesthesia & Analgesia* 103 (6) 1499-1516.

MedlinePlus (2012) Medical dictionary. http://www.nlm.nih.gov/medlineplus/mplusdictionary.html.

Mehdi J (2002) *Stochastic Models in Queueing Theory*. Second ed. Boston, USA: Academic Press.

Melton III L (1993) Hip fractures: a worldwide problem today and tomorrow. *Bone* 14 (Supp. 1) S1-S8.

Menard S (2001) *Applied Logistic Regression Analysis*. Thousand Oaks, CA, USA: Sage Publications.

Merchant R, Lui K, Ismail N, Wong H and Sitoh Y (2005) The relationship between postoperative complications and outcomes after hip fracture surgery. *Annals Academy of Medicine Singapore* 34 (2) 163-168.

Merriam-Webster (2012) Medical dictionary. http://www.merriam-webster.com/medical/.

Meyer H, Tverdal A, Falch J and Pedersen J (2000) Factors associated with mortality after hip fracture. *Osteoporosis International* 11 (3) 228-232.

Michel J, Klopfenstein C, Hoffmeyer P, Stern R and Grab B (2002) Hip fracture surgery: is the pre-operative American Society of Anesthesiologists (ASA) score of predictor of functional outcome? *Aging Clinical and Experimental Research* 14 (5) 389-394.

Mielczarek B and Uziałko-Mydlikowska J (2012) Application of computer simulation modeling in the health care sector: a survey. *Simulation* 88 (2) 197-216.

Miller C (1978) Survival and ambulation following hip fracture. *Journal of Bone and Joint Surgery* 60 930-934.

Minh D (1988) Transient solutions for some exhaustive M/G/1 queues with generalized independent vacations. *European Journal of Operational Research* 36 (2) 197-201.

Modarres R, Ouarda T, Vanasse A, Orzanco M and Gosselin P (2012) Modeling seasonal variation of hip fracture in Montreal, Canada. *Bone* 50 (4) 909-916.

Mohamed K, Copeland G, Boot D, Casserley H, Shackleford I, Sherry P and Stewart G (2002) An assessment of the POSSUM system in orthopaedic surgery. *Journal of Bone and Joint Surgery, British Volume* 84-B (5) 735-739.

Moller J, Cluitmans P, Rasmussen L, Houx P, Rasmussen H, Canet J, Rabbitt P, Jolles J, Larsen K, Hanning C, Langeron O, Johnson T, Lauven P, Kristensen P, Biedler A, Van Beem H, Fraidakis O, Silverstein J, Beneken J and Gravenstein J (1998) Long-term postoperative cognitive dysfunction in the elderly: ISPOCD1 study. *Lancet* 351 (9106) 857-861.

Moore S (2003) Capacity planning - modelling unplanned admissions in the UK NHS. *International Journal of Health Care Quality Assurance* 16 (4) 165-172.

Moppett I (2010) Letter to the Editor: New equations for predicting postoperative risk in patients with hip fracture. *Clinical Orthopaedics and Related Research* 468 (6) 1705.

Moran C, Wenn R, Sikand M and Taylor A (2005) Early mortality after hip fracture: is delay before surgery important? *Journal of Bone and Joint Surgery* 87 (3) 483-489.

Moschopoulos P (1985) The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics* 37 (1) 541-544.

Murday A, Hamilton L, Magee P and Hyde J (2000) *The conflict between service and training in cardiothoracic surgery: a report of a short-life working group of the Society of Cardiothoracic Surgeons of Great Britain and Ireland*. London, UK: Society of Cardiothoracic Surgeons of Great Britain and Northern Ireland.

Muñoz E, Tortella B, Jaker M, Sakmyster M and Kanofsky P (1994) Surgical resource consumption in an academic health consortium. *Surgery* 115 (4) 411-416.

Myers R (1990) *Classical and Modern Regression with Applications*. Second ed. Duxbury Press.

NAFW (2005) WHC [035] Hospital discharge planning guidance.

Nagelkerke N (1991) A note on a general definition of the coefficient of determination. *Biometrika* 78 (3) 691-692.

Naglie G, Tansey C, Kirkland J, Ogilvie-Harris D, Detsky A, Etchells E, Tomlinson G, O'Rourke K and Goldlist B (2002) Interdisciplinary inpatient care for elderly people with hip fracture: a randomized controlled trial. *Canadian Medical Association Journal* 167 (1) 25-32.

NCEPOD (2004) National Confidential Enquiry into Patient Outcome and Death. *Scoping Our Practice*. London: NCEPOD.

Neary W, Heather B and Earnshaw J (2003) The Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM). *British Journal of Surgery* 90 (5) 157-165.

Neary W, Prytherch D, Foy C, Heather B and Earnshaw J (2007) Comparison of different methods of risk stratification in urgent and emergency surgery. *British Journal of Surgery* 94 (10) 1300-1305.

Nelson B (1992) Statistical analysis of simulation results. In: Salvendy, G (ed.) *Handbook of Industrial Engineering*. (Second edn.) New York, USA: Wiley.

Nematy M, Hickson M, Brynes A, Ruxton C and Frost G (2006) Vulnerable patients with a fractured neck of femur: nutritional status and support in hospital. *Journal of Human Nutrition and Dietetics* 19 (3) 209-218.

NHS (2005) *Hospital Episode Statistics online.* http://www.hesonline.nhs.uk Accessed: December 2008

NHS (2008) *Reducing cancelled operations.* http://www.institute.nhs.uk/quality_and_service_improvement_tools/quality_and_service_improvement_tools/cancelled_operations.html Accessed: December 2009

NHS (2010) *Menopause.* http://www.nhs.uk/conditions/menopause/Pages/Introduction.aspx Accessed: June 2010

NHS (2012a) Data model and dictionary. http://www.connectingforhealth.nhs.uk/systemsandservices/data/nhsdmds/dmd.

NHS (2012b) *Lean thinking.* http://www.institute.nhs.uk/building_capability/general/lean_thinking.html Accessed: February 2012

NHS Management Executive (1989) The Management and Utilisation of Operating Departments (The Bevan Report). London: HMSO.

NHS Wales (2010) *Cardiff and Vale University Health Board - Our Board.* www.wales.nhs.uk Accessed: January 2011

Nicholl C (2009) Diagnosis of dementia. *British Medical Journal* 338 b1176.

Nichols K, Prytherch D, Fancourt M, Gilkison W, Kyle S and Mosquera D (2008) Risk-adjusted general surgical audit in octogenarians. *ANZ Journal of Surgery* 78 (11) 990-994.

Niteesh K, Choudhry M, Fletcher R and Soumerai S (2005) Systematic review: the relationship between clinical experience and quality of health care. *Annals of Internal Medicine* 142 (4) 260-273.

Nixon M, Jackson B, Varghese P, Jenkins D and Taylor G (2006) Methicillin-resistant Staphylococcus aureus on orthopaedic wards: incidence, spread, mortality, cost and control. *Journal of Bone and Joint Surgery, British Volume* 88 (6) 812-817.

NLIAH (2005) *Integrated Care Pathways - A guide to good practice*. Llanharan, Wales: National Leadership and Innovation Agency for Healthcare.

NOS (2007) *Osteoporosis facts and figures*. The National Osteoporosis Society.

Nosek R and Wilson J (2001) Queuing theory and customer satisfaction: A review of terminology, trends, and applications to pharmacy practice. *Hospital Pharmacy* 36 (3) 275-279.

Novack V, Jotkowitz A, Etzion O and Porath A (2007) Does delay in surgery after hip fracture lead to worse outcomes? A multicenter study. *International Journal for Quality in Health Care* 19 (3) 170-176.

O'Daly B, Walsh J, Quinlan J, Falk G, Stapleton R, Quinlan W and O'Rourke S (2010) Serum albumin and total lymphocyte count as predictors of outcome in hip fracture. *Clinical Nutrition* 29 (1) 89-93.

O'Hara D, Duff A, Berlin J, Poses R, Lawrence V, Huber E, Noveck H, Strom B and Carson J (2000) The effect of anaesthetic technique on postoperative outcomes in hip fracture repair. *Anaesthesiology* 92 (4) 947-957.

ONS (2009) Life expectancy at birth (years) by local authority in England and Wales, with 95% confidence limits, 1991-1993 to 2006-2008.

ONS (2010) *Regional trends.* www.statistics.gov.uk Accessed: August 2010

ONSMD.com (2012) *Osteoarthritis of the hip.* http://www.onsmd.com/osteoarthritis-of-the-hip/ Accessed: January 2012

Orosz G, Hannan E, Magaziner J, Koval K, Gilbert M, Aufses A, Straus E, Vespe E and Siu A (2002) Hip fracture in the older patient: reasons for delay in hospitalization and timing of surgical repair. *Journal of the American Geriatrics Society* 50 (8) 1336-1340.

Orosz G, Magaziner J, Hannan E, Morrison R, Koval K, Gilbert M, McLaughlin M, Halm E, Wang J, Litke A, Silberzweig S and Siu A (2004) Association of timing of surgery for hip fracture and patient outcomes. *Journal of the American Medical Association* 291 (14) 1738-1743.

Osnes E, Lofthus C, Meyer H, Falch J, Nordsletten L, Cappelen I and Kristiansen I (2004) Consequences of hip fracture on activities of daily life and residential needs. *Osteoporosis International* 15 (7) 567-574.

Ostrum R (2003) Orthopedic trauma surgeon: distinction or extinction? *American Journal of Orthopedics* 32 (9S) 3-4.

Overdyk F, Harvey S, Fishman R and Shippey F (1998) Successful strategies for improving operating room efficiency at academic institutions. *Anesthesia & Analgesia* 86 (4) 896-906.

Palm H, Krasheninnikoff M, Holck K, Lemser T, Foss N, Jacobsen S, Kehlet H and Gebuhr P (2012) A new algorithm for hip fracture surgery. *Acta Orthopaedica* 83 (1) 26-30.

Pandit J, Westbury S and Pandit M (2007) The concept of surgical operating list 'efficiency': a formula to describe the term. *Anaesthesia* 62 (9) 895-903.

Papadimitropoulos E, Coyte P, Josse R and Greenwood C (1997) Current and projected rates of hip fracture in Canada. *Canadian Medical Journal Association* 157 (10) 1357-1363.

Parker M (2004) Care pathways for hip fractures: a useful tool or passing fashion? *Age and Ageing* 33 (2) 93-94.

Parker M and Gurusamy K (2005) Modern methods of treating hip fractures. *Disability and Rehabilitation* 27 (18-19) 1045-1051.

Parker M and Johansen A (2006) Hip fracture. *British Medical Journal* 333 (7557) 27-30.

Parker M and Pryor G (1992) The timing of surgery for proximal femoral fractures. *Journal of Bone and Joint Surgery, British Volume* 74 (2) 203-205.

Parker M, Todd C, Palmer C, Camilleri-Ferrante C, Freeman C, Laxton C, Payne B and Rushton N (1998) Inter-hospital variations in length of hospital stay following hip fracture. *Age and Ageing* 27 333-337.

Parvizi J, Holiday A, Ereth M and Lewallen D (1999) Sudden death during primary hip arthroplasty. *Clinical Orthopaedics & Related Research* 369 39-48.

Patel M and Martin F (2008) Why don't elderly hospital inpatients eat adequately? *Journal of Nutrition, Health & Aging* 12 (4) 227-231.

Pathak G, Parker M and Pryor G (1997) Delayed diagnosis of femoral neck fractures. *Injury* 28 (4) 299-301.

Pautex S, Jacques M, Sant A, Herrmann F and Chevalley T (2005) A short comprehensive assessment to predict outcome of elderly patients after hip fracture. *Aging Clinical and Experimental Research* 17 116-120.

Pawlikowski K (1990) Steady-state simulation of queueing processes: a survey of problems and solutions. *Computing Surveys* 22 (2) 123-170.

Peberdy M, Ornato J, Larkin G, Braithwaite R, Kashner T, Carey S, Meaney P, Cen L, Nadkarni V, Praestgaard A and Berg R (2008) Survival from in-hospital cardiac arrest during nights and weekends. *Journal of the American Medical Association* 299 (7) 785-792.

Pedrazzoni M, Alfano F, Malvi C, Ostanello F and Passeri M (1993) Seasonal variation in the incidence of hip fractures in Emilia-Romagna and Parma. *Bone* 14 (Supp. 1) S57-S63.

Persson M and Persson J (2010) Analysing management policies for operating room planning using simulation. *Health Care Management Science* 13 (2) 182-191.

Peskun C, Walmsley D, Waddell J and Schemitsch E (2012) Effect of surgeon fatigue on hip and knee arthroplasty. *Canadian Journal of Surgery* 55 (2) 81-86.

Petermann A, Tadvi J and Calthorpe D (2003) Does early surgery improve the outcome of hip fracture surgery? A prospective study. *Journal of Bone and Joint Surgery, British Volume* 85-B (Supp. I) 37.

Pfeiffer E (1975) A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society* 23 (10) 433-441.

Pidd M (2003) *Tools for Thinking: Modelling in Management Science*. Second ed. John Wiley & Sons.

Pioli G, Barone A, Giusti A, Oliveri M, Pizzonia M, Razzano M and Palummeri E (2006) Predictors of mortality after hip fracture: results from 1-year follow-up. *Aging Clinical and Experimental Research* 18 (5) 381-387.

Piziak V and Rajab M (2011) An effective team approach to improve postoperative hip fracture care. *Journal of Orthopaedic Trauma* 25 (2) 72-75.

Pollaczek F (1930) Über eine Aufgabe der Wahrscheinlichkeitstheorie. I. *Mathematische Zeitschrift* 32 (1) 64-100.

Pollard T, Newman J, Barlow N, Price J and Willett K (2006) Deep wound infection after proximal femoral fracture: consequences and costs. *Journal of Hospital Infection* 63 (2) 133-139.

Porter R, Miller C, Grainger D and Palmer S (1990) Prediction of hip fracture in elderly women: a prospective study. *British Medical Journal* 301 (6753) 638-641.

Prytherch D, Sirl J, Weaver P, Schmidt P, Higgins B and Sutton G (2003) Towards a national clinical minimum data set for general surgery. *British Journal of Surgery* 90 (10) 1300-1305.

Prytherch D, Whiteley M, Higgins B, Weaver P, Prout W and Powell S (1998) POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. *British Journal of Surgery* 85 (9) 1217-1220.

Pugh I (2011) *Osteoporosis risk factors.* www.bonematters.co.uk Accessed: January 2011

Radcliff T, Henderson W, Stoner T, Khuri S, Dohm M and Hutt E (2008) Patient risk factors, operative care, and outcomes among older community-dwelling male veterans with hip fracture. *Journal of Bone and Joint Surgery* 90 (1) 34-42.

Rademakers L, Vainas T, van Zutphen S, Brink P and van Helden S (2007) Pressure ulcers and prolonged hospital stay in hip fracture patients affected by time-to-surgery. *European Journal of Trauma and Emergency Surgery* 33 (3) 238-244.

Rae H, Harris I, McEvoy L and Todorova T (2007) Delay to surgery and mortality after hip fracture. *ANZ Journal of Surgery* 77 (10) 889-891.

Ramanathan T, Moppett I, Wenn R and Moran C (2005) POSSUM scoring for patients with fractured neck of femur. *British Journal of Anaesthesia* 94 (4) 430-433.

Ranhoff A, Martinsen M, Holvik K and Solheim L (2011) Use of warfarin is associated with delay to surgery for hip fracture in older patients. *Hospital Practice* 39 (1) 37-40.

Rapp K, Becker C, Lamb S, Icks A and Klenk J (2008) Hip fractures in institutionalized elderly people: incidence rates and excess mortality. *Journal of Bone and Mineral Research* 23 (11) 1825-1831.

Raw D, Beattie J and Hunter J (2003) Anaesthesia for spinal surgery in adults. *British Journal of Anaesthesia* 91 (6) 886-904.

RCN (2004) *Falls (full NICE guideline). Clinical guideline 21*. National Collaborating Centre for Nursing and Supportive Care, Royal College of Nursing.

RCOA (2010) Information for patients: Risks associated with your anaesthetic.

Rethnam U, Ramesh B and Clay N (2009) Orthopaedic trauma service utilisation - Is there more than what meets the eye? *Injury Extra* 40 (10) 187.

Ricci W, Gallagher B, Brandt A, Schwappach J, Tucker M and Leighton R (2009) Is after-hours orthopaedic surgery associated with adverse outcomes? A prospective comparative study. *Journal of Bone and Joint Surgery* 91 (9) 2067-2072.

Ricketts D, Hartley J, Patterson M, Harries W and Hitchin D (1994) An orthopaedic theatre timings survey. *Annals of The Royal College of Surgeons of England* 76 (3) 200-204.

Ridgeway S, Wilson J, Charlett A, Kafatos G, Pearson A and Coello R (2005) Infection of the surgical site after arthroplasty of the hip. *Journal of Bone and Joint Surgery, British Volume* 87 (6) 844-850.

Rix T and Bates T (2007) Pre-operative risk scores for the prediction of outcome in elderly people who require emergency surgery. *World Journal of Emergency Surgery* 2 16.

Roberts H, Pickering R, Onslow E, Clancy M, Powell J, Roberts A, Hughes K, Coulson D and Bray J (2004) The effectiveness of implementing a care pathway for femoral neck fracture in older people: a prospective controlled before and after study. *Age and Ageing* 33 (2) 178-184.

Robertson M, Althoff R, Hafez A and Pauls D (2008) Principal components analysis of a large cohort with Tourette syndrome. *The British Journal of Psychiatry* 193 (1) 31-36.

Robinson S (2004) *Simulation - The Practice of Model Development and Use*. Chichester, UK: John Wiley & Sons, Inc.

Robinson S (2007) A statistical process control approach to selecting a warm-up period for a discrete-event simulation. *European Journal of Operational Research* 176 (1) 332-346.

Roche J, Wenn R, Sahota O and Moran C (2005) Effect of comorbidities and postoperative complications on mortality after hip fracture in elderly people: prospective observational cohort study. *British Medical Journal* 331 (7529) 1374.

Rodriguez-Fernandez P, Adarraga-Cansino D and Carpintero P (2011) Effects of delayed hip fracture surgery on mortality and morbidity in elderly patients. *Clinical Orthopaedics and Related Research* 469 (11) 3218-3221.

Ryan J, Ghani M, Staniforth P, Bryant G and Edwards S (1996) "Fast tracking" patients with a proximal femoral fracture. *Journal of Accident & Emergency Medicine* 13 (2) 108-110.

Röder F, Schwab M, Aleker T, Mörike K, Thon K and Klotz U (2003) Proximal femur fracture in older patients - rehabilitation and clinical outcome. *Age and Ageing* 32 (1) 74-80.

Sanders D, Grayling M and Lillie H (2008) Defining efficiency requires more fidelity. *Anaesthesia* 63 (2) 204-205.

Sargent R (2000) Verification, validation and accreditation of simulation models. In Joines, J, Barton, R, Kang, K and Fishwick, P (Eds.), *Proceedings of the Winter Simulation Conference* (pp. 50-59). Orlando, USA.

SAS (2002-2003) 9.1.3. Cary, North Carolina, USA: SAS Institute Inc.

Sawka A, Boulos P, Beattie K, Thabane L, Papaioannou A, Gafni A, Cranney A, Zytaruk N, Hanley D and Adachi J (2005) Do hip protectors decrease the risk of hip fracture in institutional and community-dwelling elderly? A systematic review and meta-analysis of randomized controlled trials. *Osteoporosis International* 16 (12) 1461-1474.

Schellenberg J, Victora C, Mushi A, de Savigny D, Schellenberg D, Mshinda H, Bryce J and Tanzania Integrated Management of Childhood Illness MCE Baseline Household Survey Study Group (2003) Inequities among the very poor: health care for children in rural southern Tanzania. *Lancet* 361 (9357) 561-566.

Schilling P, Campbell Jr D, Englesbe M and Davis M (2010) A comparison of in-hospital mortality risk conferred by high hospital occupancy, differences in nurse staffing levels, weekend admission, and seasonal influenza. *Medical Care* 48 (3) 224-232.

Schilling P, Goulet J and Dougherty P (2011) Do higher hospital-wide nurse staffing levels reduce in-hospital mortality in elderly patients with hip fractures: A pilot study. *Clinical Orthopaedics and Related Research* 469 (10) 2932-2940.

Schmidt A, Leighton R, Parviz J, Sems A and Berry D (2009) Optimal arthoplasty for femoral neck fractures: is total hip arthoplasty the answer? *Journal of Orthopaedic Trauma* 23 (6) 428-433.

Schofield W, Rubin G, Piza M, Lai Y, Sindhusake D, Fearnside M and Klineberg P (2005) Cancellation of operations on the day of intended surgery at a major Australian referral hospital. *Medical Journal of Australia* 182 (12) 612-615.

Scholl M and Kleinrock L (1983) On the M/G/1 queue with rest periods and certain service-independent queueing disciplines. *Operations Research* 31 (4) 705-719.

Schruben L, Singh H and Tierney L (1983) Optimal tests for initialization bias in simulation output. *Operations Research* 31 (6) 1167-1178.

Schwartz S and Yeh Y (1982) On the distribution function and moments of power sums with lognormal components. *Bell System Technical Journal* 61 1441-1462.

Semel J, Gray J, Ahn H, Nasr H and Chen J (2010) Predictors of outcome following hip fracture rehabilitation. *Physical Medicine and Rehabilitation* 2 (9) 799-805.

Sernbo I and Johnell O (1993) Consequences of a hip fracture: a prospective study over 1 year. *Osteoporosis International* 3 (3) 148-153.

Sexson S and Lehner J (1987) Factors affecting hip fracture mortality. *Journal of Orthopaedic Trauma* 1 (4) 298-305.

Shabat S, Heller E, Mann G, Gepstein R, Fredman B and Nyska M (2003) Economic consequences of operative delay for hip fractures in a non-profit institution. *Orthopedics* 26 (12) 1197-1199.

Shah M, Aharonoff G, Wolinsky P, Zuckerman J and Koval K (2001) Outcome after hip fracture in individuals ninety years of age and older. *Journal of Orthopaedic Trauma* 15 (1) 34-39.

SHFA (2008) *Clinical decision-making: is the patient fit for theatre?* Edinburgh, UK: ISD Scotland Publications.

Shiga T, Wajima Z and Ohe Y (2008) Is operative delay associated with increased mortality of hip fracture patients? Systematic review, meta-analysis, and meta-regression. *Canadian Journal of Anesthesia* 55 (3) 146-154.

Shukla S, Nixon M, Acharya M, Korim M and Pandey R (2009) Incidence of MRSA surgical-site infection in MRSA carriers in an orthopaedic trauma unit. *Journal of Bone and Joint Surgery, British Volume* 91 (2) 225-228.

Siegmeth A, Gurusamy K and Parker M (2005) Delay to surgery prolongs hospital stay in patients with fractures of the proximal femur. *Journal of Bone and Joint Surgery, British Volume* 87-B (8) 1123-1126.

SIGN (2002) *Prevention and management of hip fracture in older people*. Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network.

Simul8 (1993-2010©) Version 17.0 Build 2191. Simul8 Corporation.

Simunovic N, Devereaux P and Bhandari M (2011) Surgery for hip fractures: does surgical delay affect outcomes? *Indian Journal of Orthopaedics* 45 (1) 27-32.

Simunovic N, Devereaux P, Sprague S, Guyatt G, Schemitsch E, Debeer J and Bhandari M (2010) Effect of early surgery after hip fracture on mortality and complications: systematic review and meta-analysis. *Canadian Medical Association Journal* 182 (15) 1609-1616.

Sipilä S, Heikkinen E, Cheng S, Suominen H, Saari P, Kovanen V, Alén M and Rantanen T (2006) Endogenous hormones, muscle strength, and risk of fall-related fractures in older women. *The Journals of Gerontology: Series A* 61 (1) 92-96.

Siris E, Miller P, Barrett-Connor E, Faulkner K, Wehren L, Abbott T, Berger M, Santora A and Sherwood L (2001) Identification and fracture outcomes of undiagnosed low bone mineral density in postmenopausal women. *Journal of the American Medical Association* 286 (22) 2815-2822.

Smektala R, Endres H, Dasch B, Maier C, Trampisch H, Bonnaire F and Pientka L (2008) The effect of time-to-surgery on outcome in elderly patients with proximal femoral fractures. *BMC Musculoskeletal Disorders* 9 171.

Smith E, Parvizi J and Purtill J (2011) Delayed surgery for patients with femur and hip fractures - risk of deep venous thrombosis. *Journal of Trauma* 70 (6) E113-E116.

Smith N, Harris K, Salem K and Kurian J (2008) Integrated care pathway for hip fractures - a help or a hindrance? *Journal of Integrated Care Pathways* 12 88.

Småbrekke A, Espehaug B, Havelin L and Furnes O (2004) Operating time and survival of primary total hip replacements: An analysis of 31 745 primary cemented and uncemented total hip replacements from local hospitals reported to the Norweigan Arthroplasty Register 1987-2001. *Acta Orthopaedica Scandinavica* 75 (5) 524-532.

Snedecor G and Cochran W (1991) *Statistical Methods*. Eighth ed. Wiley-Blackwell.

Sokolovic E, Biro P, Wyss P, Werthemann C, Haller U, Spahn D and Szucs T (2002) Impact of the reduction of anaesthesia turnover time on operating room efficiency. *European Journal of Anaesthesiology* 19 (8) 560-563.

Spangler W, Strum D, Vargas L and May J (2004) Estimating procedure times for surgeries by determining location parameters for the Lognormal model. *Health Care Management Science* 7 (2) 97-104.

Spratt S (1998) *Heuristics for the startup problem.* University of Virginia.

SPSS© (2007) 16.0. Chicago, Illinois, USA: SPSS Inc.

Stat::Fit (1995-2001©) Version 2.0 10.6 Simul8. South Kent, Connecticut, USA: Geer Mountain Software Corporation.

Stepaniak P, Heij C, Mannaerts G, de Quelerij M and de Vries G (2009a) Modeling procedure and surgical times for current procedural terminology-anesthesia-surgeon combinations and evaluation in terms of case-duration prediction and operating room efficiency: A multicenter study. *Anesthesia & Analgesia* 109 (4) 1232-1245.

Stepaniak P, Mannaerts G, de Quelerij M and de Vries G (2009b) The effect of the Operating Room Coordinator's risk appreciation on operating room efficiency. *Anesthesia & Analgesia* 108 (4) 1249-1256.

Sterling R (2011) Gender and race/ethnicity differences in hip fracture incidence, morbidity, mortality, and function. *Clinical Orthopaedics and Related Research* 469 (7) 1913-1918.

Stewart S, Rosario M, Brown T, Carter A, Leckman J, Sukhodolsky D, Katsovitch L, King R, Geller D and Pauls D (2007) Principal components analysis of obsessive-compulsive disorder symptoms in children and adolescents. *Biological Psychiatry* 61 (3) 285-291.

Stroup D, Berlin J, Morton S, Olkin I, Williamson G, Rennie D, Moher D, Becker B, Sipe T and Thacker S (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *Journal of the American Medical Association* 283 (15) 2008-2012.

Stüber G (2000) *Principles of Mobile Communication* Second ed. Netherlands: Kluwer Academic Publishers.

Sund R, Riihimäki J, Mäkelä M, Vehtari A, Lüthje P, Huusko T and Häkkinen U (2009) Modeling the length of the care episode after hip fracture: does the type of fracture matter? *Scandinavian Journal of Surgery* 98 (3) 169-174.

Sutton R, Bann S, Brooks M and Sarin S (2002) The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit. *British Journal of Surgery* 89 (6) 763-768.

Svensson O, Strömberg L, Öhlén G and Lindgren U (1996) Prediction of the outcome after hip fracture in elderly patients. *Journal of Bone and Joint Surgery, British Volume* 78-B (1) 115-118.

Sweet S and Grace-Martin K (2008) *Data Analysis with SPSS*. Third ed. Allyn & Bacon.

Sweetnam D, Williams J and Britton D (1994) An audit of the effect of a 24-hour emergency operating theatre in a district general hospital. *Annals of the Royal College of Surgeons of England* 76 (2S) 56-58.

Szulc P, Duboeuf F, Schott A, Dargent-Molina P, Meunier P and Delmas P (2006) Structural determinants of hip fracture in elderly women: re-analysis of data from the EPIDOS study. *Osteoporosis International* 17 (2) 231-236.

Szyszkowicz S and Yanikomeroglu H (2009) Fitting the modified-power-lognormal to the sum of independent lognormals distribution. *IEEE Global Telecommunications Conference* 5823-5828.

Söderqvist A, Ekström W, Ponzer S, Pettersson H, Cederholm T, Dalén N, Hedström M, Tidermark J and Stockholm Hip Fracture Group (2009) Prediction of mortality in elderly patients with hip fractures: a two-year prospective study of 1,944 patients. *Gerontology* 55 (5) 496-504.

Tadj L (2003) On an M/G/1 quorum queueing system under the *T*-policy. *Journal of the Operational Research Society* 54 (5) 466-471.

Tadj L, Choudhury G and Rekab K (2012) A two-phase quorum queueing system with Bernoulli vacation schedule, setup, and N-policy for an unreliable server with delaying repair. *International Journal of Services and Operations Management* 12 (2) 139-164.

Takagi H (1991) *Queueing analysis: a foundation of performance evaluation. Volume I - Vacation and priority systems*. Amsterdam, The Netherlands: Elsevier.

Tang X, Luo Z and Gardiner J (2012) Modeling hospital length of stay by Coxian phase-type regression with heterogeneity. *Statistics in Medicine* 31 (14) 1502-1516.

Thakar C, Alsousou J, Hamilton T and Willett K (2010) The cost and consequences of proximal femoral fractures which require further surgery following initial fixation. *Journal of Bone and Joint Surgery, British Volume* 92 (12) 1669-1677.

The National Kidney Federation (2011) *Medical Info.* http://www.kidney.org.uk/Medical-Info/pd/pdmonitor.html Accessed: December 2011

Theis J-C (2006) Hip fractures: a deadly and silent epidemic. *New Zealand Medical Journal* 119 (1234) U1981.

Thomas S, Ord J and Pailthorpe C (2001) A study of waiting time for surgery in elderly patients with hip fracture and subsequent in-patient hospital stay. *Annals of the Royal College of Surgeons of England* 83 (1) 37-39.

Thomas T and Stevens R (1974) Social effects of fractures of the neck of the femur. *British Medical Journal* 3 (5928) 456-458.

Thyagarajan D, Sunderamoorthy D, Haridas S, Beck S, Praveen P and Johansen A (2009) MRSA colonisation in patients admitted with hip fracture: implications for prevention of surgical site infection. *Acta Orthopaedica Belgica* 75 (2) 252-257.

Tian N and Zhang Z (2006) *Vacation Queueing Models - Theory and Applications*. New York: Springer.

Tidermark J (2003) Quality of life and femoral neck fractures. *Acta Orthopaedica Scandinavica Supplementum* 74 (309) 1-42.

Torgerson D, Iglesias C and Reid D (2001) The economics of fracture prevention. In: Barlow, D H, Francis, R M and Miles, A (eds.) *The effective management of osteoporosis*.

Torkki P, Marjamaa R, Torkki M, Kallio P and Kirvelä O (2005) Use of anesthesia induction rooms can increase the number of urgent orthopedic cases completed within 7 hours. *Anesthesiology* 103 (2) 401-405.

Treasure T, Valencia O, Sherlaw-Johnson C and Gallivan S (2002) Surgical performance measurement. *Health Care Management Science* 5 (4) 243-248.

Trisoft (2006) Trisoft Press Release.

Trisoft (2009a) *TheatreMan - Theatre management system.* http://www.trisoft.co.uk/html/nhs-theatreman.html Accessed: August 2009

Trisoft (2009b) Theatreman: theatre and day surgery management.

Troosters T, Gosselink R and Decramer M (1999) Six minute walking distance in healthy elderly subjects. *European Respiratory Journal* 14 270-274.

Tucker J, Barone J, Cecere J, Blabey R and Rha C (1999) Using queueing theory to determine operating room staffing needs. *Journal of Trauma* 46 (1) 71-79.

Tuggle D, Lucas C, Coscia R, Hurst J, Meredith J, Middleton J, Rinker C, Vlahos A and Wilberger J (2010) Mathematical modeling to define optimum operating room staffing for pediatric patients in trauma centers. *Journal of Pediatric Surgery* 45 (12) 2431-2435.

Turrentine F, Wang H, Simpson V and Jones R (2006) Surgical risk factors, morbidity, and mortality in elderly patients. *Journal of the American College of Surgeons* 203 (6) 865-877.

Tyler D, Pasquariello C and Chen C-H (2003) Determining optimum operating room utilization. *Anesthesia & Analgesia* 96 (4) 1114-1121.

Urwin S, Parker M and Griffiths R (2000) General *versus* regional anaesthesia for hip fracture surgery: a meta-analysis of randomized trials. *British Journal of Anaesthesia* 84 (4) 450-455.

Utley M, Gallivan S, Davis K, Daniel P, Reeves P and Worrall J (2003a) Estimating bed requirements for an intermediate care facility. *European Journal of Operational Research* 150 (1) 92-100.

Utley M, Gallivan S and Jit M (2005) How to take variability into account when planning the capacity for a new hospital unit. In: Vissers, J and Beech, R (eds.) *Health Operations Management: Patient Flow Logistics in Health Care*. Abingdon, UK: Routledge.

Utley M, Gallivan S, Treasure T and Valencia O (2003b) Analytical methods for calculating the capacity required to operate an effective booked admissions policy for elective inpatient services. *Health Care Management Science* 6 97-104.

Van der Meer R, Rymaszewski L, Findlay H and Curran J (2005) Using OR to support the development of an integrated musculo-skeletal service. *Journal of the Operational Research Society* 56 (2) 162-172.

Van Hoang H, Silverstone F, Leventer S, Wolf-Klein G and Foley C (1998) The effect of nutritional status on length of stay in elderly hip fracture patients. *Journal of Nutrition Health and Aging* 2 (3) 159-161.

van Staa T, Dennison E, Leufkens H and Cooper C (2001) Epidemiology of fractures in England and Wales. *Bone* 29 (6) 517-522.

van Staa T, Selby P, Leufkens H, Lyles K, Sprafka J and Cooper C (2002) Incidence and natural history of Paget's disease of bone in England and Wales. *Journal of Bone and Mineral Research* 17 (3) 465-471.

Varshney M, Kumar A, Khan S and Rastogi S (2009) Functional and radiological outcome after delayed fixation of femoral neck fractures in pediatric patients. *Journal of Orthopaedics and Traumatology* 10 (4) 211-216.

Vasilakis C, Lecznarowicz D and Lee C (2008) Modelling Health Systems: delivering care to older patients with hip fracture. Strategic Promotion of Ageing Research Capacity (SPARC).

Vassilacopoulos G (1989) Testing for initialization bias in simulation output. *Simulation* 52 (4) 151-153.

Verbeek D, Ponsen K, Goslings J and Heetveld M (2008) Effect of surgical delay on outcome in hip fracture patients: a retrospective multivariate analysis of 192 patients. *International Orthopaedics* 32 (1) 13-18.

Vertelis A, Robertsson O, Tarasevicius S and Wingstrand H (2009) Delayed hospitalization increases mortality in displaced femoral neck fracture patients. *Acta Orthopaedica* 80 (6) 683-686.

Vestergaard P, Rejnmark L and Mosekilde L (2007) Has mortality after a hip fracture increased? *Journal of the American Geriatrics Society* 55 (11) 1720-1726.

Vidán M, Sánchez E, Gracia Y, Marañón E, Vaquero J and Serra J (2011) Causes and effects of surgical delay in patients with hip fracture: a cohort study. *Annals of Internal Medicine* 155 (4) 226-233.

Villar R, Allen S and Barnes S (1986) Hip fractures in healthy patients: operative delay versus prognosis. *British Medical Journal* 293 (6556) 1203-1204.

Wachtel R and Dexter F (2009) Reducing tardiness from scheduled start times by making adjustments to the operating room schedule. *Anesthesia & Analgesia* 108 (6) 1902-1909.

Wachtl S, Jakob R and Gautier E (2003) Ten-year patient and prosthesis survival after unipolar hip hemiarthroplasty in female patients over 70 years old. *The Journal of Arthroplasty* 18 (5) 587-591.

WAG (2008a) [003122] Mid-Year Population Estimates (2001 onwards), by local authority (single year of age, gender, Welsh LAs).

WAG (2008b) [013906] NHS beds by specialty.

WAG (2009a) [SB 31/2009] Population aged 50 and over in Wales, 1991 to 2031.

WAG (2009b) Wales's population: A demographic overview 2009. Cardiff, UK: Welsh Assembly Government.

WAG (2010a) *[007455] Small area population and Local Authority population data.* http://www.statswales.wales.gov.uk Accessed: December 2010

WAG (2010b) [013875] Delay reason by NHS organisation.

WAG (2010c) [013876] Rate of delay per 10,000 population over 75+ by LAA.

WAG (2010d) [013881] Delay stage by NHS organisation.

WAG (2010e) [013883] Delay length by NHS organisation.

WAG (2010f) *NHS beds by NHS region and Trust (NHS region, Trust, hospital, specialty).* http://www.statswales.wales.gov.uk/ Accessed: September 2010

WAG (2010g) [SDR 157/2010] Household projections for Wales (2008-based): Summary report. Cardiff, UK: Welsh Assembly Government.

Wales Audit Office (2006) Making better use of NHS day surgery in Wales. Cardiff: Wales Audit Office.

Waljee J, Greenfield L, Dimick J and Birkmeyer J (2006) Surgeon age and operative mortality in the United States. *Annals of Surgery* 244 (3) 353-362.

Walley G, Orendi J, Bridgman S, Davis B, Ahmed E and Maffulli N (2009) Methicillin resistant Staphylococcus aureus (MRSA) is not always caught on the orthopaedic ward. *Acta Orthopaedica Belgica* 75 (2) 245-251.

Wang K-H and Ke J-C (2000) A recursive method to the optimal control of an M/G/1 queueing system with finite capacity and infinite capacity. *Applied Mathematical Modelling* 24 (12) 899-914.

Weatherall M (2001) Does a delay in transfer to a rehabilitation unit for older people affect outcome after fracture of the proximal femur? *The New Zealand Medical Journal* 114 (1145) 547-579.

Wei N (2004) *Degenerative bone disease.* http://www.arthritis-treatment-and-relief.com/degenerative-bone-disease.html Accessed: October 2008

Welch P (1983) The statistical analysis of simulation results. In: Lavenberg, S (ed.) *Computer Performance Modeling Handbook*. New York, USA: Academic Press.

Weller I, Wai E, Jaglal S and Kreder H (2005) The effect of hospital type and surgical delay on mortality after surgery for hip fracture. *Journal of Bone and Joint Surgery, British Volume* 87-B (3) 361-366.

White Jnr K (1997) An effective truncation heuristic for bias reduction in simulation output. *Simulation* 69 (6) 323-334.

White Jnr K, Cobb M and Spratt S (2000) A comparison of five steady-state truncation heuristics for simulation. In Joines, J, Barton, R, Kang, K and Fishwick, P (Eds.), *Proceedings of the Winter Simulation Conference* (pp. 755-760). Orlando, USA.

WHO (2002) The World Health Report 2002 - reducing risks, promoting healthy life. World Health Organization.

WHO (2007) WHO Global Report on Falls Prevention in Older Age. Geneva, Switzerland: World Health Organization.

WHO (2008) Invest in your bones: Osteoporosis in the workplace. World Health Organization / International Osteoporosis Foundation.

Wilson J, Charlett A, Leong G, McDougall C and Duckworth G (2008) Rates of surgical site infection after hip replacement as a hospital performance indicator: analysis of data from the English mandatory surveillance system. *Infection Control and Hospital Epidemiology* 29 (3) 219-226.

Wood D, Ions G, Quinby J, Gale D and Stevens J (1992) Factors which influence mortality after subcapital hip fracture. *Journal of Bone and Joint Surgery, British Volume* 74 (2) 199-202.

Worthington D (1991) Hospital waiting list management models. *Journal of the Operational Research Society* 42 (10) 833-843.

Wright D, Blanckley S, Stewart G and Copeland G (2008) The use of orthopaedic POSSUM as an audit tool for fractured neck of femur. *Injury* 39 (4) 430-435.

Wu J, Mehta N and Zhang J (2005) A flexible Lognormal sum approximation method. *IEEE Global Telecommunications Conference* 6 3413-3417.

Wullink G, van Houdenhoven M, Hans E, van Oostrum J, van der Lans M and Kazemier G (2007) Closing emergency operating rooms improves efficiency. *Journal of Medical Systems* 31 (6) 543-546.

X-Rates© (2012) http://www.x-rates.com/d/GBP/EUR/hist2008.html Accessed: February 2012

Yadin M and Naor P (1963) Queueing systems with a removable service station. *Operational Research Quarterly* 14 (4) 393-405.

Yaghoubian A, Kaji A, Ishaque B, Park J, Rosing D, Lee S, Stabile B and de Virgilio C (2010) Acute care surgery performed by sleep deprived residents: Are outcomes affected? *Journal of Surgical Research (online)* 1-5.

Yamamoto S, Suto Y, Kawamura H, Hashizume T, Kakurai S and Sugahara S (1983) Quantitative gait evaluation of hip diseases using principal component analysis. *Journal of Biomechanics* 16 (9) 717-726.

Youde J, Husk J, Lowe D, Grant R, Potter J and Martin F (2009) The national clinical audit of falls and bone health: the clinical management of hip fracture patients. *Injury* 40 (11) 1226-1230.

Youm T, Aharonoff G, Zuckerman J and Koval K (2000) Effect of previous cerebrovascular accident on outcome after hip fracture. *Journal of Orthopaedic Trauma* 14 (5) 329-334.

Young W, Seigne R, Bright S and Gardner M (2006) Audit of morbidity and mortality following neck of femur fracture using the POSSUM scoring system. *New Zealand Medical Journal* 119 (1234) U1986.

Young Y, Xiong K and Pruzek R (2011) Longitudinal functional recovery after postacute rehabilitation in older hip fracture patients: the role of cognitive impairment and implications for long-term care. *Journal of the American Medical Directors Association* 12 (6) 431-438.

Yousri T, Khan Z, Chakrabarti D, Fernandes R and Wahab K (2011) Lean thinking: Can it improve the outcome of fracture neck of femur patients in a district general hospital? *Injury* 42 (11) 1234-1237.

Yücesan E (1993) Randomization tests for initialization bias in simulation output. *Naval Research Logistics* 40 643-663.

Zhang J and Yu K (1998) What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *Journal of the American Medical Association* 280 1690-1691.

Zhang Z, Vickson R and van Eenige M (1997) Optimal two-threshold policies in an M/G/1 queue with two vacation types. *Performance Evaluation* 29 (1) 63-80.

Zhou Z and Fan S (2010) Letter to the Editor: New equations for predicting postoperative risk in patients with hip fracture. *Clinical Orthopaedics and Related Research* 468 (4) 1181-1182.

Zuckerman J, Skovron M, Koval K, Aharonoff G and Frankel V (1995) Postoperative complications and mortality associated with operative delay in older patients who have a fracture of the hip. *Journal of Bone and Joint Surgery* 77-A (10) 1551-1556.

APPENDICES

APPENDIX A: PROBABILITY DISTRIBUTION FUNCTIONS

Distribution	Probability distribution function	Theoretical mean	Theoretical standard deviation
Beta	$f(t) = \dfrac{t^{\alpha-1}(1-t)^{\beta-1}}{\displaystyle\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,du}$	$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\sqrt{\alpha\beta}}{(\alpha+\beta)\sqrt{(\alpha+\beta+1)}}$
Binomial	$f(x) = {}^nC_x\, p^x(1-p)^{n-x}$	np	$\sqrt{np(1-p)}$
Erlang	$f(t) = \dfrac{k\mu(k\mu t)^{k-1}e^{-k\mu t}}{(k-1)!}$	$\dfrac{1}{\mu}$	$\dfrac{1}{\sqrt{k}\,\mu}$
Gamma	$f(t) = \dfrac{1}{\theta^k\,\Gamma(k)}t^{k-1}e^{-\frac{t}{\theta}}$	$k\theta$	$\sqrt{k}\,\theta$
Lognormal	$f(t) = \dfrac{1}{\sigma t\sqrt{2\pi}}e^{-\frac{(\ln t-\mu)^2}{2\sigma^2}}$	$e^{\mu+\frac{\sigma^2}{2}}$	$\sqrt{(e^{\sigma^2}-1)(e^{2\mu+\sigma^2})}$
Negative Exponential	$f(t) = \mu e^{-\mu t}$	$\dfrac{1}{\mu}$	$\dfrac{1}{\mu}$
Normal	$f(t) = \dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	μ	σ
Poisson	$f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$	λ	$\sqrt{\lambda}$
Uniform (continuous)	$f(t) = \begin{cases} \dfrac{1}{b-a} & a \le t \le b \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{1}{2}(a+b)$	$\dfrac{1}{\sqrt{12}}(b-a)$

APPENDIX B: VARIABLE LISTS

Table B3.2.1a: Supplementary information for variables taken from the Cardiff Hip Fracture Survey

Variable name	Description and values
Admdelay	Delay to admission, days between fracture and admission
Admfrom_d(+*number*)*	Place admitted from; 1 = Own home 4 = Nursing home 7 = Acute hospital 2 = Sheltered housing 5 = Permanent hospital inpatient 8 = Other 3 = Residential care 6 = Rehabilitation unit
Age	Age at admission in years
ASAnew_n	American Society of Anaesthesiologists (ASA) grade; 1 = ASA grade I or II 2 = ASA grade III 3 = ASA grade IV
Finlos	University Health Board length of stay (days)
Fractype_d(+*number*)*	Type of fracture; 1 = Undisplaced intracapsular 4 = Trochanteric, two fragment 2 = Displaced intracapsular 5 = Trochanteric, multi fragment 3 = Basocervical 6 = Subtrochanteric
Livealon	Patient living alone; 1 = Yes 2 = No 3 = Institutional care
Mentalst	Mental state on admission; 1 = Normal 2 = Known dementia 3 = Confusion
Mobility	Mobility score pre-fracture; 1 = Able to shop 2 = Able to get out but unable to shop 3 = Housebound
Opdelay	Operative delay; 0 = Operation within two days of admission 1 = Operation after 2 days of admission
Optypenew_d(+*number*)*	Type of operation; 1 = No operation / conservative treatment 3 = Screws 6 = Total hip arthroplasty 4 = Intramedullary nail 7 = Other 2 = Dynamic hip screw 5 = Hemiarthroplasty

Pathfrac_d*(+number)**	Pathological fracture; 1 = No 3 = Malignant primary bony tumour 5 = Paget's disease 2 = Malignant secondary bony tumour 4 = Bone cyst 6 = Other
SexM	Sex; 0 = Female 1 = Male
Side	Side of fracture; 0 = Right 1 = Left
Survival_ac	Indicator of survival at end of acute ward stay; 0 = Patient survives 1 = Patient does not survive
Survival_fin	Indicator of survival at end of UHB stay; 0 = Patient survives 1 = Patient does not survive
WAASP	WAASP (Weight, Appetite, Ability to eat, Stress factors, Pressure sores/wounds) category on admission 1 = WAASP score 1-2 2 = WAASP score 3-6 3 = WAASP score 7+
Walkaid0	Walking aids used pre-fracture 1 = None 2 = One aid (stick, crutch) 3 = Two aids 4 = Frame 5 = Wheelchair / bed-bound
Walking0	Walking ability pre-fracture 1 = Outside, alone 3 = Inside, alone 5 = Wheelchair / bed-bound 2 = Outside, with someone 4 = Inside, with someone
Wardlos	Acute ward length of stay (days)

* These variables comprise of several 'dummy' variables in order to account for their nominal status. For example, *admfrom_d1* takes a value of 1 if the patient was admitted from their own home or a value of 0 if they were not.

Table B7.1.2a: Fracture neck of femur operations and their associated OPCS-4 codes

OPCS-4 code	Description
W19.1C	Internal fixation subcapsular fracture A/O cannulated screws
W19.1D	Open reduction with internal fixation subtrochanteric fracture femur / dynamic hip screw
W19.1E	Open reduction with internal fixation intertrochanteric fracture / dynamic hip screw
W20.1B	Dynamic condylar screw and plate
W20.1H	Revision of dynamic condylar screw and plate
W24.21	Proximal femoral nail
W24.2A	Closed intramedullary nail (fully locked)
W24.2B	Closed intramedullary nail (locked proximally)
W24.2C	Closed intramedullary nail (locked distally)
W24.2D	Closed intramedullary nail (unlocked)
W24.2E	Closed intramedullary fixation with nancy nails
W37.1	Total prosthetic replacement of hip joint using cement
W39.4	Attention total hip replacement
W46.1	Primary replacement of head of femur using cement
W46.1B	Primary cemented hip – Thompson stem
W46.1C	Primary cemented bipolar hemiarthroplasty
W46.1E	Hemiarthroplasty
W46.1F	Logic cemented hemiarthroplasty
W47.1A	Uncemented hip – Austin Moore stem
W47.1B	Uncemented hemiarthroplasty hip – Austin Moore
W57.4A	Girdlestone's procedure

APPENDIX C: GLOSSARY OF MEDICAL TERMS

Acetabulum	A concave surface of the pelvis. The head of the femur meets the pelvis at the acetabulum, forming the hip joint.
Albumin	The main protein of plasma. Albumin levels are often tested to evaluate nutritional status.
Anaemia	A condition in which there is an abnormally low number of red blood cells in the bloodstream.
Anorexia	A serious psychological eating disorder characterised by noticeably reduced appetite or total aversion to food.
Anthropometric	Of or relating to anthropometry; the study of human body measurement for use in classification and comparison.
Appendectomy	Surgical removal of the appendix.
Avascular necrosis	The loss of bone tissue due to a restriction of blood supply, leading to persistent hip pain; also known as osteonecrosis.
Arthritis	A relatively common group of conditions that causes damage to joints and bones, characterised by symptoms including restricted movements of the joints, pain and stiffness.
Arthroplasty (for fractured NoF)	A surgical procedure in which the hip joint is replaced by a prosthetic implant. It is the construction of a new moveable hip joint; both the acetabulum and the femoral head are replaced in this procedure.
Basocervical	Intra- or extracapsular.
Bone cyst	A fluid-filled cavity in the bone. It is benign (non-cancerous), but weakens the bone and makes it more likely to fracture. There is no known cause.
Cerebral dysfunction	Functional disorder of the brain.
Cerebrovascular accident (stroke)	The sudden death of some brain cells due to lack of oxygen when the blood flow to the brain is impaired by blockage or rupture of an artery to the brain.
Cholecystectomy	Surgical removal of the gall bladder.
Chronic obstructive pulmonary disease	A lung disease which leads to damaged airways in the lungs, causing them to become narrower, with symptoms including chronic cough, wheezing and tightness of the chest.
Clostridium difficile	A bacterium which lives in the gut. Symptoms of *Clostridium difficile* infection range from mild diarrhoea to a life-threatening bowel inflammation.
Combined peripheral nerve block technique	Injection of local anaesthetic near the nerve(s) that control sensation and movement of a specific area. Typically used for surgeries of upper or lower extremities.
Condylar	Rounded prominence at the end of a bone.

Confusion	Change in mental status in which a person is not able to think with his or her usual level of clarity, characterised by disorientation regarding time, place, person, or situation.
Crohn's disease	A chronic gastrointestinal disorder that causes inflammation of the lining of the digestive tract.
Debridement	The act of debriding: removing dead, contaminated or adherent tissue or foreign material.
Dementia	Significant loss of intellectual abilities such as memory capacity, severe enough to interfere with social or occupational functioning.
Displaced	Removed from the usual or proper place.
Diuretics	Drugs which increase the amount of water passed out from the kidneys, and consequently an increase in urine excretion.
Extracapsular	Situated or occurring outside a capsule (of a joint).
Femoral head	The rounded extremity of the femur (thigh bone); part of the hip joint. It is supported by the neck of femur.
Haemoglobin	The oxygen-carrying pigment and predominant protein in red blood cells.
Hemiarthroplasty (for fractured NoF)	A surgical procedure which replaces one half of the joint with an artificial surface and leaves the other part in its natural (pre-operative) state. The head of the femur is removed and replaced with a metal or composite prosthesis.
Human immuno-deficiency virus (HIV)	One of a group of viruses known as retroviruses, which kills or damages cells of the body's immune system; affects approximately 40 million people worldwide.
Intertrochanteric	Between the two trochanters of the femur.
Intracapsular	Situated or occurring within a capsule (of a joint).
Laparoscopic (surgery)	Minimally invasive ("keyhole") surgery which allows the surgeon to access the inside of the abdomen and pelvis.
Lymphocyte	Type of white blood cell in the immune system; a total lymphocyte count is often used to assess nutritional status.
Malignant	Tending to metastasise; cancerous.
Metastasis	The spread of a disease-producing agent (e.g. cancerous cells) from the initial primary site of the disease to another part of the body; the process by which such spreading occurs.
Methicillin-resistant *Staphylococcus aureus* (MRSA)	Any of several bacterial strains of the genus *Staphylococcus aureus* that are resistant to a broad range of conventional antibiotics, such as penicillin and methicillin; the most prevalent type of hospital-acquired infection in the United Kingdom.
Myocardial infarction	The death of heart muscle from the sudden blockage of a coronary artery by a blood clot. More commonly known as a heart attack.

Neck of femur (NoF)	The column of bone connecting the femoral head and the shaft of the femur.
Neuraxial block technique	Injection of local anaesthetic into the epidural or subarachnoid spaces.
Olecranon	The bone in the forearm that forms the pointed portion of the elbow; the bony projection of the ulna (a bone in the forearm) behind the elbow joint.
Paget's disease	A chronic disorder that typically results in enlarged and deformed bones. The excessive breakdown and formation of bone tissue that occurs with this disease can cause bone to weaken, resulting in bone pain, arthritis, deformities and fractures.
Pertrochanteric fracture	A fracture through the intertrochanteric region of the femur; a form of extracapsular hip fracture.
Pneumonia	Inflammation of the lung tissue, usually caused by an infection.
Pressure sores	Type of injury that affects areas of the skin and underlying tissue, caused when the affected area of skin is placed under too much pressure. Also known as pressure ulcers or bedsores.
Primary tumour	A tumour that is in the original site where it first arose.
Prosthesis	An artificial extension that replaces a missing body part.
Proximal	Relating to where an appendage joins the body.
Pulmonary	Relating to the lungs.
Secondary tumour	A tumour that develops as a result of metastasis.
Sexual dimorphism	The existence of physical differences between the sexes other than differences in the sex organs; the difference in form between male and female members of the same species.
Statins	Class of drugs used to lower cholesterol levels; a type of treatment for heart conditions.
Subtrochanteric	Situated or occurring below the trochanter.
Thrombosis; thromboembolism	Formation of a blood clot within a blood vessel; obstruction of a blood vessel caused by thrombosis.
Trochanter; trochanteric	A rough prominence at the upper part of the femur serving for the attachment of muscles; relating to the trochanter.
Undisplaced	Not removed from the usual or proper place.
Warfarin	Anticoagulant drug; used to prevent and treat the formation of harmful blood clots within the body by thinning the blood and/or dissolving clots.

APPENDIX D: SUPPLEMENTARY MATERIAL
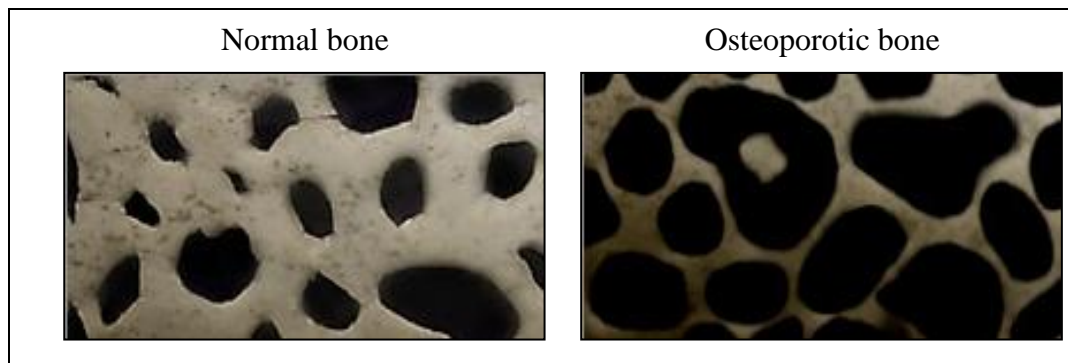
Figure D1.1.3a: Comparison of bone density between normal and osteoporotic bones
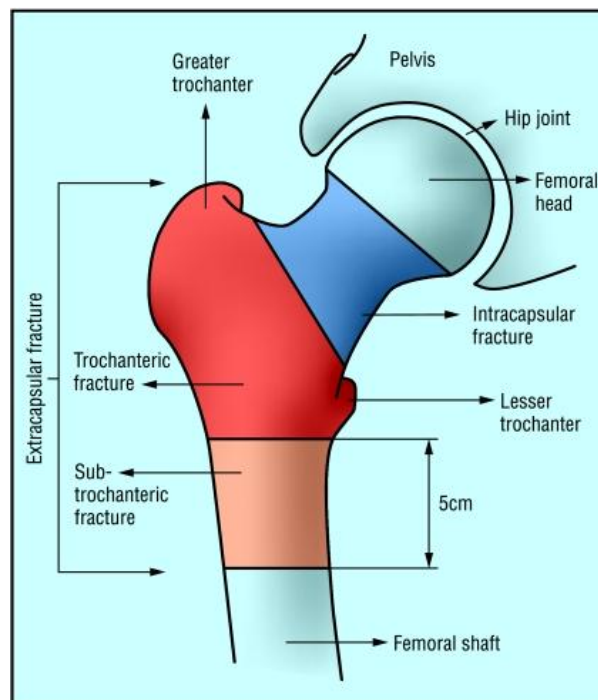
Figure D1.4.1a: Classification of hip fractures
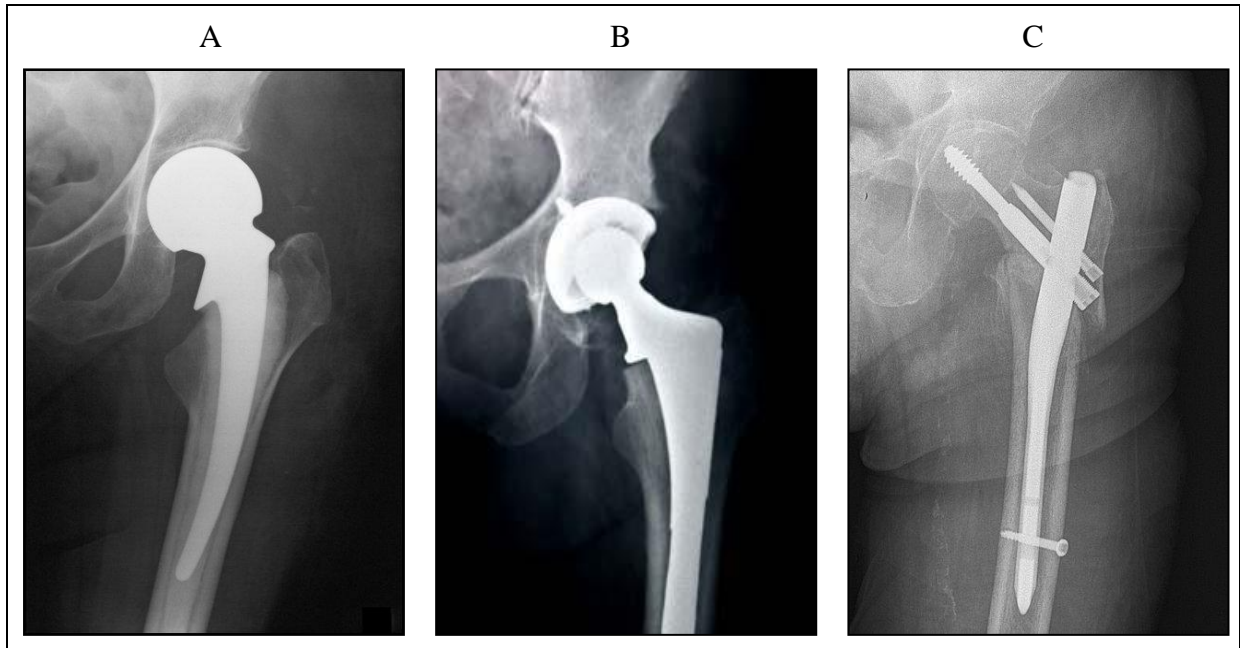
Figure D1.4.1b: Radiographs of (A) a prosthesis following hemiarthroplasty surgery, (B) a prosthesis following total hip replacement surgery and (C) a prosthesis following intramedullary nail surgery

Table D3.2.3a: Collinearity diagnostics for the multivariate linear regression model (VIF – variance inflation factor)

Variable	Tolerance	VIF
Intercept	-	-
Admfrom_d1	0.6809	1.4686
Admfrom_d4	0.8664	1.1543
Admfrom_d5	0.9659	1.0354
Mobility	0.6053	1.6521
Mentalst	0.7691	1.3003
WAASP	0.8003	1.2495
Opdelay	0.9651	1.0362
Age	0.7878	1.2693
SexM	0.9460	1.0571
Optypenew_d3	0.5078	1.9694
Optypenew_d6	0.9182	1.0891
Fractype_d1	0.5408	1.8492
Fractype_d5	0.9309	1.0743
Fractype_d6	0.9556	1.0464

Table D3.2.4a: Collinearity diagnostics for the multivariate linear regression model, surviving patients only

Variable	Tolerance	VIF
Intercept	-	-
Admfrom_d4	0.9071	1.1024
Admfrom_d5	0.9721	1.0288
Admfrom_d7	0.9509	1.0516
Mobility	0.6535	1.5302
Mentalst	0.7829	1.2773
WAASP	0.7775	1.2862
Opdelay	0.9511	1.0515
Age	0.7902	1.2655
SexM	0.8570	1.1668
ASAnew_n	0.9294	1.0760
Optypenew_d2	0.7887	1.2679
Optypenew_d3	0.4913	2.0355
Optypenew_d6	0.8980	1.1136
Fractype_d1	0.5419	1.8454
Fractype_d5	0.8665	1.1540

Table D3.3.2a: Splitting criteria for CART analysis on length of stay

Node number	Splitting variable	Splitting condition	Node number	Splitting variable	Splitting condition
0	-	-	32F	Optypenew_n	2, 4, 7
1	Mentalst	1	33F	Admfrom	4, 5, 8
2	Mentalst	2, 3	34	Admfrom	1, 2, 3, 6, 7
3	Walking0	1	35F	Age	≤ 76.5
4	Walking0	2, 3, 4, 5	36	Age	> 76.5
5	Opdelay	0	37	Age	≤ 87.5
6	Opdelay	1	38	Age	> 87.5
7F	Fractype	1, 3	39	Admfrom	1, 2, 6
8	Fractype	2, 4, 5, 6	40	Admfrom	3, 7
9	WAASP	1	41F	Fractype	1, 3, 4, 5
10	WAASP	2, 3	42F	Fractype	2, 6
11	ASAnew_n	1	43F	Optypenew_n	5, 7
12F	ASAnew_n	2, 3	44F	Optypenew_n	2, 3, 4, 6
13F	Age	≤ 75	45F	Age	≤ 91.5
14F	Age	> 75	46F	Age	> 91.5
15F	Side	1			
16F	Side	0			
17F	Optypenew_n	3, 4, 6, 7			
18	Optypenew_n	2, 5			
19F	WAASP	2			
20F	WAASP	1, 3			
21F	Optypenew_n	3, 6			
22	Optypenew_n	2, 4, 5, 7			
23	Opdelay	0			
24	Opdelay	1			
25F	Fractype	1, 2			
26	Fractype	3, 4, 5, 6			
27F	Walkaid0	2, 3, 5			
28F	Walkaid0	1, 4			
29	Walkaid0	1, 2, 3, 5			
30F	Walkaid0	4			
31F	Optypenew_n	5			

Table D4.3.2a: Splitting criteria for CART analysis on mortality using the Gini Index procedure method

Node number	Splitting variable	Splitting condition	Node number	Splitting variable	Splitting condition
0	-	-	30	Walkaid0	1, 3, 4
1	Walking0	2, 3, 4	31	Admfrom	3
2	Walking0	1, 5	32F	Admfrom	4, 5, 8
3	ASAnew_n	1, 2	33F	Optypenew_n	2, 3
4	ASAnew_n	3	34F	Optypenew_n	4, 5, 6, 7
5	Mentalst	1	35F	Fractype	1, 2
6	Mentalst	2, 3	36F	Fractype	3, 4, 5, 6
7	ASAnew_n	1	37	Admfrom	1, 2, 4, 5, 6
8	ASAnew_n	2	38F	Admfrom	3, 7, 8
9	Walking0	3	39	Optypenew_n	4, 5, 6
10F	Walking0	2, 4	40	Optypenew_n	2, 3, 7
11F	Optypenew_n	2, 3	41	Fractype	1, 2, 4, 5
12F	Optypenew_n	4, 5, 6, 7	42F	Fractype	3, 6
13F	Fractype	3, 4	43	SexM	0
14	Fractype	1, 2, 5, 6	44	SexM	1
15	Walking0	3	45F	Walkaid0	2, 4
16	Walking0	2, 4	46	Walkaid0	1, 3, 5
17F	Mobility	1, 2	47F	ASAnew_n	1
18F	Mobility	3	48F	ASAnew_n	2, 3
19F	WAASP	2	49F	Opdelay	0
20F	WAASP	1, 3	50F	Opdelay	1
21	Admfrom	1, 2, 6, 7	51	WAASP	1, 2
22	Admfrom	3, 4, 5, 8	52F	WAASP	3
23	WAASP	1, 2	53F	Fractype	5
24	WAASP	3	54F	Fractype	1, 2, 3, 4, 6
25F	Opdelay	0			
26F	Opdelay	1			
27F	Walkaid0	1, 2, 3			
28F	Walkaid0	4			
29F	Walkaid0	2			

Table D4.3.2b: Splitting criteria for CART analysis on mortality using the Information Entropy procedure method

Node number	Splitting variable	Splitting condition	Node number	Splitting variable	Splitting condition
0	-	-	31	Walkaid0	1
1	Age	≤ 83.5	32F	Walkaid0	2, 3, 4, 5
2	Age	> 83.5	33F	Fractype	1, 3, 4
3	Age	≤ 65.5	34F	Fractype	2, 5
4	Age	> 65.5	35	Walking0	2, 3, 4
5F	WAASP	1, 2	36	Walking0	1, 5
6F	WAASP	3	37	WAASP	1, 2
7	Walking0	2, 3, 4	38	WAASP	3
8	Walking0	1, 5	39	SexM	0
9	ASAnew_n	1, 2	40F	SexM	1
10F	ASAnew_n	3	41	Opdelay	0
11	Mobility	1, 2	42	Opdelay	1
12	Mobility	3	43	Mentalst	1
13	Fractype	2, 4	44F	Mentalst	2, 3
14F	Fractype	1, 3, 5, 6	45F	ASAnew_n	2
15F	WAASP	2	46F	ASAnew_n	1, 3
16F	WAASP	1, 3	47F	Fractype	1, 3, 4
17F	ASAnew_n	1	48F	Fractype	2, 5, 6
18	ASAnew_n	2	49F	Admfrom	2, 4, 5, 6
19F	Admfrom	2, 3, 4, 5	50	Admfrom	1, 3, 7, 8
20F	Admfrom	1, 6, 7, 8	51F	Fractype	4, 5
21	Optypenew_n	4, 5, 6	52	Fractype	1, 2, 3, 6
22	Optypenew_n	2, 3, 7	53F	ASAnew_n	2
23	Opdelay	0	54F	ASAnew_n	1, 3
24	Opdelay	1	55	Optypenew_n	4, 5
25F	ASAnew_n	1	56	Optypenew_n	2, 3, 6, 7
26F	ASAnew_n	2, 3	57F	Walkaid0	2
27F	ASAnew_n	1	58F	Walkaid0	1, 3, 4, 5
28F	ASAnew_n	2, 3	59F	Fractype	1, 4
29F	WAASP	2	60F	Fractype	2, 3, 5, 6
30	WAASP	1, 3			

Figure D6.3.4a: Distribution of operative delay in hours for ASAIII patients who are delayed for more than 48 hours against the Negative Exponential distribution
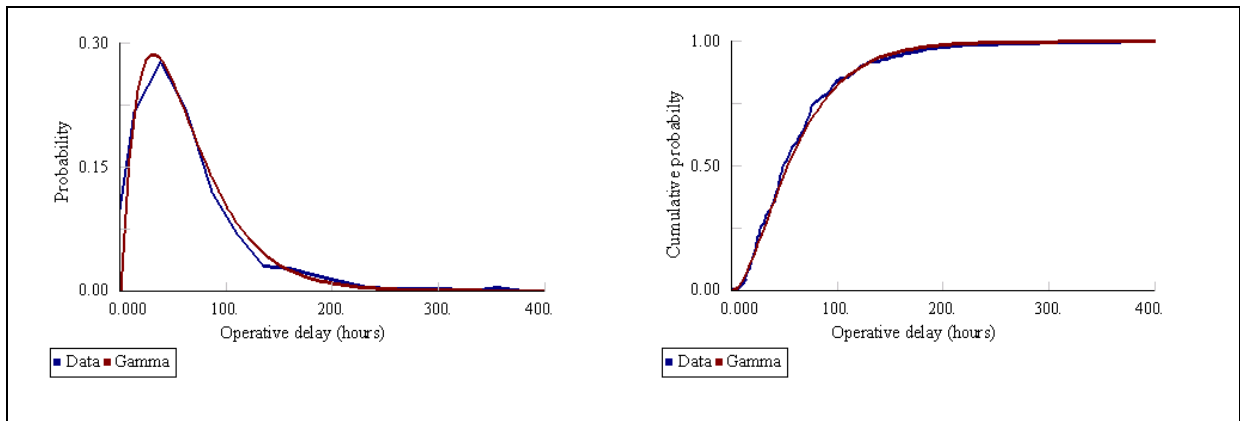
Figure D6.3.4b: Distribution of operative delay in hours for all ASA3 patients against the Gamma distribution with parameters min = 1, α = 1.971 and β = 31.745
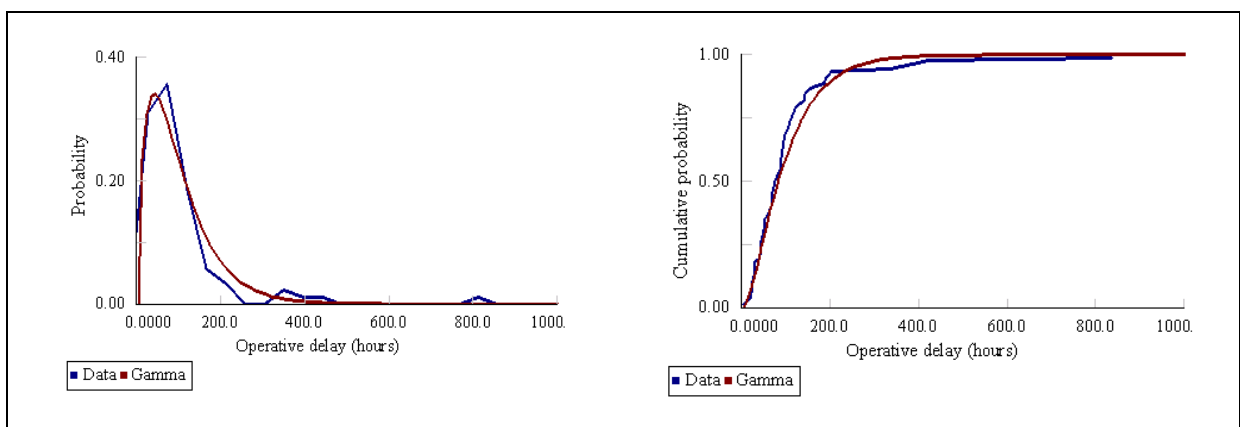
Figure D6.3.4c: Distribution of operative delay in hours for all ASAIV patients against the Gamma distribution with parameters min = 1, α = 1.629 and β = 59.978

1. Sample from the appropriate Gamma distribution to gain a value for operative delay, d.

2. If $d \leq 48$, patient is classified as not delayed; take d as operative delay value, go to END. If $d > 48$, go to 3.

3. Decide type of delay and sample from the appropriate Negative Exponential distribution to get total delay, t (d is discarded).

4. Calculate administrative delay, a. To avoid the potential problem of the sampled administrative delay exceeding the total delay, the random number used to determine administrative delay is scaled to be in the range $[0, r]$, where $r = F_x(t) = P(X \leq t)$ and F_x represents the cumulative distribution function of the time spent administratively delayed.

5. Calculate clinical delay c, where $c = t - a$. Go to END.

END

Figure D6.3.4d: Method of calculating delay for ASAIII and ASAIV patients
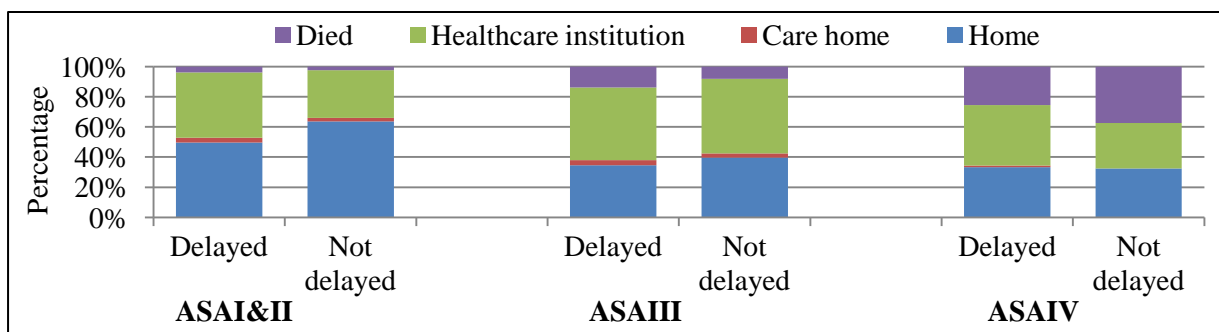
Figure D6.3.7a: Acute discharge destinations for patients admitted from home
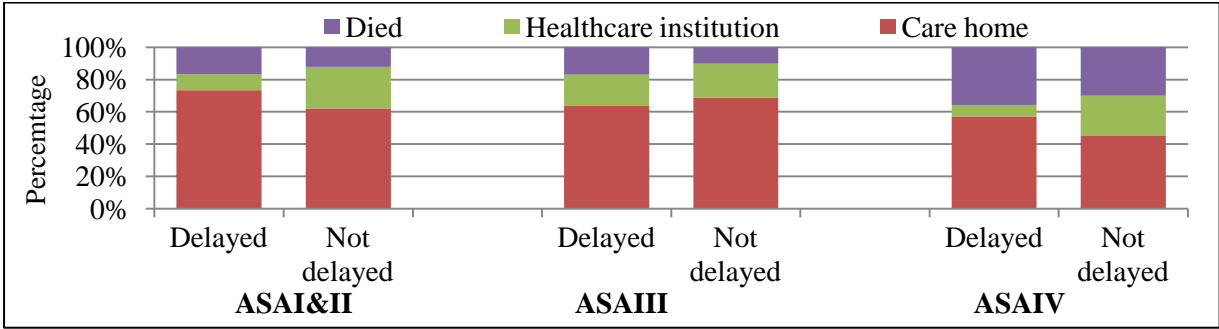
Figure D6.3.7b: Acute discharge destinations for patients admitted from a care home
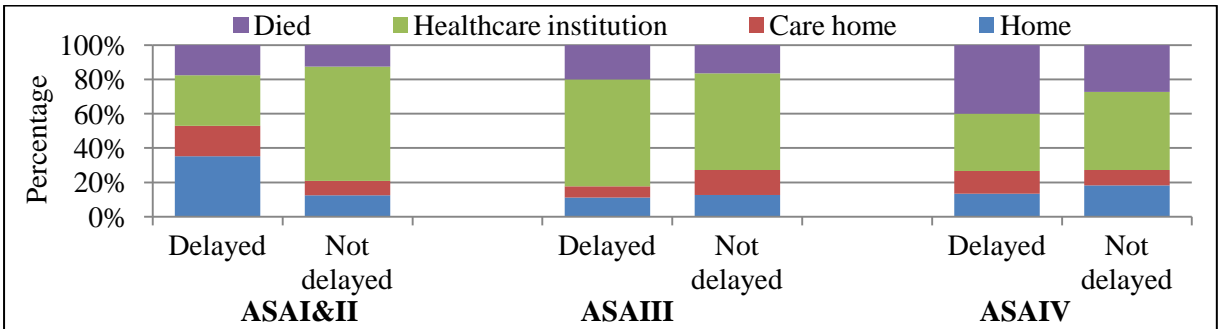
Figure D6.3.7c: Acute discharge destinations for patients admitted from a healthcare institution
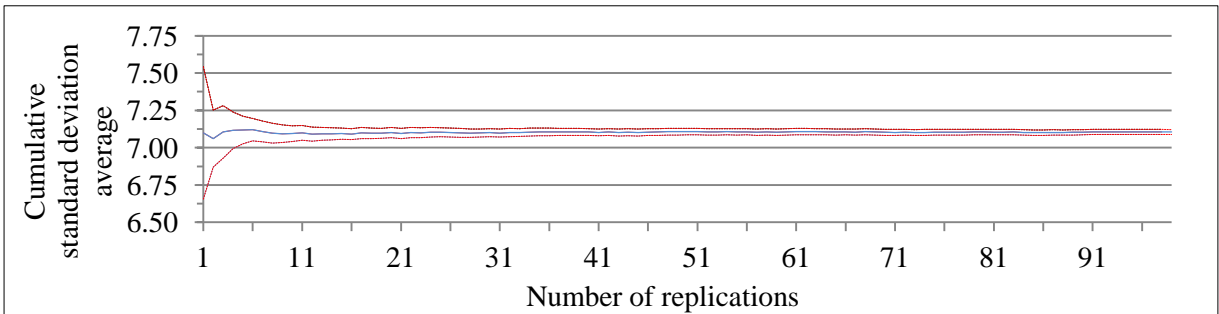
Figure D6.3.11a: Results of standard deviation of bed occupancy
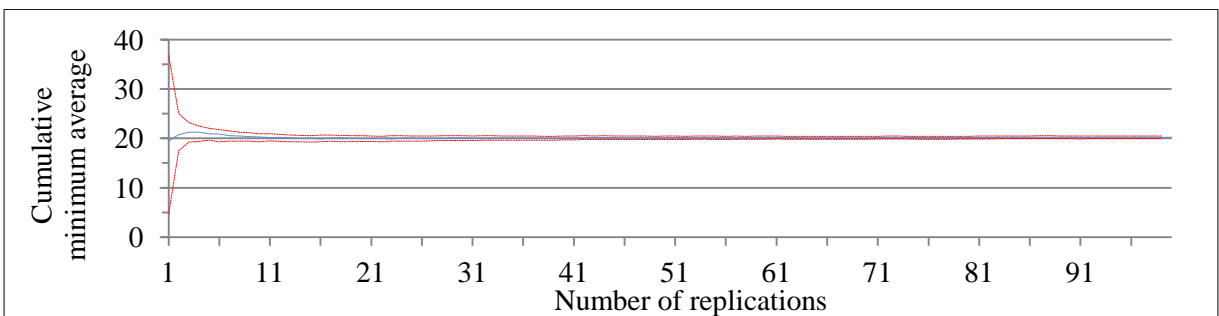
Figure D6.3.11b: Results of minimum bed occupancy
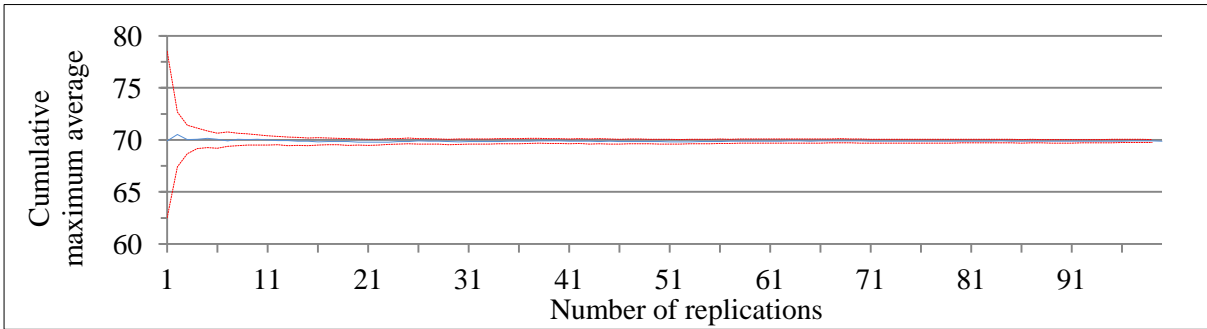
Figure D6.3.11c: Results of maximum bed occupancy
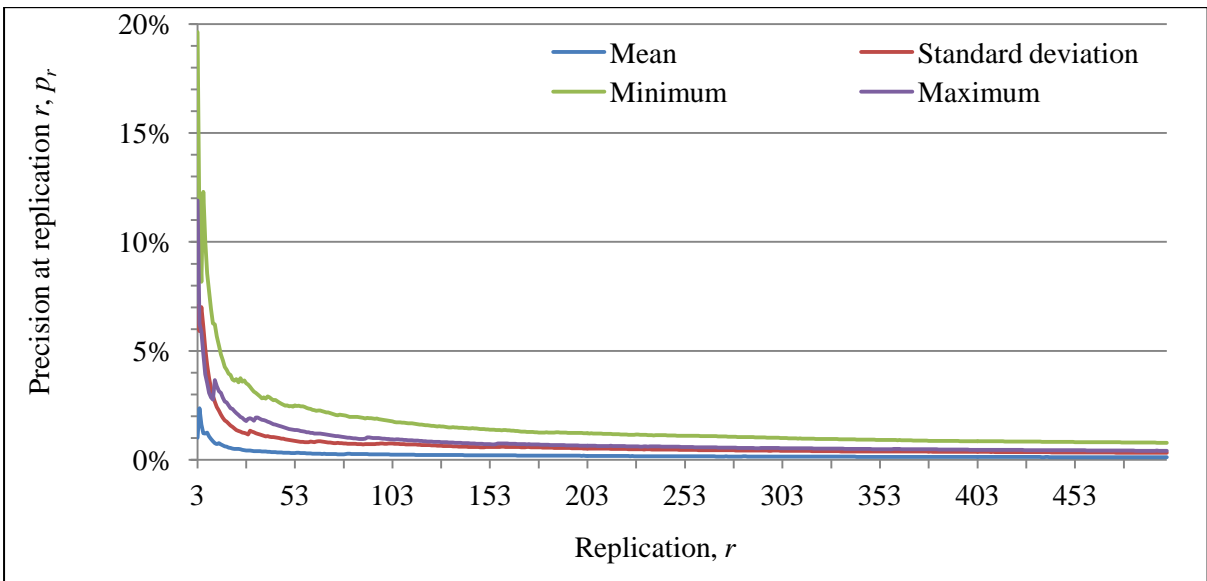
Figure D6.3.11d: Precision values obtained for various bed occupancy measures
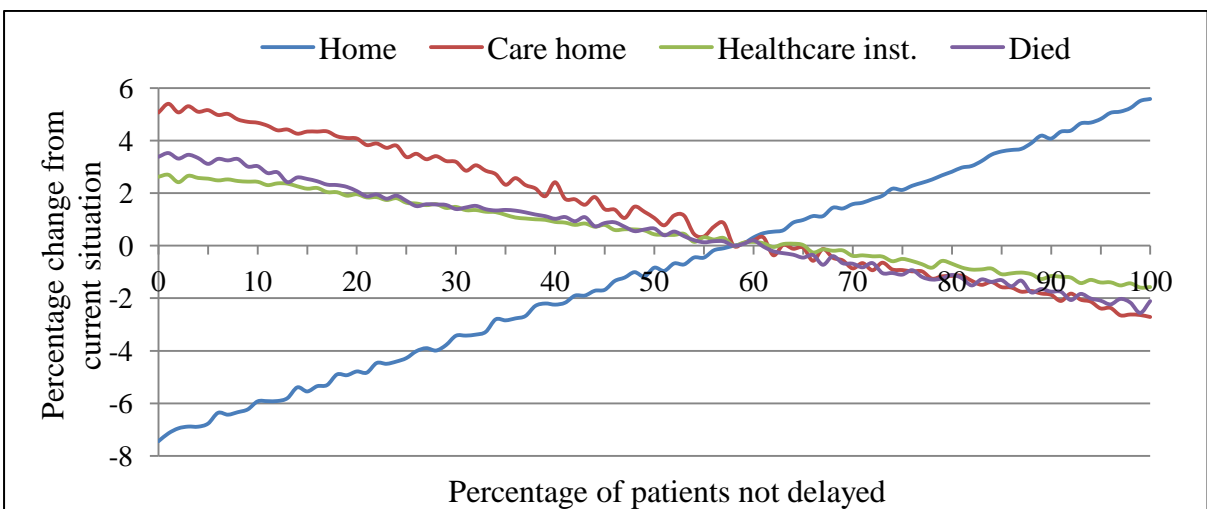
Figure D6.5.1a: Results of changing the percentage of ASA grade I&II patients who are not delayed on acute discharge destination
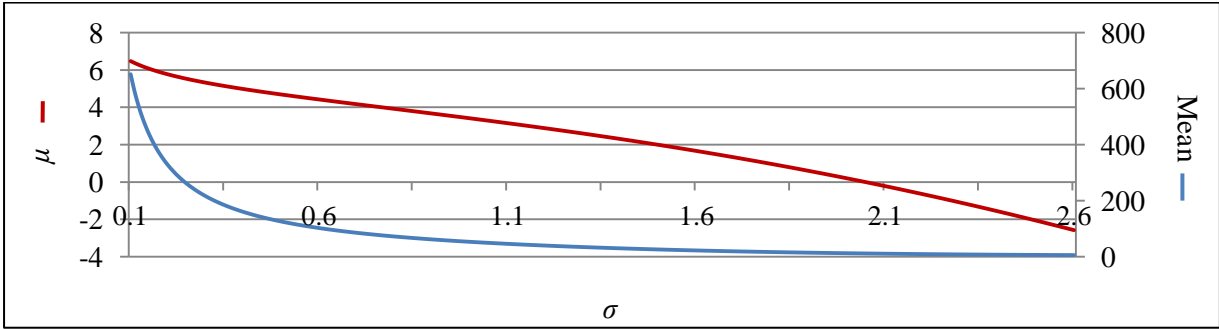
Figure D6.6.2a: The relationship of μ, σ and the mean for the Lognormal distribution with a fixed standard deviation of 65.1 and minimum of 3
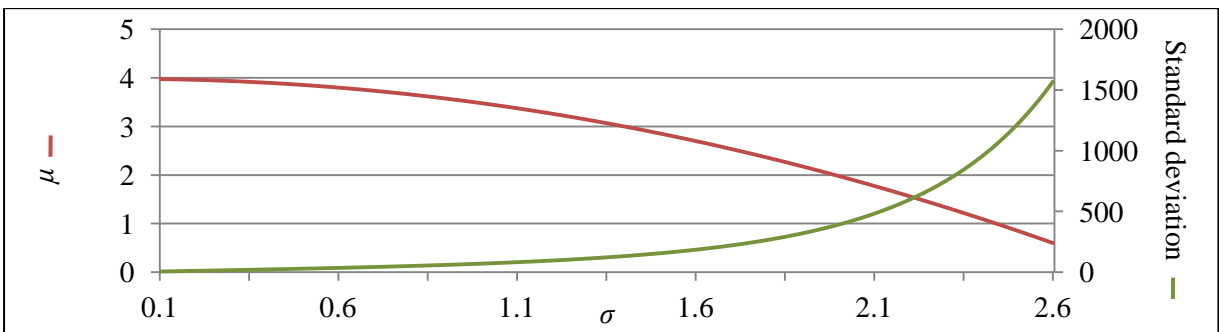
Figure D6.6.2b: The relationship of μ, σ and the standard deviation for the Lognormal distribution with a fixed mean of 56.4 and minimum of 3
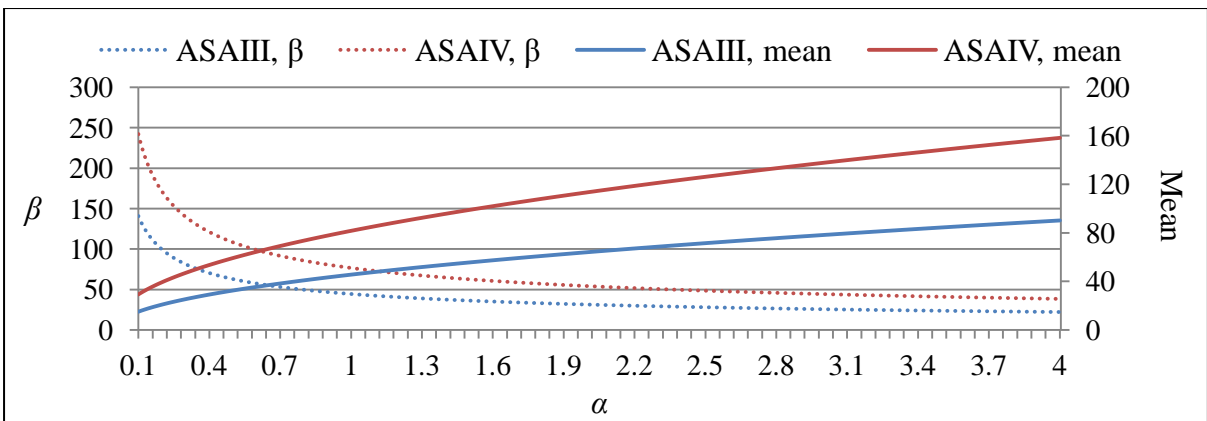
Figure D6.6.3a: The relationship of α, β and the mean for the Gamma distribution with a fixed standard deviation
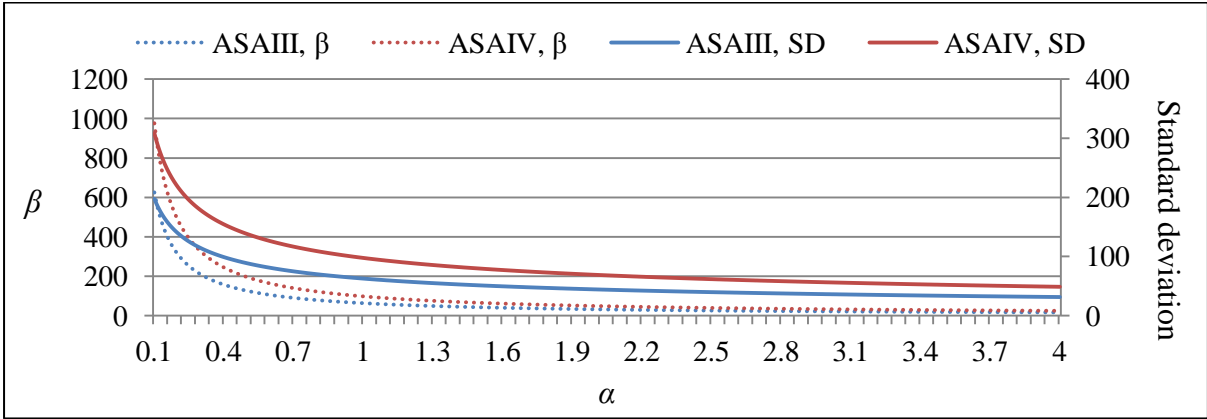
Figure D6.6.3b: The relationship of α, β and the standard deviation for the Gamma distribution with a fixed mean
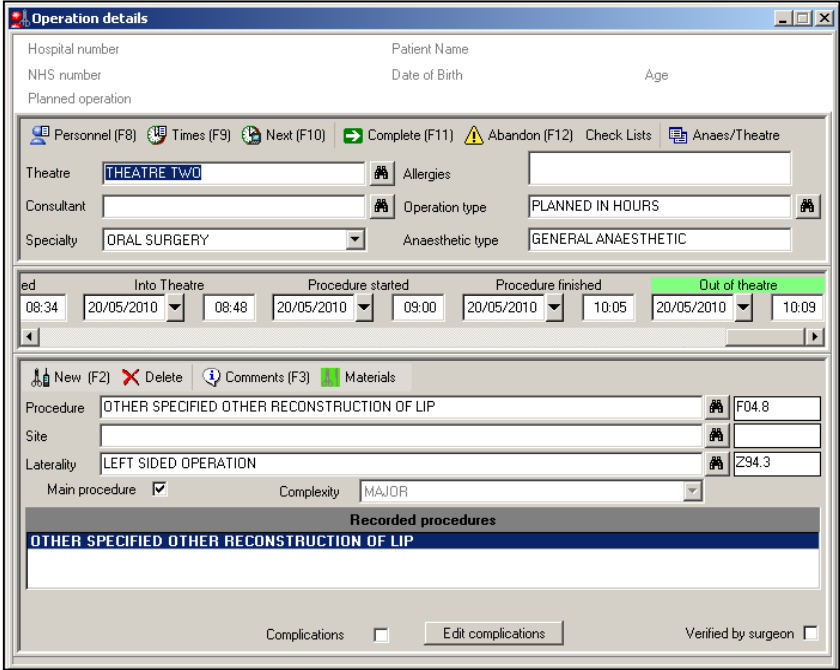
Figure D7.1.1a: TheatreMan screenshot
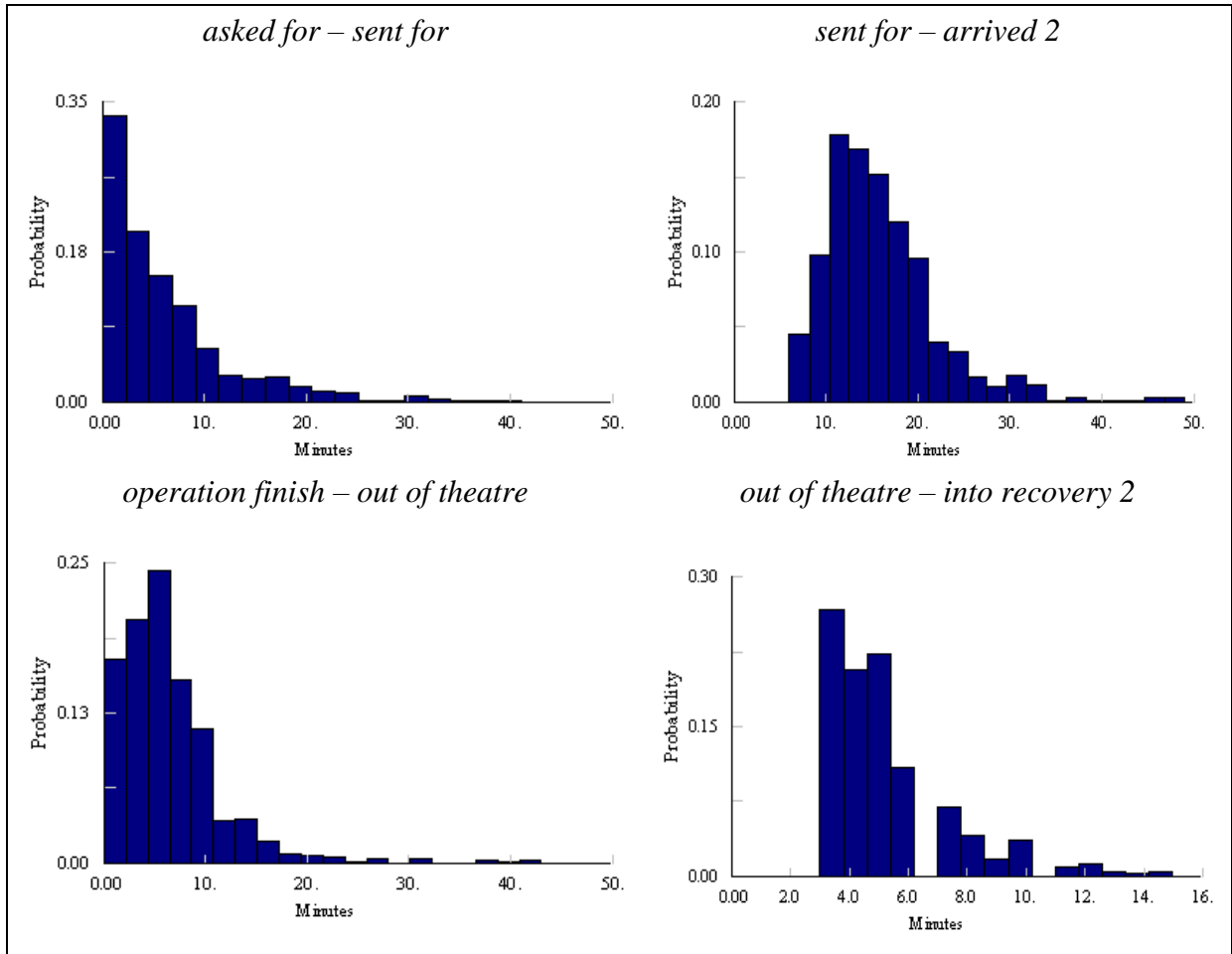
Figure D7.2.1a: Distribution of time spent in various theatre intervals for trauma hip surgery
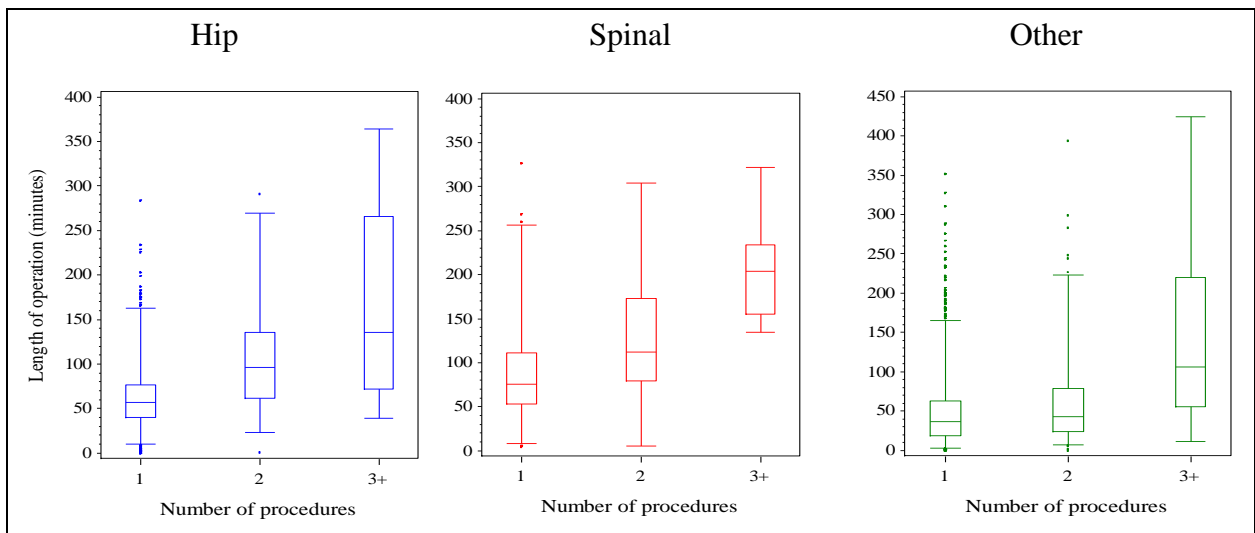
Figure D7.4a: Distribution of operation time by operation type
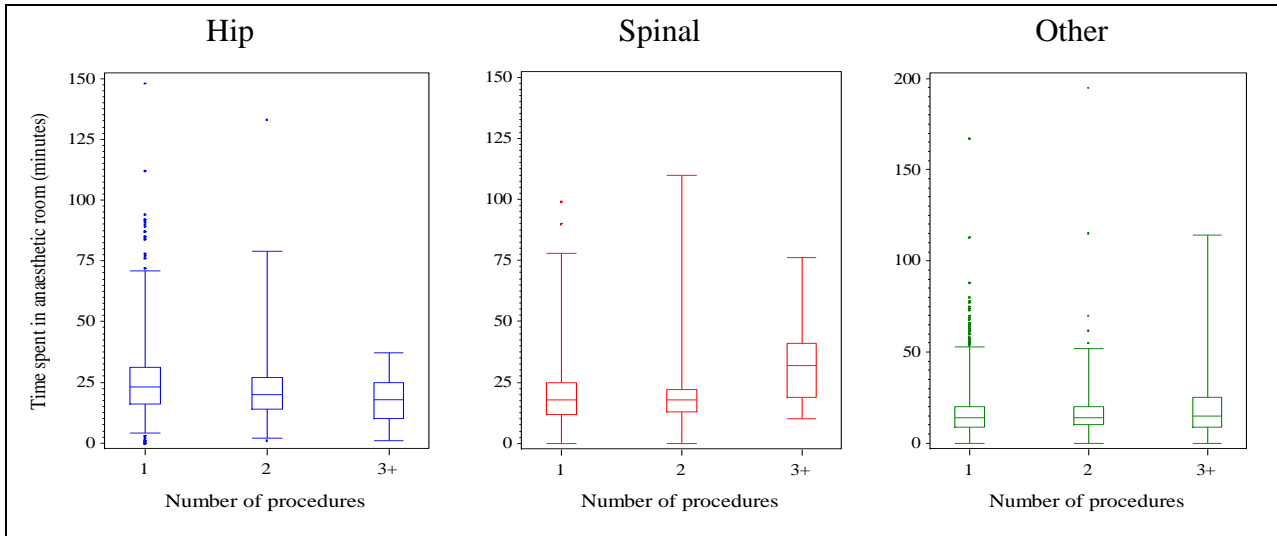
Figure D7.4b: Distribution of anaesthetic time by operation type
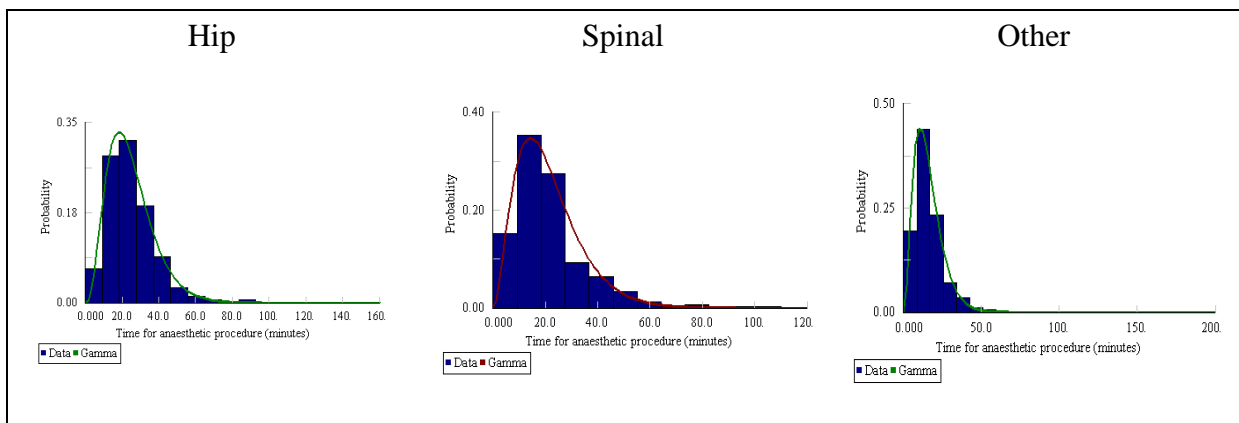
Figure D8.2.2a: Gamma fits for time for anaesthetic procedure, by operation type
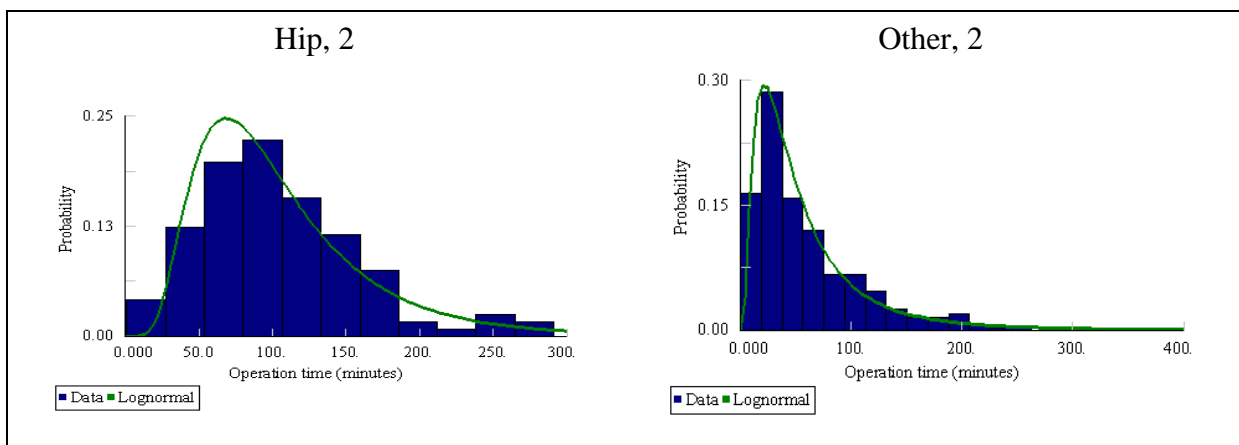
Figure D8.2.3a: Lognormal fits for operation time, by operation type and number of procedures

427
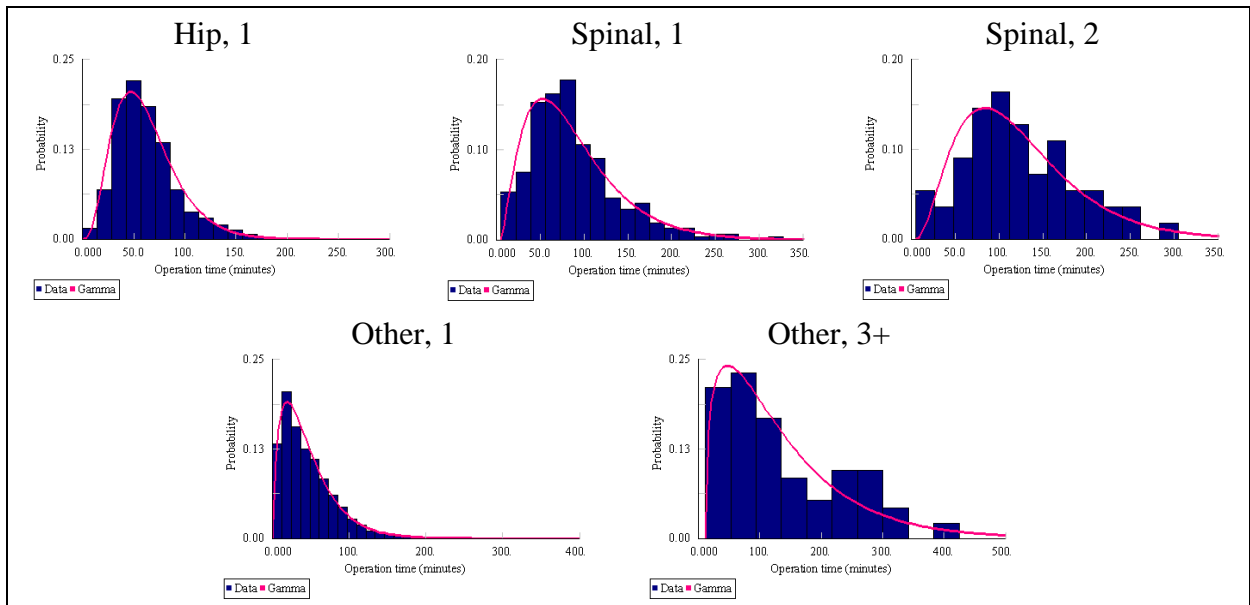
Figure D8.2.3b: Gamma fits for operation time, by operation type and number of procedures
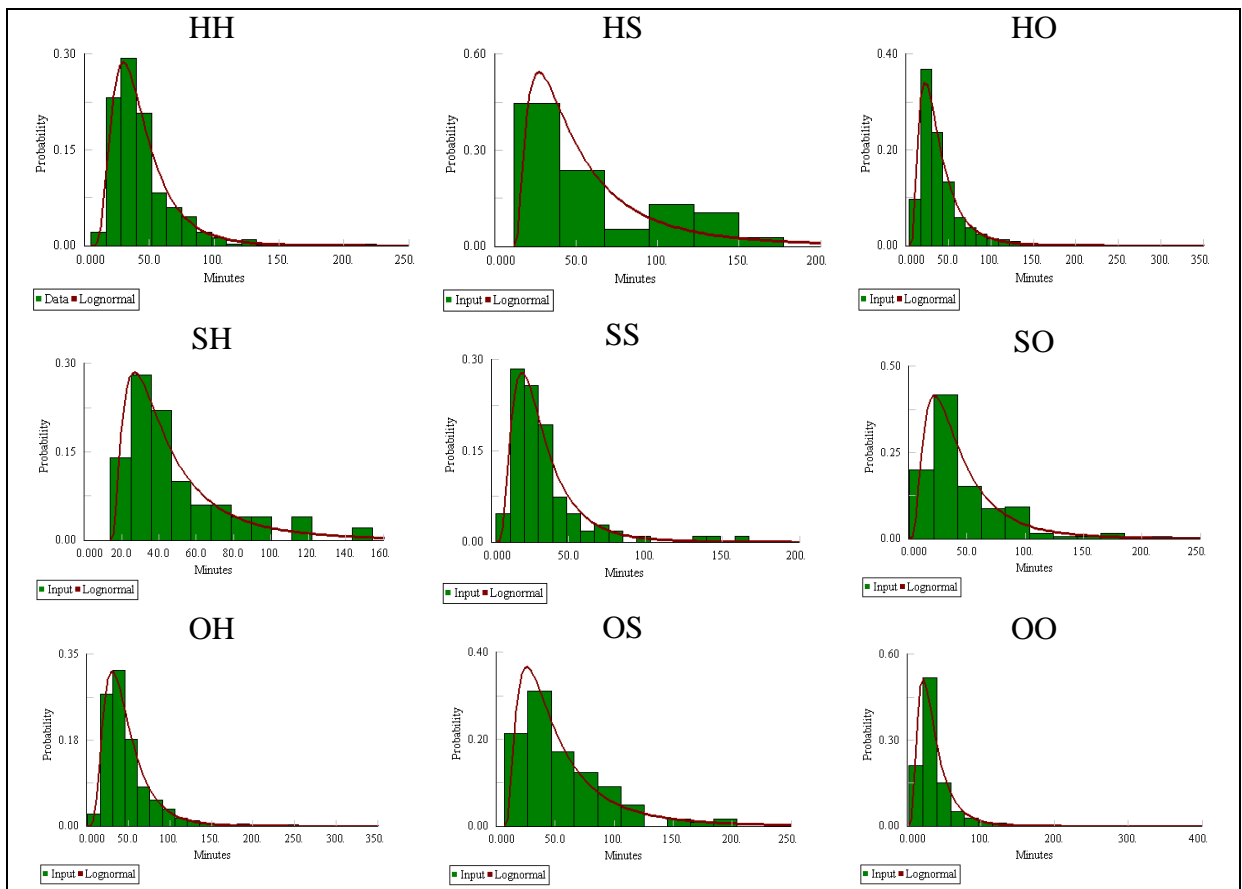
Figure 8.2.4a: Lognormal fits for distribution of turnover time, by sequence of operations

428
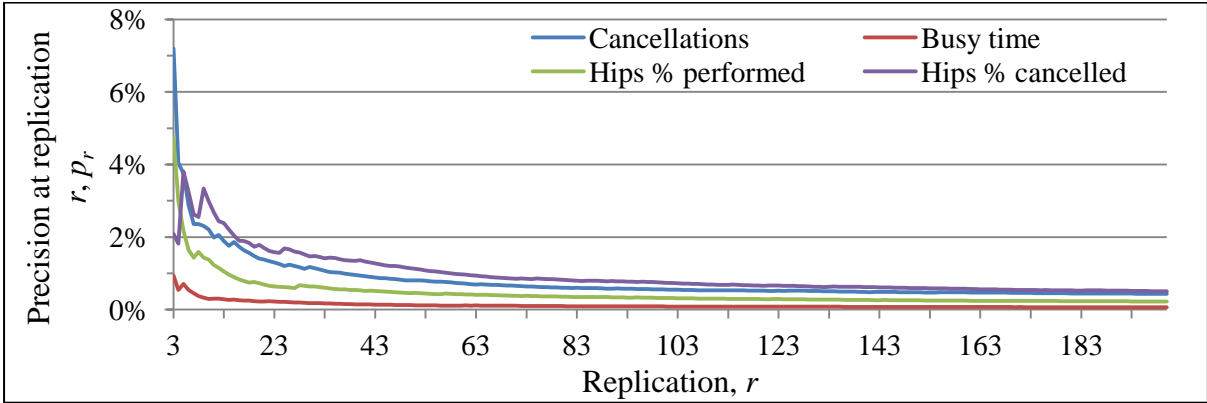
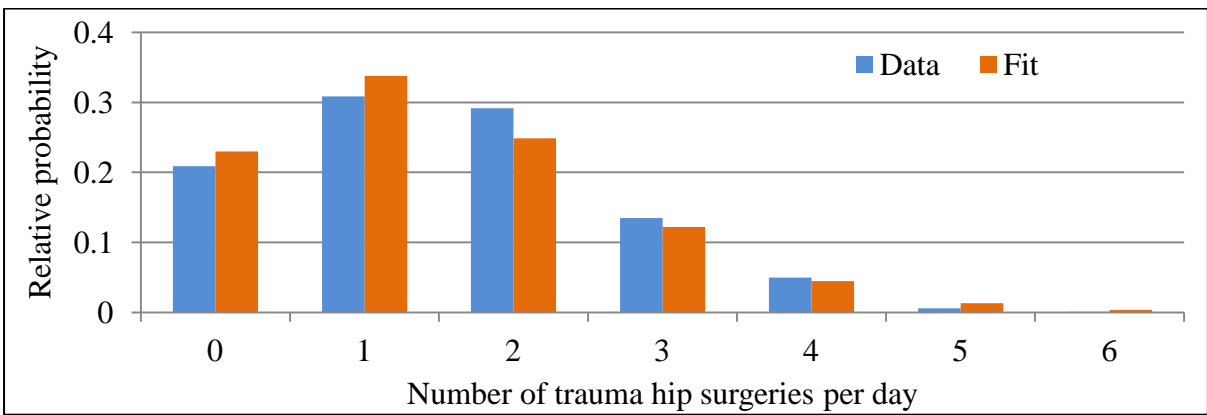**Figure D8.2.7a:** Precision values obtained for various trauma theatre model measures

Figure D10.1a: Number of trauma hip surgeries per day against the Poisson distribution with parameter $\lambda = 1.47$
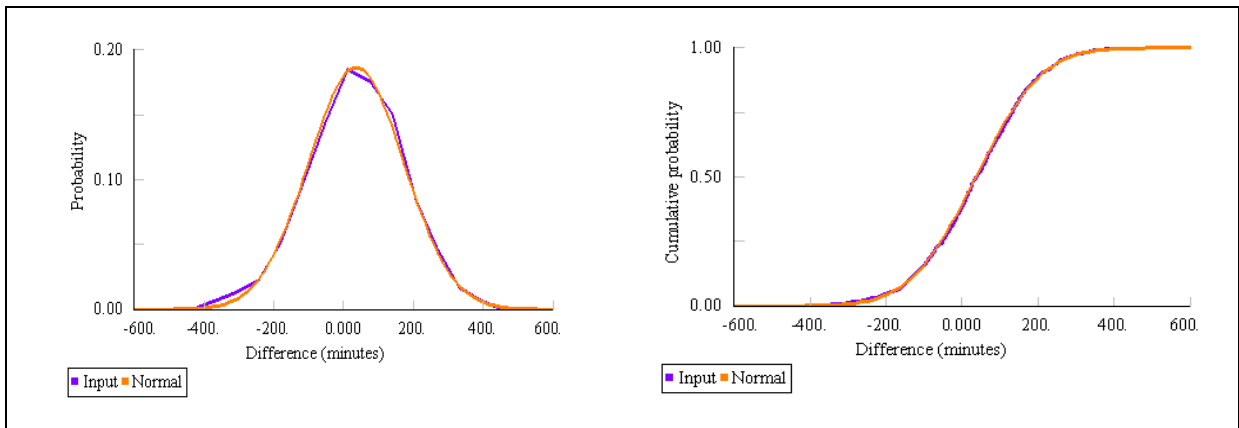

Figure D11.5.4a: PDF and CDF of the differences between actual theatre usage and theatre time predicted by the mean of the Lognormal distribution given by the Fenton-Wilkinson approximation, against the Normal distribution with parameters $\mu = 37.88$ and $\sigma = 136.9$
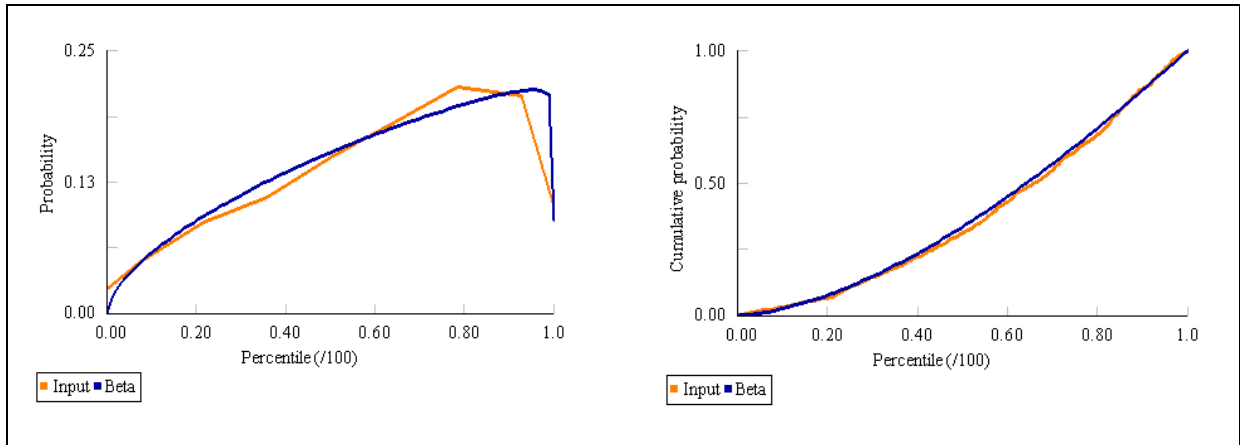
Figure D11.5.4b: PDF and CDF of percentiles which correctly predict theatre time by the Fenton-Wilkinson approximation, against the Beta distribution with parameters min = 0, max = 1, p = 1.618 and q = 1.029