Morbidity following Orthopaedic Surgery

Dr Elizabeth ASHBY

BA MB Chir MA (Cantab) FRCS (Tr & Orth)

UCL

Doctor of Medicine, 2014

I, Elizabeth ASHBY, confirm that the work contained in this thesis is my own. Information from other sources is indicated below:

Chapter 1: Nil

Chapter 2 and Chapter 3:
Data collection by Sr Claire Matejowsky and Sr Maj Mutch

Chapter 4:
Data collection by Mr Rahul Patel

Chapter 5 and Chapter 6:
Data collection by UCLH wound surveillance team lead by Dr APR Wilson

Chapter 7: Nil

Table of contents

Table of tables

Table of figures

Table of images

Abstract

Morbidity following hip and knee arthroplasty has previously been poorly recorded. This is the first time the Post-Operative Morbidity Survey (POMS) has been used for this purpose. The POMS identifies clinically significant morbidity using indicators of organ system dysfunction rather than traditional diagnostic categories.

The most common types of morbidity following hip and knee arthroplasty are infection and renal morbidity. Pulmonary, pain and gastro-intestinal morbidity are less common. Cardiovascular, wound, neurological and haematological morbidity are least common.

Many arthroplasty patients remain in hospital without morbidity. The POMS identifies these patients and thus has potential as a prospective bed utilisation tool. To be used for this purpose, the POMS must identify all clinically significant morbidity. Mobility is an important factor for safe discharge of arthroplasty patients. Addition of a 'mobility' domain could improve the utility of POMS as a bed utilisation tool following orthopaedic surgery.

This study showed no association between post-operative morbidity defined by the POMS and longer-term patient-reported outcome measures (PROMs). This study does not support the POMS as an early surrogate marker of long-term PROMs in orthopaedic patients.

The wound domain of the POMS has a high specificity, reasonable sensitivity, high negative predictive value and low positive predictive value compared to the inpatient ASEPSIS (**A**dditional treatment, **S**erous discharge, **E**rythema, **P**urulent exudate, **S**eparation of deep tissues, **I**solation of bacteria, inpatient **S**tay over 14 days) score. The wound domain of POMS could be replaced with a validated definition of wound infection such as ASEPSIS.

On the same series of orthopaedic patients, surgical site infection (SSI) rate according to the Centres for Disease Control (CDC) definition was 15.45%, according to the Nosocomial Infection National Surveillance Scheme (NINSS) definition was 11.32% and according to the ASEPSIS definition was 8.79%. This highlights the need for a consistent definition of SSI.

Acknowledgements

Prof Mike Grocott for extreme patience and continuing support

Prof Fares Haddad for encouragement and opportunity

Mr Adrian Toutoungi, proofreader and my rock

Chapter 1: Background

1.1 Introduction

In this chapter I will describe how surgical outcome is measured for interventions on the hip and knee. I will start by explaining why surgical outcome is measured and I will describe the introduction of outcome measures in the United Kingdom. I will proceed to explain how these measures have evolved with time.

I will then discuss how the quality of surgery can be measured. I will describe how quality can be categorised and the advantages and disadvantages of reporting each category.

Following this I will discuss how outcome measures are adjusted to allow for patient factors. I will describe two risk classification systems that are commonly used in orthopaedic studies.

Next I will describe outcome measures used to assess orthopaedic surgery. I will focus on hip and knee arthroplasty. I will discuss mortality rates, hospital length of stay, post-operative morbidity and quality of life measures. In the morbidity discussion I will pay particular attention to surgical site infection. I will divide quality of life measures into generic outcomes, disease-specific outcomes and joint-specific outcomes. The joint-specific tools will be further

divided into surgeon-reported and patient-reported measures.

Following this I will discuss the possibility of using post-operative outcome measures as bed utilisation tools. Finally I will discuss the ways in which outcome measures are evaluated and validated.

1.2 Why measure the outcome of orthopaedic interventions?

The World Health Organisation (WHO) estimates that over 234.2 million surgical procedures are performed worldwide each year[1]. Due to high morbidity and mortality rates, the WHO recommended that surgical safety should be a global public health concern and surveillance of all types of surgery should be established.

The aim of surgery is either to increase quality of life or to prolong life. A small proportion of elective orthopaedic surgery aims to prolong life e.g. excision of primary bone tumours. However, the aim of most elective orthopaedic surgery is to improve quality of life. For example, lower limb arthroplasty (e.g. total hip replacement) aims to improve joint pain and mobility.

All surgery carries risk and can have a significant impact on a patient's life. There are physical (e.g. pain, fatigue), psychological (e.g. anxiety, depression) and social (e.g. loss of income due to work absence) implications of surgery.

Mortality is a risk of all surgery but is low in most elective orthopaedic procedures.

There are both moral and economic reasons to continually strive to improve surgical outcome. Morally, a surgical procedure should offer a patient benefit with the lowest possible risk. From an economic point of view, in order to maximise the cost-effectiveness of surgery, the maximum benefit must be achieved whilst minimising risk. The risk of surgery comes from the operation itself and from peri-operative interventions.

Measures of outcome are required to monitor the effectiveness of surgical interventions. These measures need to evaluate both the benefits and risks of surgery. They have several purposes. Firstly, they are used in clinical trials to compare different interventions. In orthopaedic surgery, they can be used to compare joint prostheses, different methods of prosthesis fixation and varying surgical techniques. Outcome measures are also used to assess peri-operative care including the use of prophylactic antibiotics, the effect of increased physiotherapy input and alternate regimes of post-operative analgesia. Secondly, outcome measures are used for audit purposes to compare individuals, departments, hospitals and regions. In this way outliers can be identified. This enables good practice to be highlighted and propagated and remedial action to be initiated where practice is sub-standard. Thirdly, regular feedback of outcome measures to surgeons makes them more aware of their own performance, and has been shown to improve results[2]. Fourthly,

outcome measure results can be used to guide the distribution of resources. In departments with poor results, increased training or new equipment may be needed. Finally, results from outcome measures can be used to guide patients and purchasers in their choice of surgeon and institution. Some studies show that patients are reluctant to use healthcare performance indicators[3] and that they are more concerned about other factors (such as the proximity of a hospital to their home) than outcome results[4]. However, a more recent study indicates that patients do wish to make an informed choice regarding their healthcare based on outcome data[5]. The 'NHS Choices' website now provides the general public with information regarding hip and knee arthroplasty procedures at different hospitals. Information such as average length of hospital stay, waiting time for surgery and surgical site infection rates is available to view.

1.3 The use of performance indicators in the UK

The National Health Service (NHS) was formed in the UK in 1948. This service provides free universal healthcare at the point of contact, paid for by general taxation. Until recently healthcare delivery was mainly self-regulated. However, due to budget restrictions and recent gross failures in the self-regulatory system (such as the Harold Shipman case), accountability has become an integral part of the NHS.

The push for accountability started in 2000 when the government announced an increase in NHS spending. In an attempt to ensure that extra spending resulted in a better service, hospitals were required to publish performance indicators.

Initially, most of the data published by the Healthcare Commission was process focused and a large emphasis was put on the prompt delivery of services. There was little initial focus on the outcome of services.

In 2003 the Department of Health introduced 'Payment by Results'[6]. Prior to this, hospitals were paid a fixed annual sum for providing a service. This sum did not reward efficiency or activity. 'Payment by Results' aimed to change this by paying a fixed 'tariff' for each procedure performed. This encouraged increased activity with increased payment but it did not incentivise higher 'quality'.

In 2008, the Department of Health published a report entitled 'High Quality Care For All'[7]. This report stipulated that all hospitals must collect and publish surgical outcome data by 2010. This data covers three domains: safety (e.g. surgical site infection rates), clinical effectiveness (e.g. patient-centred quality-of-life assessments) and personal experience (e.g. respectful treatment of patients). The payment for each procedure is now modified based on outcome data. Thus, the government is now rewarding 'quality' of healthcare, as well as 'quantity'.

1.4 How is the quality of orthopaedic surgery measured?

Performance indicators include information regarding the 'structure' of a service, the 'process' by which a service is delivered and the 'outcome' of a service.

Data regarding the ‘structure’ of a service provides information about the place where the service is delivered. Examples in orthopaedic surgery include the number of operating theatres, the ‘quality’ of surgical equipment and the ‘quality’ of the operating theatres e.g. whether a laminar airflow system is present. ‘Process’ data provides information about how efficiently a service is delivered. In orthopaedic surgery, examples include the waiting time for a surgical intervention and the level of physiotherapy provision following surgery. ‘Outcome’ data provides information about the risks and benefits of surgery. Outcome indicators used in orthopaedic surgery include surgical site infection rates and patient-centred outcome measures (both of which are now reported on the NHS Choices website).

The Department of Health introduced ‘High Quality of Care for All’[7] to improve the ‘quality’ of medical services but there was considerable debate over which of the quality measurements (structure, process or outcome) should be used. ‘Structural’ factors remain relatively stable over time. They allow for the provision of high quality care but do not directly result in high quality care. For these two reasons ‘structural’ factors are not deemed a fair way to reward quality.

‘Process’ measures are easy to record and have been used in the NHS as measures of ‘quality’ with financial implications. For example, hospitals used to receive a financial penalty for any patient in the Accident and Emergency Department who did not receive treatment within 4 hours. However ‘process’

measures have their limitations. There is no sense in rewarding a process if it does not result in a better outcome. A better outcome with poor processing is preferable to a poor outcome with efficient processing. Rewarding 'process' measures alone can lead to a decline in quality of care. This is seen when managers are rewarded for 'processing' increased numbers of operative interventions. Patients may be operated on by less experienced surgeons, there may be less time for safety checks, older equipment may be used since there is insufficient time to sterilise newer equipment, and non-orthopaedic theatres may be used. All these factors can lead to a poorer outcome. If 'process' factors are to be rewarded, a link to 'outcome' must be proven.

'Outcome' measures are generally accepted as a good way of measuring 'quality'. However the data is often not readily available and requires resources to collect it. 'Process' data may be used until 'outcome' data is available, but 'outcome' data should be the ultimate way to assess any intervention. 'Outcome' measures must be continually validated to ensure that they provide a true representation of the benefits or failings of a surgical intervention.

The Department of Health splits 'quality' into three main categories: safety, clinical effectiveness and personal experience. Safety is not clearly defined but can be surmised to mean 'do as least harm as possible' i.e. reduce risks to a minimum. In orthopaedic surgery, there are certain events or risks that should never happen. These are known as 'never events' e.g. wrong side surgery. There are other risks that can never be completely eradicated e.g. deep vein

thrombosis and surgical site infection. The incidence of these risks should be continually audited and regular action taken in an attempt to reduce rates further.

Clinical effectiveness looks at the ‘success’ of a procedure. Such outcomes include mortality and morbidity rates. Effectiveness can be reported by the surgeon (e.g. the range of motion in a replaced joint) or the patient (e.g. quality of life outcome measures). Personal experience investigates what the patient thought about their treatment. Did they feel that they were treated with dignity and respect? Did they feel that medical staff communicated well with them?

From the three different ways of classifying ‘quality’, a matrix can be formed (table 1). The matrix includes examples of quality indicators used in orthopaedic surgery. The rest of this thesis will concentrate on ‘outcome’ measures. There will be no further discussion about ‘structure’ and ‘processes’, except where they act as a representative marker for an ‘outcome’ e.g. length of hospital stay.

Table 1. Matrix showing different ways to measure 'quality' in healthcare with orthopaedic examples

	Structure	Process	Outcome
Safety	Laminar flow theatre	Thoroughness of theatre cleaning	Surgical site infection rates
Effectiveness	Type of operating equipment	Turn-around time in theatre	Post-operative morbidity survey (POMS)
Expectation	Quality of food provision	Waiting time for surgery	Patient-centred outcome measures

In order for 'quality' data to be meaningful, it must be accurate. Some institutions may deliberately 'improve' their figures for both financial purposes and to maintain a good reputation. This is fraud. Other institutions may 'select' their patients to improve their overall outcome data. This is known as 'gaming'. For example, high-risk patients (e.g. diabetics) may be discouraged from having lower limb arthroplasty surgery due to a higher surgical site infection rate. This change in patient population undergoing surgery was seen in New York when the outcome of cardiac surgery was first published[8].

1.5 Risk adjustment for surgical outcome Measures

The outcome of surgery is not simply a result of the effectiveness of care (both structure and process). Patient factors and random variation also play a role[9]. In order to adjust for confounding patient factors, risk adjustment tools have

been established. These tools aim to eliminate patient factors, making outcome a measure of the effectiveness of care and random variation alone. However, in practice not all patient factors can be excluded since some factors cannot be taken into account. Random variation cannot be excluded.

Several risk adjustment tools have been developed. They aim to predict which patients are at risk of higher morbidity and mortality. Most also aim to quantify this risk. Two tools commonly used to adjust for risk in orthopaedic surgery are the American Society of Anesthesiologists (ASA) Physical Status Classification and the Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM).

The American Society of Anesthesiologists Physical Status Classification (ASA-PS) was first published in 1941[10] and has been modified several times since. The ASA-PS divides patients into one of six groups (table 2).

Table 2. American Society of Anesthesiologists Physical Status Score (ASA 2008)

ASA grade	Criteria
I	Normal healthy patient
II	Patient with mild systemic disease
III	Patient with severe systemic disease
IV	Patient with severe systemic disease that is a constant threat to life
V	Moribund patient who is not expected to survive without the operation
VI	Declared brain-dead patient whose organs are being removed for donation

The ASA grade has been shown to correlate with morbidity in patients undergoing hip fracture surgery[11]. ASA-PS was originally devised as a predictor of anaesthetic risk for epidemiological studies and not as a predictor of outcome. ASA-PS is not suitable for predicting outcome on an individual basis.

First described in 1992, the POSSUM score is composed of eighteen components[12]. Twelve of the components relate to pre-operative physiological status and six relate to the severity of surgery. The scores for these two groups of components are entered into logistic regression equations to calculate the overall risk of morbidity and mortality. POSSUM is useful at predicting 'high risk' patients. It should not be used in isolation to suggest that an operation on a 'high-risk' individual is not worth performing. Other factors must be taken into account in this situation and it must be remembered that the figures for the POSSUM score are derived from population statistics, not individuals.

The lowest possible POSSUM mortality score for any intervention is 1.08%. It is well known that many surgical procedures have mortality rates lower than this[13]. It has been shown that POSSUM over-predicts the mortality rate by up to 6-fold in all patients with a predicted mortality rate of less than 10%[14]. For this reason, a new model of POSSUM, known as P-POSSUM was developed[15]. This predicts mortality rates more accurately than POSSUM[15]. A predictor of morbidity for P-POSSUM was not developed due to poor recording of post-operative morbidity. A further version of POSSUM has been produced specifically for orthopaedic patients[16].

1.6 Orthopaedic post-operative outcome Measures

1.6.1 Introduction

Describing the overall clinical impact of orthopaedic interventions can be done in a variety of ways, all of which have their limitations. Outcome measures aim to detect physiological (e.g. level of fitness), pathological (e.g. joint fixed flexion deformity), psychological (e.g. depression) and/or social (e.g. inability to work) factors. Outcomes may be measured in the short-term and long-term. There is no clear universal definition of 'short-term'. For the purpose of this thesis 'short-term' will refer to the duration of the inpatient stay.

1.6.2 Death

Surgical mortality has strengths as an outcome measure: it is easy to define and diagnose, accurate records of death are kept and it is a very important measure of outcome to both surgeons and patients alike. When considering death rates over longer time scales, it is important to compare this to the 'background' death rate of a similar population.

A National Joint Registry in England and Wales was established in 2003. It is now mandatory that data from all arthroplasty procedures be submitted to this registry. Mortality rates following primary hip arthroplasty according to the 2013 report are 0.25% at 30 days (95% CI 0.24% - 0.26%), 0.51% at 90 days (95%

CI 0.50% - 0.53%) and 1.52% at 1 year (95% CI 1.49% - 1.56%). Mortality rates following knee arthroplasty are 0.20% at 30 days (95% CI 0.19% - 0.21%), 0.36% at 90 days (95% CI 0.34% - 0.37%) and 1.15% at 1 year (95% CI 1.12% - 1.17%).

One drawback of mortality as a measure of outcome in comparative trials is its relative infrequency. For most types of orthopaedic surgery it is not a useful comparative index: the event rate is so low that very large numbers of operations would have to be compared to demonstrate a meaningful difference in outcome.

1.6.3 Length of hospital stay

Length of hospital stay is sometimes used as a surrogate for clinical outcome. It is easy to define and is recorded accurately by most hospital systems. The use of this measure as a surrogate marker of outcome is based on two assumptions: firstly, that all patients are discharged at same 'point' of recovery i.e. at a standard level of well-being; and secondly, that all patients who have reached this level of well-being are discharged. It is well known that these two assumptions are not true and length of hospital stay is influenced by many factors other than the health status of the patient[17,18].

1.6.4 Post-operative morbidity

Post-operative complications and morbidity following surgery are poorly recorded[19]. The post-operative morbidity survey (POMS)[20] was developed to

provide a true reflection of post-operative morbidity. It can be used to record general post-operative morbidity such as chest infection, urinary tract infection and pulmonary embolism as well as morbidity related to the surgical site such as wound infection.

The POMS has been validated as a measure of post-operative morbidity[21]. During this validation process, the authors examined whether post-operative morbidity was a set of unrelated, disparate phenomena, or whether they were linked by a common underlying pathology. It was thought that this pathology could be a mild variant of Multiple Organ Dysfunction Syndrome (MODS) resulting from a mild form of Systemic Inflammatory Response Syndrome (SIRS). A low level of internal consistency was found between the POMS domains, suggesting that a single underlying pathology is not being measured. Therefore the POMS should not be used as a one-dimensional scale and a score derived from the summation of POMS domains is not valid.

The POMS has been used to show that intra-operative gastric-to-end tidal carbon dioxide measurement may be a useful prognostic index of post-operative morbidity[22]. In another study, the POMS was used to find predictors of total morbidity burden on days 3, 5 and 8 after cardiac surgery[23]. Pre-operative albumin and haemoglobin levels, as well as weight, were found to be independently predictive of post-operative morbidity. This suggests that interventions aimed to improve these pre-operatively may lead to reduced post-operative morbidity and reduced health-care costs following cardiac surgery.

The POMS is also being used in a multicentre, prospective, blinded observational cohort study to investigate the effect of neoadjuvant chemotherapy and chemoradiotherapy on exercise capacity and outcome following upper gastrointestinal cancer surgery[24]. The primary endpoints are physical fitness and one-year mortality following surgery. Secondary endpoints are post-operative morbidity (assessed using the POMS on day 5) and patient related quality of life.

The POMS was designed to identify morbidity of a type and severity that could delay discharge from hospital. The data collection process is simple, allowing for routine screening of large numbers of patients. The POMS assesses indicators of organ system dysfunction (e.g. inability to tolerate an enteral diet) rather than traditional diagnostic categories (e.g. deep vein thrombosis). The survey assesses nine domains of morbidity (Table 3) using readily available data from observation charts, medication charts, patient notes, routine blood test results and direct questioning and observation of the patient. The POMS requires no additional investigations. The POMS has been shown to be reliable, valid and acceptable to patients[21].

Table 3. Criteria for a positive POMS score

Variable	Criteria for positive result
Pulmonary	Requires supplementary oxygen or ventilatory support
Infection	Currently on antibiotics or temperature over 38°C in the last 24 hours
Renal	Oliguria (<500ml/day), elevated creatinine (>30% pre-op level), catheter in-situ (for non-surgical reason)
Gastrointestinal	Unable to tolerate enteral diet for any reason
Cardiovascular	Diagnostic tests or treatment within the last 24 hours for: myocardial infarction, hypotension (requiring pharmacological therapy or fluids >200ml/hour), atrial/ventricular arrhythmia or cardiogenic pulmonary oedema
Central nervous system	Presence of new focal deficit, coma, confusion, delirium
Wound complications	Wound dehiscence requiring surgical exploration or drainage of pus from operative wound with or without isolation of organisms
Haematological	Requirement of blood transfusion, platelets, fresh frozen plasma or cryoprecipitate within the last 24 hours
Pain	Wound pain requiring parenteral opioids or regional anaesthesia

POMS has been used in outcomes[25] and effectiveness research[26]. In elective orthopaedic patients the POMS has shown that an increase in the revised cardiac index leads to an increase in non-cardiac post-operative morbidity[27]. The POMS has also shown that chronic kidney disease in orthopaedic patients leads to prolonged morbidity and increased hospital stay in a substantial

minority of patients[28]. A further study uses the POMS to show that lower socioeconomic status in orthopaedic patients does not lead to increased post-operative morbidity but does lead to a prolonged hospital stay with no morbidity present[29].

One study used the POMS to identify post-operative morbidity in a variety of surgical specialities[30]. This study included 289 orthopaedic patients. On post-operative day 3 the most common forms of post-operative morbidity were pain (30.8%), pulmonary morbidity (30.1%), infection (26.6%), renal morbidity (24.9%) and gastrointestinal morbidity (20.1%). The incidence of morbidity decreased with time. By post-operative day 15 the incidence of pain was 0.7%, pulmonary morbidity 1.7%, infection 7.6%, renal morbidity 1.0% and gastrointestinal morbidity 1.0%. This heterogeneous group included both elective orthopaedic and trauma patients. No further sub analysis of the data was performed.

The relationship between short-term generic clinical outcome and long-term quality of life outcome is not yet understood. One significant problem with quality of life measures is the time taken to collect data. The POMS provides early post-operative information but it is unknown if there is any correlation with longer-term measures. If short-term measures could predict longer-term quality of life, they could be used as early surrogate markers for longer-term function and well-being. At present both short-term and long-term outcome measures are needed to assess the success of any intervention.

1.6.5 Surgical Site Infection

Surgical site infection (SSI) is one form of post-operative morbidity. It is a major risk in orthopaedic surgery. SSI causes pain and can lead to wound dehiscence and generalized sepsis. Further surgery and admission to intensive care may be necessary. A patient with a SSI spends twice the average length of time in hospital[31]. SSI is therefore not only distressing for the patient; it is also an economic burden for the health care provider.

Superficial wound infection can spread to deeper tissues including bone. A deep infection diagnosed within the first six weeks of primary joint replacement (hip or knee) is treated with repeated joint washouts, replacement of polyethylene components and intravenous antibiotics. These interventions may be sufficient to eradicate the infection. If a deep infection is diagnosed after the sixth post-operative week, revision joint replacement surgery is normally required (one-stage or two-stage) together with a prolonged course of intravenous antibiotics[32]. If the infection cannot be eradicated, life-long antibiotics to suppress the infection, joint arthrodesis (fusion) or even limb amputation may be required.

Wound surveillance in Orthopaedic Surgery became mandatory in the NHS in England in 2004. Reported SSI rates depend on the method used for diagnosis, case mix, the thoroughness of surveillance and documentation and the length of patient follow-up. Patient follow-up is essential in any wound

surveillance program since half of SSIs present after hospital discharge[33]. Therefore SSI rates cannot be defined as a ‘short-term’ outcome measure. Long-term follow-up of patients must be established to ensure true rates are reported.

There is a misconception that SSIs are simple to define and diagnose. Several definitions of SSI have been proposed and diagnosis varies between surgeons. Diagnosis cannot rely solely on microbiology results since this would delay the initiation of appropriate treatment. There would also be false positive results from contaminants and false negative results when organisms fail to grow in the culture medium.

SSIs were traditionally diagnosed using the hallmarks of pain (dolor), redness (rubor), heat (calor), swelling (tumor) and impairment of function. As surgeons became increasingly accountable for their practice, more reliable and reproducible methods of diagnosing SSI became necessary. Three SSI definitions in use today are the American Centres for Disease Control (CDC) definition, the English Nosocomial Infection National Surveillance Scheme (NINSS) definition and the English ASEPSIS definition. ASEPSIS is an acronym for Additional treatment, Serous discharge, Erythema, Purulent exudate, Separation of deep tissues, Isolation of bacteria and Stay as inpatient prolonged over fourteen days.

The CDC definition[34] is used worldwide to classify wound infections. It includes any wound infection within 30 days of surgery or one year if an implant is present. The CDC definition divides SSIs into incisional and organ/space infections. Incisional SSIs are further divided into superficial and deep infections (see table 4). Although widely used, the CDC definition is weak since three out of the four diagnostic criteria are subjective. On psychometric evaluation CDC has been shown to be unreliable[35].

Table 4. CDC definition of surgical site infection

	Superficial infection (involving skin and superficial tissues)	**Deep infection** (involving the fascial and muscle layers)
Time period	Occurs within 30 days of surgery	Occurs within 30 days of surgery or within 1 year if implant present
Site	Involves only the skin and superficial tissue	Related to the surgical site and involves deep tissues
Further criteria	Must fulfil one of the following: • Purulent discharge from superficial incision • Organisms isolated from incision • Pain, tenderness, swelling, redness or heat around the incision AND the incision deliberately opened by a surgeon (unless cultures are negative) • Diagnosis by a surgeon or physician	Must fulfil one of the following: • Purulent discharge from deep incision • Spontaneous dehiscence or deliberate opening of a deep incision, following fever or pain or tenderness around the wound (unless cultures are negative) • Abscess involving a deep incision • Diagnosis by a surgeon or physician

The UK NINSS definition of SSI is based on CDC with two significant modifications. Firstly, pus cells must be present for a wound culture to be classified as positive. Secondly, a surgeon's diagnosis of infection is excluded as a sufficient criterion to diagnose SSI. These changes were implemented to improve the objectivity of CDC but reproducibility of NINSS remains low[36].

ASEPSIS is a quantitative wound scoring method developed in 1986[37]. It provides a numerical score that indicates the severity of wound infection. The score is calculated using objective criteria based on the wound's physical appearance (e.g. erythema and serous exudate) and the clinical consequences of infection (e.g. prolonged hospital stay and readmission) (tables 5 and 6). A score of over 10 indicates an increasing probability and severity of infection (table 7). The original ASEPSIS scoring method was psychometrically tested and found to be objective and repeatable[38]. The most recent revised version has not undergone the same evaluation.

Table 5. Points scale used to calculate ASEPSIS score

Criterion		Points
Additional treatment	Antibiotics	10
	Drainage of pus under local anaesthetic	5
	Debridement of wound under general anaesthetic	10
Serous discharge		0-5
Erythema		0-5
Purulent exudates		0-10
Separation of deep tissues		0-10
Isolation of bacteria		10
Stay in hospital over 14 days		5

Table 6. Point scale for ASEPSIS wound inspection score

	Proportion of wound affected					
	0%	>0 -19%	20-39%	40-59%	60-79%	80-100%
Serous exudate	0	1	2	3	4	5
Erythema	0	1	2	3	4	5
Purulent exudates	0	2	4	6	8	10
Separation of deep tissues	0	2	4	6	8	10

Table 7. Interpretation of total ASEPSIS score

ASEPSIS score	Meaning
0-10	No infection. Normal healing.
11-20	Disturbance of healing.
21-30	Minor infection
31-40	Moderate infection
41 and over	Severe infection

Scoring methods provide more detailed and objective information regarding SSI than CDC and NINSS but they are more costly, complicated and time-consuming to perform. The average time taken to collect the data and calculate an overall ASEPSIS score is 59 minutes[39].

One of the domains for the POMS score is 'wound infection'. A positive result is defined as 'wound dehiscence requiring surgical exploration or drainage of pus from operative wound with or without isolation of organisms'. This definition could under-estimate the true wound infection rate. The POMS 'wound infection' domain has not previously been compared to other definitions of surgical site infection to assess accuracy.

1.6.6 Quality of life outcome measures

Historically, the success or failure of an orthopaedic intervention was assessed and reported by the operating surgeon. This trend is now changing with greater emphasis on Patient Reported Outcome Measures (PROMs). There are

several PROMs available. They fall into one of three broad categories: generic, disease-specific and joint-specific. Generic surveys aim to investigate all aspects of quality of life and can be used to assess any medical or surgical intervention. Disease-specific tools concentrate on disability relating to a particular condition and aim to elucidate the impact of a single disease entity on a patient's quality of life. Joint-specific tools are used to assess the impact of one particular joint on the patient's quality of life. This thesis concentrates on patient outcome following hip and knee joint replacement surgery (arthroplasty). Therefore, hip-specific and knee-specific assessment tools will be described later in this section.

Analysis of PROMs data has traditionally been performed at the group level. Many recent studies have focused on individual patient reported outcomes. Individuals can be assessed in two ways: by responder criteria or by state-attainment criteria. With responder criteria, each patient is classified as a responder or a non-responder to treatment based on whether the change in health status exceeds a pre-defined threshold. With state-attainment criteria, a patient is classified not on the basis of change, but on whether a certain level of low symptom severity is attained. Research in both areas is experimental but may provide more relevant results than group level studies.

1.6.6.1 Generic quality of life outcome measures

Generic PROMs aim to assess all dimensions of health-related quality of life. The World Health Organization Quality of Life Group has recommended that

5 dimensions are assessed in any generic quality of life survey: physical health, psychological health, social relationship perceptions, function and well-being. Commonly used generic PROMs in orthopaedic literature include the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36), the Medical Outcomes Study 12-Item Short Form Health Survey (SF-12), the Nottingham Health Profile (NHP) and the European Quality of Life 5-Dimension (Euroqol) questionnaire.

1.6.6.1.1 The Medical Outcomes Study 36-Item Short Form Health Survey (SF-36)

SF-36 is a multi-purpose questionnaire used to measure general health status[40]. It was originally developed in American English. A United Kingdom English version is now available. The questionnaire refers to health over the previous 4 weeks but a more acute version that refers to health over the previous week is available. The questionnaire contains 36 questions each of which has between 2 and 6 answers. Each answer is awarded a score of between 0 (indicating poor health) and 100 (indicating good health). The questions are grouped into one of eight health concepts: bodily pain (BP), physical functioning (PF), role limitations due to physical health (RP), general health (GH), mental health (MH), vitality (VT), social functioning (SF) and role limitations due to emotional health (RE). There is also a health transition question that does not contribute to any of the 8 domains.

The 8 health concepts can be further amalgamated into 2 higher order groups. These are known as the Physical Component Summary (PCS) and the Mental Component Summary (MCS). The PCS is calculated from the BP, PF, RP and GH scores. It is most responsive to treatments that alter physical symptoms such as hip arthroplasty. The MCS is calculated from the MH, VT, SF and RE scores and is most responsive to drugs and therapies that target psychiatric disorders. Three of the scales (VT, GH and SF) have significant correlation with both the physical and mental summary measures.

SF-36 takes approximately 10 minutes to complete. It is proven to be suitable for self-administration, computerized administration or administration by an interviewer either in person or by telephone. Scores are calculated by summated ratings and standardised SF-36 algorithms. Individual question scores are summated without standardisation or weighting. Standardisation is avoided by using questions with roughly similar means and standard deviations, and weighting is avoided by selecting equally representative questions.

SF-36 has been evaluated in several studies. It is proven to be valid and reliable[41,42], sensitive and reproducible[40]. It has been used in over 4,000 publications assessing over 200 different diseases. SF-36 has been specifically investigated in patients undergoing hip arthroplasty where it was shown to be both valid[43] and reliable[44]. However, these studies also showed that SF-36 has minor 'floor' and 'ceiling' effects[45,46]. 'Floor' effect refers to the situation where a questionnaire is unable to measure a negative value that is lower than the

range provided in the choice of answers. In this situation, if a patient reports the lowest value for a question and then deteriorates further, the deterioration will not be detected by the questionnaire. 'Ceiling' effect refers to the opposite situation, where a questionnaire is unable to measure a positive value that is higher than the range provided in the choice of answers. In this situation, if a patient reports the highest value for a question and then improves, the improvement will not be detected by the questionnaire.

1.6.6.1.2 The Medical Outcomes Study 12-Item Short Form Health Survey (SF-12)

SF-12[47] is an abridged version of SF-36 containing 12 out of the 36 questions. SF-12 questions can be amalgamated to produce profiles of the eight SF-36 health concepts but only if the sample size is sufficiently large. SF-12 scores are calculated using weighted algorithms (i.e. the questions in SF-12 contribute different values to the overall score, unlike SF-36) and a computer program is available for this calculation.

The main advantage of SF-12 over SF-36 is that it is shorter. It is therefore quicker for patients to complete and quicker for research personnel to record and analyse data. A disadvantage is that a computer program is necessary for scoring each survey. A further disadvantage of SF-12 is that it has less construct validity and sensitivity than SF-36, producing less precise scores for the 8-scale health profile[47]. This is less important in large group studies since

the confidence intervals are largely determined by sample size but could result in insignificant findings in smaller studies.

1.6.6.1.3 Nottingham Health Profile (NHP)

The Nottingham Health Profile (NHP) is a self-administered questionnaire that takes 5 to 10 minutes to complete. It was developed in United Kingdom English and consists of 2 parts. Part I contains 38 'yes/no' items covering 6 dimensions: pain, physical mobility, emotional reactions, energy, social isolation and sleep. Part II contains 7 'yes/no' questions concerning activities of daily living. Each part is scored using weighted values giving a score of 0 (no problems at all) to 100 (presence of maximal problems). The two parts can be used independently. The NHP is proven to be internally consistent, valid, reproducible and sensitive[48]. No psychometric analysis of the NHP has been performed on patients undergoing hip or knee arthroplasty.

Overall, the NHP has one major disadvantage when compared to SF-36 concerning the response format. The NHP uses dichotomous 'yes/no' responses, where as SF-36 has several choices for each response. This allows the SF-36 to detect positive as well as negative states of health. The NHP often explores only ill health. A patient with an initial acceptable NHP score who makes an obvious clinical improvement may fail to show a change in the NHP score. Furthermore, a false negative response is more likely with the NHP when a patient with good function must respond on a scale that only assesses dysfunction. The dichotomous NHP response format produces higher 'ceiling'

effects in all dimensions when compared to SF-36[49]. They both have equal minor 'floor' effects.

1.6.6.1.4 EuroQol

The EuroQol (EQ-5D)[50] questionnaire contains 2 pages. There are 15 questions on the first page regarding 5 aspects of general health: mobility, self-care, usual activities, pain and depression. Each question has 3 possible answers: ‘no problem’, ‘moderate problem’ or ‘extreme problem’. The second page of EuroQol aims to elucidate the overall health of the patient. It contains a visual analogue scale with 0 indicating the worst possible health and 100 indicating the best possible health.

EuroQol was designed to be self-administered and takes 5 minutes to complete. It has been shown in studies to be both valid[51] and reliable[52]. However, EuroQol suffers from ‘ceiling’ effects due to the restricted response format. This effect is partially overcome by the use of the visual analogue scale on the second page. There is limited psychometric analysis of the questionnaire in patients undergoing hip and knee arthroplasty. Test-retest reliability has been shown[53] and there is some evidence of construct validity and responsiveness[54].

1.6.6.2 Disease-specific quality of life outcome measures

Disease-specific PROMs provide patient-centred information about a particular disease. This allows comparison of different surgical and medical treatment

options for that particular disease entity. They can be used in both research and clinical practice.

PROMs commonly used to assess hip and knee arthritis are the Western Ontario and MacMaster Universities (WOMAC) Osteoarthritis Index and the Arthritis Impact Measurement Scales (AIMS). These measures are specific for arthritis and can be used to assess any joint and any intervention.

1.6.6.2.1 The Western Ontario and MacMaster Universities (WOMAC) Osteoarthritis Index

The WOMAC Index was developed in Canadian English for patients with osteoarthritis[55]. The original version contained 5 subscales (WOMAC 5.0) but only three of these were retained for further development (WOMAC 3.0). Globalisation of WOMAC resulted in several refinements and the production of WOMAC 3.1. This has been the standard form of the Index for several years. It is self-administered and contains 24 questions covering three dimensions: pain, stiffness and physical function. Other versions of the Index are available with differing numbers of questions and dimensions to meet different measurement needs. The standard version uses a 48-hour timeframe but the Index is sufficiently robust to tolerate a variation in timescale of 24 hours to 1 month. The WOMAC Index is available in a 5-point Likert, 100mm visual analogue and 11-point numerical rating format.

The WOMAC Index is valid, reliable, responsive, easy to complete and simple to score[56]. It has been used in several hundred publications and has been translated and linguistically validated in over 65 languages. Most clinical research uses the Likert and visual analogue versions of WOMAC 3.1. WOMAC has been extensively evaluated in patients undergoing hip arthroplasty where it has been shown to be responsive[57], have high internal consistency and acceptable test-retest reliability[58]. However, it shows post-operative ceiling effects for the pain and stiffness subscales in patients undergoing hip arthroplasty[59], the same as with SF-36 and Euroqol. One study showed WOMAC to be superior in responsiveness to generic outcome measures[60] but disease-specific and generic outcome measures are generally used for different purposes. The use of both instruments together provides more information than using either individually.

1.6.6.2.2 The Arthritis Impact Measurement Scale (AIMS)

The Arthritis Impact Measurement Scale (AIMS)[61] was developed in American English to measure outcome in patients undergoing treatment for rheumatic disease. It has since been shown to be sensitive to clinical improvement in patients with osteoarthritis.

AIMS 2[62] is a modified version of the original AIMS, developed in American English. It was designed to be self-administered and takes 20 minutes to complete. The questionnaire contains 78 questions exploring 12 concepts: mobility, walking and bending, hand and finger function, arm function, self-

care tasks, household tasks, social activity, support from family and friends, arthritis pain, work, level of tension and mood. The scores of each question within a health concept are simply added. The range of scores for each concept depends upon the number of questions it contains. In order to express each scale in the same units, a normalization procedure is performed so that each concept is expressed as a value from 0 to 10 with 0 representing good health status and 10 representing poor health status.

The 12 concepts can be combined into 3 or 5 component models of health status. The 3-component model groups the concepts into general categories of physical function, psychological status and pain. The 5-component model combines the concepts into measures of lower limb function, upper limb function, affect, symptoms and social interaction. AIMS2 has been psychometrically evaluated and been shown to be both valid and reliable[61].

A further version of AIMS has been produced specifically for patients undergoing hip arthroplasty[63]. This questionnaire consists of 57 items that are scored and weighted to produce 4 subscales: physiologic function, self-concept, role function and interdependence. Responsiveness, content and construct validity have been proved for this version of AIMS[63] but reliability and other forms of validity remain unproven.

1.6.6.3 Joint-specific outcome measures

Historically, the operating surgeon assessed outcome following hip and knee arthroplasty. For hip surgery, tools such as the Harris Hip Score[64] and Charnley Score[65] were used. For knee surgery, the American Knee Society Score[66] and the Bristol Knee Score[67] were used. These scores were derived from clinical and radiological data and ultimately depended on the judgment of the surgeon. Patients and surgeons often differ in opinion[68] and it became apparent that methods were needed to elicit the patient's perception of their surgery. This lead to the design of joint-specific PROMs. For the hip, the Oxford Hip Score (OHS)[69], the Hip disability and Osteoarthritis Outcome Score (HOOS)[70] and University of California at Los Angeles (UCLA) hip score[71] were developed. For the knee, the Knee disability and Osteoarthritis Outcome Score (KOOS)[72] and the Oxford Knee Score (OKS)[73] were devised. These measures are commonly used in orthopaedic literature. This creates a problem when different PROMs are grouped together in meta-analyses. The heterogeneity of the responses may create bias and result in incorrect conclusions being drawn[74].

1.6.6.3.1 Harris Hip Score

The Harris Hip Score[64] was developed in American English as a means of assessing patients following hip arthroplasty. The assessment is made up of 8 questions and a physical examination. The questions cover 3 dimensions: pain (with a maximum score of 44), function (with a maximum score of 33) and level of activity (with a maximum score of 14). The physical examination assesses

hip range of motion and a maximum of 9 points can be awarded. The point score in each section are summated to give a maximum possible score of 100. A score of 90-100 is ‘excellent’, 80-89 is ‘good’, 70-79 is ‘fair’, 60-69 is ‘poor’ and a score of less than 60 is a ‘fail’.

The original version of the assessment was performed entirely by the surgeon. This has now been modified to create a PROM. This contains 7 questions regarding hip pain, walking aids, limping, walking distance, climbing stairs, putting on socks and shoes, and sitting. Each question has between 3 and 7 answers that are expressed on a Likert-type scale. The scores from each question are summated to give a total score of between 0 and 100, where 0 represents the best result. The Harris Hip Score was found to be both valid and reliable in the assessment of outcome of total hip replacement[75].

1.6.6.3.2 Charnley Score

Devised in 1972 in the United Kingdom, the Charnley Score[65] grades hip pain, mobility and walking on a 6-point scale. Walking is only assessed in patients who have no other condition that may affect their walking ability. The surgeon performs the assessment. Lower scores indicate greater disability. Scores for different treatment groups can either be averaged or state-attainment criteria can be used where the number of patients scoring 5 or 6 in each group can be compared.

This assessment is simple to perform but reflects the opinions of both the surgeon and the patient. There is little psychometric testing of the Charnley Score to support its use.

1.6.6.3.3 Oxford Hip Score

The Oxford Hip Score (OHS)[69] is a joint-specific PROM designed to assess disability in patients undergoing total hip replacement. The survey was developed in the United Kingdom and takes approximately 5 minutes to complete. It assesses the patient's health over the previous 4 weeks.

The OHS contains 12 questions each with 5 Likert-type response choices. The questions assess pain and functional ability from the patient's perspective. The overall score ranges from 12-60 with a higher score indicating increasing disability.

The psychometric properties of the OHS have been rigorously examined. It has been used extensively in orthopaedic literature and is highly sensitive to patients undergoing primary total hip replacement[59] and revision total hip replacement[76]. The OHS is internally consistent, reproducible and valid[77].

1.6.6.3.4 The Hip Disability and Osteoarthritis Outcome Score

The Hip disability and Osteoarthritis Outcome Score (HOOS)[70] is a 40-item PROM based on the Knee disability and Osteoarthritis Outcome Score (KOOS). The questionnaire is self-administered and takes 8-10 minutes to complete.

Each question has 5 possible answers and is scored 0 - 4. The questions are grouped into 5 subscales: pain, other symptoms, activities of daily living, sport and hip related quality of life. The score for each subscale is simply the sum of the individual question scores. HOOS is then transformed onto a 0 - 100 scale with 100 indicating the best possible outcome.

HOOS contains all the WOMAC Likert 3.0 questions in an identical form so can also be used to calculate WOMAC scores. HOOS has been shown to be both valid and responsive and 2 of the subscales (pain and other symptoms) have been shown to be more responsive than WOMAC[70].

1.6.6.3.5 The University of California at Los Angeles Hip Scale

The University of California at Los Angeles (UCLA) hip scale[71] is a PROM that has been routinely used to assess post-operative outcome in arthroplasty patients. More recently it has been used to assess hip arthroscopy outcome.

The scale explores 4 dimensions: pain, walking, function and activity. There are 10 points on the scale with 10 indicating the best outcome and 1 indicating the worst. There is little psychometric evidence in the literature validating the UCLA hip scale following arthroplasty, but many studies still use it as a measure of outcome.

1.6.6.3.6 The American Knee Society Score

The American Knee Society Score (AKSS)[66] consists of 2 parts. The first part is a knee score that considers pain, stability and range of movement. The maximum number of points available for this section is 100 for a well-aligned knee with no pain, 125º of flexion and negligible antero-posterior and medio-lateral instability. There are point deductions for flexion contractures, extension lag, pain, limited range of motion, instability and mal-alignment.

The second part of the AKSS is a function score that utilises walking distance and stair climbing as its main parameters. The maximum score is 100. This is given to those who can walk an unlimited distance unaided and ascend stairs normally. There are point deductions for walking and stair-climbing limitations and the use of walking aids.

The AKSS puts patients into 1 of 3 categories: those with no contra-lateral knee disease; those with contra-lateral knee disease; and those with multiple joint disease. The AKSS is designed so the knee score is independent of other joint disease, but the function score can be influenced by pathology in other joints. The AKSS has been shown to be sensitive[78], and have high inter- and intra-observer reliability[79].

1.6.6.3.7 The Bristol Knee Score

The Bristol Knee Score consists of 4 subscales: pain, function, movement and

deformity. There is a maximum score of 15 for pain, 20 for function, 10 for range of movement and 15 for deformity. A total score of 41-50 is deemed 'excellent', a score of 36-40 'good', 30-35 'fair' and a score of less than 30 'poor'. The Bristol Knee Score is not widely used in the literature and there are no validation studies for its use following arthroplasty.

1.6.6.3.8 The Knee disability and Osteoarthritis Outcome Score

The Knee disability and Osteoarthritis Outcome Score (KOOS)[72] was developed as an extension of the WOMAC index to evaluate short-term and long-term symptoms and function in patients with knee injury and osteoarthritis. It was developed at Lund University, Sweden and the University of Vermont, USA. American English and Swedish versions were developed simultaneously. The score is made up of 5 subscales: pain (9 items), other symptoms (7 items), function in daily living (17 items), function in sport and recreation (15 items) and knee-related quality of life. KOOS has been validated in patients undergoing total knee replacement[80]. The questionnaire has been modified for the foot and ankle (FAOS) and the hip (HOOS).

KOOS is self-administered and takes approximately 10 minutes to complete. A Likert-type scale is used for all answers. There are 5 answer options ranging from 0 (no problems) to 4 (extreme problems). Scores are transformed to a 0-100 scale with 0 representing extreme problems and 100 representing no problems. An aggregate score is not calculated since the 5 dimensions do not measure the same entity and therefore the 5 dimensions must be considered

separately. KOOS has been shown to be valid[81], sensitive to change[82] and responsive[72].

1.6.6.3.9 Oxford Knee Score

The Oxford Knee Score (OKS) was devised in the United Kingdom [73]. It consists of 12 questions, each of which has a choice of 5 answers. Each question was originally scored from 1 to 5, with 1 representing the best outcome and 5 the worst. This gave an overall possible score of 12 to 60 with 12 representing the best outcome.

This scoring system was deemed too confusing, so now each question is scored from 0 to 4 with 4 representing the best possible outcome. This gives an overall possible score of 0 to 48 with 48 representing the best possible score.

The OKS has been shown to be short, practical, reliable, valid and sensitive to change[83]. However, it is also influenced by proximal pathology such as hip or spine problems[84].

1.7 The use of Outcome Measures as Bed Utilisation Tools

Appropriately timed discharge of patients following surgery is essential for optimal patient care and efficient hospital functioning. A patient that is discharged too early is at risk of under-diagnosis of post-operative morbidity

with consequent adverse outcome. A patient whose discharge is delayed is at risk of developing a hospital-associated complication (e.g. hospital-acquired infection) and incurs an unnecessary cost to the health-care provider. Post-operative patients should be discharged at the earliest safe opportunity to reduce such complications and minimise the cost of each inpatient episode. Appropriate discharge timing should increase the throughput of patients and reduce waiting times.

Many patients remain in hospital with no medical indication[18]. One study showed that 31% of post-operative patients remained in hospital inappropriately[85]. Payment by Results aims to reduce this figure by rewarding efficiency and encouraging increased activity.

In order to improve efficiency, hospitals must first recognise inappropriate bed occupancy. The Post-Operative Morbidity Survey (POMS)[21] is the only validated prospective method of measuring short-term post-operative morbidity in the literature. The POMS was designed to identify morbidity of a type and severity that could delay discharge from hospital.

The most commonly used tool to assess appropriate bed utilisation in the literature is the Appropriateness Evaluation Protocol (AEP)[86]. The AEP is a retrospective tool that relies on data from the inpatient record. It has been shown to be valid and reliable in some studies[86] but not in others[87]. The POMS is a prospective tool that could be used in real time to assist with appropriate

patient discharge. This is in contrast to the AEP, which is a retrospective tool that can only be used to evaluate past events. Data for the POMS is collected directly from contemporary data sources during the inpatient stay where as the AEP relies solely on past patient records and is therefore dependent on completeness and accuracy of record keeping for reliable functioning.

The AEP has been used in several European countries to examine bed utilization[88 89 90 91]. In the USA, over 98% of post-operative inpatients had morbidity defined by the POMS[21]. This implies that patients with a POMS score of zero are fit for discharge. Therefore, as well as providing useful clinical research and audit data, the POMS may have utility for assessing and improving hospital bed utilisation.

1.8 Validation of outcome measures

Outcome measures must fulfil certain psychometric criteria to be good descriptors of clinical or quality-of-life phenomena. They must be reliable, valid and sensitive to change. They should be acceptable to patients, simple, easy to use and score, and preferably short.

Reliability is a term used inconsistently in the literature. It is a measure of the degree to which subjects can be distinguished from each other. It can be defined as the ratio of the variance between subjects to the total variance. A reliability value of zero indicates a completely unreliable measure and a reliability value of one indicates a perfectly reliable measure.

Reliability is dependent on the relationship between the measurement error and the variability between subjects. Therefore, internal consistency and reproducibility are both components of reliability. Internal consistency determines whether a survey measures a single variable. The test used to measure internal consistency is Cronbach's alpha[92]. This summarizes the internal correlation of all questions in a survey onto a single scale. The higher the alpha coefficient, the greater the likelihood the questionnaire is tapping into a single variable and is free from random error.

Reproducibility investigates if a questionnaire produces the same results if repeated under the same conditions. Inter-observer reliability (agreement between two or more observers on the same occasion), intra-observer reliability (same observer on separate occasions), and test-retest reliability (stability of the measure over time in the same subject) are all aspects of reproducibility. Paired sets of data can be compared using the kappa coefficient or the coefficient of reliability according to the method of Bland & Aitman[93]. A higher coefficient indicates a more reproducible questionnaire.

Validity examines whether a questionnaire measures what it purports to measure. Several types of validity exist: content and face validity, criterion (convergent/concurrent) validity and construct validity. Face and content validity assess whether a survey fully investigates the intended topic of interest. Content validity can be increased by conducting exploratory interviews with

patients prior to writing the questionnaire. This will elucidate the priorities and concerns of patients rather than imposing clinical assumptions. Face and content validity are subjective measures with no statistical methods to assess them.

Criterion validity examines how a new measure relates to an established “gold standard” in the field. This approach can only be used when a “gold standard” exists which begs the question of why a new measure is being developed. For measures where no “gold standard” exists, construct validity examines the extent to which the results from the questionnaire support a predefined hypothesis. It can be measured using Pearson correlation coefficients between the total score for the questionnaire and other measures considered to be associated with the underlying notion being investigated.

Construct validity investigates if a single concept is being measured by the questionnaire. If construct validity is proven, scores can be combined to produce one overall score. Construct validity is tested by calculating the correlation between scale scores.

Sensitivity to change or responsiveness indicates whether the survey is able to detect clinically significant changes. It can be assessed by comparing outcome scores before and after an intervention and can be defined as the difference between the mean preoperative and post-operative scores divided by the

standard deviation of the preoperative scores. An effect size of one is equal to a change of one standard deviation in the sample.

1.9 Plan of investigation

This research project begins with an observational cohort study investigating morbidity following hip and knee arthroplasty. It describes the prevalence and pattern of post-operative morbidity for different types of hip and knee arthroplasty. It investigates whether there is any difference in post-operative morbidity between the different arthroplasty groups.

This initial study revealed that many post-operative arthroplasty patients remained in hospital with no apparent morbidity. This led me to consider the use of the POMS as a bed utilisation tool and its ability to identify inappropriate bed occupancy following hip and knee arthroplasty. The number of inappropriate bed occupancy days for each type of arthroplasty is presented together with the potential cost saving if patients with no morbidity were discharged at the earliest opportunity. The reasons patients remained in hospital with no identifiable morbidity are reported. Particular attention is given to 2 groups of patients: those who developed morbidity following a period with no morbidity and patients re-admitted to the same hospital within one year of surgery. Inappropriate bed utilisation rates are compared to other bed utilisation studies.

After assessing the utility of the POMS as a bed utilisation tool, I went on to assess whether it could act as an early surrogate marker for patient–reported outcome measures (PROMs). One disadvantage of PROMs is the time taken to collect and analyse data. If an earlier surrogate marker for long-term PROMs could be found, operative and peri-operative procedures could be assessed more quickly. At present there is no evidence that morbidity is associated with longer-term patient reported outcome. I decided to investigate whether there is any association between POMS results in the early post-operative period and PROMs scores at 18 months post-surgery. Traditionally, length of hospital stay has been used as an early marker of the success of a procedure. I decided to also examine whether there is any association between length of hospital stay and PROMs scores to justify its use for this purpose.

At the same time as investigating the POMS, I was given access to surgical site infection data on all trauma and orthopaedic inpatients at the same institution. Surgical site infection data was available on the entire study population in the POMS study. I decided to compare the 'wound' domain of the POMS to the ASEPSIS score. ASEPSIS is a validated measure of surgical site infection. I calculated the sensitivity, specificity, positive predictive value, negative predictive value and overall accuracy of the 'wound' domain of the POMS compared to the ASEPSIS score.

Discrepancies between the 'wound' domain of the POMS and the ASEPSIS score led me to consider how reliable and comparable other methods of

diagnosing wound infection are. I decided to assess if three commonly used methods of diagnosing surgical site infection (CDC, NINSS and ASEPSIS) report similar rates of infection in the same series of patients. Different institutions use different methods to assess SSI. If published SSI rates are to be meaningful, different definitions of SSI must give similar values to ensure that comparisons between surgeons and hospitals are fair.

1.10 Summary

1. Outcome measures evaluate the benefits and risks of surgery.
2. Outcome measures are necessary for research and audit purposes and to guide patients in their choice of surgeon and institution.
3. The routine collection and reporting of outcome measures started in the UK in 2000 and has gradually evolved over the last fourteen years.
4. Quality of surgery can be classified in different ways. Quality is commonly divided into three dimensions: structure, process and outcome. The Department of Health divided quality into three different dimensions: safety, effectiveness and expectation.
5. When interpreting outcome data, adjustment must be made for patient risk factors. Two commonly used risk adjustor scores in orthopaedic literature are ASA-PS and POSSUM.
6. Post-operative outcome measures include mortality rates, morbidity rates and patient-reported outcome measures.
7. The Post-Operative Morbidity Survey is the only validated measure of

post-operative morbidity in the UK. It recognises morbidity of a severity that prevents discharge from hospital. As such, it may have utility as a bed utilisation tool.

8. Quality of life outcomes can be subdivided into generic measures, disease-specific measures and joint-specific measures. They can be reported by surgeons or patients. Patient-reported measures are less prone to bias.
9. Outcome measures must be psychometrically evaluated to ensure they are reliable, valid and sensitive to change.

Chapter 2 Post-operative morbidity in Orthopaedic Patients

2.1 Introduction

In this chapter I will report results from a prospective observational cohort study investigating morbidity following hip and knee arthroplasty. In the introduction I will describe the common types of hip and knee arthroplasty including hip resurfacing, partial knee replacement, total joint replacement and revision joint replacement. I will report how morbidity is described in comparative arthroplasty studies.

I will then proceed to describe the methodology of this cohort study. I will describe the study setting, the study population, the data collection process and data analysis. The same methodology and data is used for the bed utilisation analysis presented in Chapter 3.

Following this I will present the results. I will describe the characteristics of the study population. I will report the prevalence and pattern of post-operative morbidity for each type of arthroplasty. Statistical analysis will assess if there is any difference in morbidity between the different arthroplasty groups.

I will conclude with a summary and discussion of the findings. I will compare my results to other reports of morbidity following arthroplasty in the literature.

I will comment on the limitations of the study and suggest areas for further investigation.

2.1.1 Reporting of morbidity following arthroplasty in peer-reviewed journals

Morbidity following major surgery is inconsistently and poorly reported in the literature[94]. In some randomised controlled trials comparing joint prostheses, reported in peer-reviewed journals, there is no mention of post-operative morbidity[95]. However, this is unusual and most trials make an attempt to report 'post-operative complications'[96]. These complications usually refer to joint specific problems such as dislocation or deep prosthetic infection rather than morbidity in other organ systems e.g. the respiratory or cardiovascular systems.

There is no consistency in the reporting of post-operative morbidity in orthopaedic literature. Papers report different events as a 'complication' and different time-scales are used. One complication may be reported in one paper, and a different complication in another paper. This makes direct comparisons between papers both inaccurate and misleading.

The CONSORT (Consolidated Standards of Reporting Trials) statement[97] provides a standardised framework for randomised controlled trials. There is evidence that this framework improves the standard of randomised controlled trials but the majority of surgical research does not adhere to it[98]. An extension

to the CONSORT statement[99] suggests that data regarding 'harm' should be reported. A standardised definition of 'harm' should be used and the method of collecting 'harm' data should be clearly reported. The statement emphasised that the personnel performing the assessment should be identified and the frequency and time frame of data collection reported. At present, the only validated method of recording post-operative 'harm' is the Post-Operative Morbidity Survey (POMS)[21]. The POMS has not previously been used to describe and compare morbidity between different types of hip and knee arthroplasty.

2.1.2 Types of arthroplasty

'Arthroplasty' is an alternative word for joint replacement. Joint replacements are usually performed for arthritis: osteoarthritis and inflammatory arthritis. There are more rare indications such as metabolic bone disease, post-traumatic arthritis, avascular necrosis and tumour. Arthroplasty is usually performed when a patient has significant joint pain, not relieved with simple analgesia, that is affecting their activities of daily living.

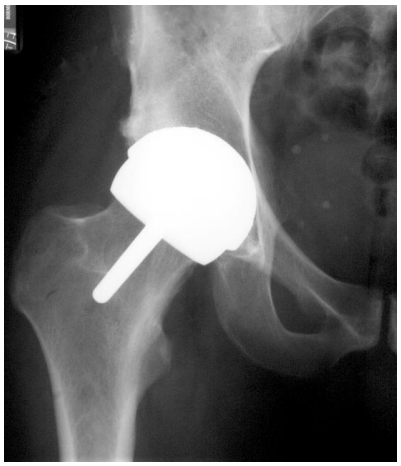
This study looks at three types of hip arthroplasty (hip resurfacing, total hip replacement and revision total hip replacement) and three types of knee arthroplasty (unicondylar knee replacement, total knee replacement and revision total knee replacement). A brief description of each type of arthroplasty is given below.

2.1.2.1 Hip resurfacing

Hip resurfacing involves placing a metal cap over the head of the femur and a metal lining in the acetabulum (Images 1 and 2). The articular bearing surface is metal-on-metal. Resurfacing evolved as an alternative to total hip replacement with several perceived advantages. Firstly, there is restoration of normal anatomy allowing for a good range of hip motion and a stable joint with a low chance of dislocation. Secondly, there is preservation of more femoral bone than in total hip replacement. This can make revision surgery easier in the future.

Image 1. Photograph of hip resurfacing implant

Image 2. X-ray showing hip resurfacing implant



There are several risks associated with hip resurfacing and patient selection and counselling are very important. Hip resurfacing should only be performed on young patients with no evidence of osteoporosis and no bone cysts in the femoral neck. This reduces the risk of femoral neck fracture. The success of hip resurfacing is not only dependent on patient-factors but also on surgeon-factors. The outcome of surgery is highly dependent on the operating surgeon's experience[100] and also on correct alignment of the implants[101].

There have been recent controversies regarding hip resurfacing. High levels of metal ions have been found in some patients' blood[102]. It has been hypothesised this may lead to an increased risk of cancer but at present there is no evidence to support this[103]. The formation of 'pseudotumours' due to an immune-mediated response has also been reported[104]. These tumours destroy soft tissue including muscle around the hip joint making revision surgery very difficult. These problems have lead many surgeons to abandon hip resurfacing over the last two years. The National Joint Registry shows that the failure rate of resurfacing implants at 7 years following surgery is 11.81%[105].

2.1.2.2 Total hip replacement

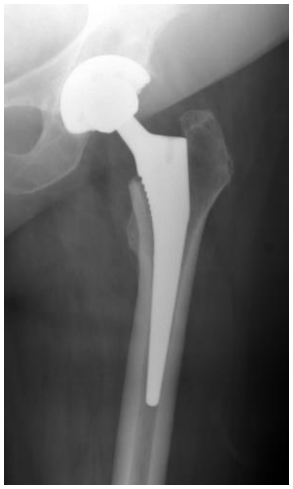
Total hip replacement is the most commonly performed orthopaedic procedure and generally improves both hip pain and mobility[106]. It involves replacement of the entire femoral head and replacement of the acetabulum. The prosthetic head is attached to a stem that is inserted into the femoral shaft.

Total hip implants can be cemented into bone or an uncemented technique can be used. Uncemented components are coated in an osteoconductive layer (hydroxyapetite) to encourage bony ingrowth. There are 3 commonly used bearing surfaces (the articular surfaces on the femoral head and acetabulum): metal-on-polyethylene (Images 3 and 4), metal-on-metal and ceramic-on-ceramic. They vary in their wear properties and longevity. Metal-on-metal total hip replacements have recently become less popular due to the same problems encountered with metal-on-metal hip resurfacing.

The National Joint Registry reports an overall failure rate of 3.08% for cemented implants at 7 years following surgery and a failure rate of 5.46% for uncemented implants at 7 following surgery[105].

Image 3. Photograph of metal femoral component of a total hip replacement

Image 4. X-ray showing metal-on-polyethylene total hip replacement



2.1.2.3 Revision total hip replacement

Revision total hip replacement is performed for failure of a primary total hip replacement. There are many reasons for revision surgery including pain, aseptic loosening, instability of the joint, component mal-alignment, infection of the primary implant and wear of the original components[107].

Revision surgery usually takes longer to perform than a primary procedure. A larger incision is generally needed and there is greater blood loss. Specialist surgical instruments and implants are often required, together with different forms of bone graft and metal augments. Revision surgery usually more invasive and technically more demanding than primary surgery and complication rates are generally reported to be higher[108-110].

2.1.2.4 Unicondylar knee replacement

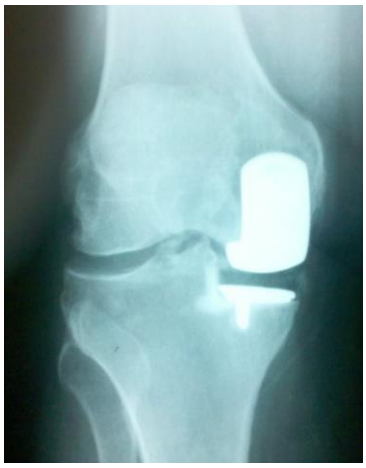
Unicondylar knee replacement is a 'half' knee replacement. It is indicated for patients with isolated arthritis in the medial or lateral compartment of the knee. It can be performed in the presence of patello-femoral arthritis. The underlying concept is to preserve as much of the normal knee as possible and to replace only the damaged parts. Knee kinematics are retained due to the preservation of all knee ligaments. One in four patients with knee osteoarthritis is a suitable candidate for unicondylar knee replacement[111].

Medial unicondylar knee replacements are much more common than lateral replacements. The femoral condyle and tibial plateau are resurfaced with metal components. A polyethylene spacer is inserted between the two (Images 5 and 6).

When compared to total knee arthroplasty, UKR has been shown to provide greater patient satisfaction[112], result in a shorter hospital stay[113], cause less blood loss and hence decreased blood transfusion requirement[114], result in a more speedy recovery[115] and cost less for the implant[113]. UKR preserves more bone stock than total knee arthroplasty, making revision surgery easier. The National Joint Registry reports failure rates of 16.64% of this type of arthroplasty at 7 years following surgery[105].

Image 5. Photograph of medial unicondylar knee implant

Image 6. X-ray of unicondylar knee arthroplasty



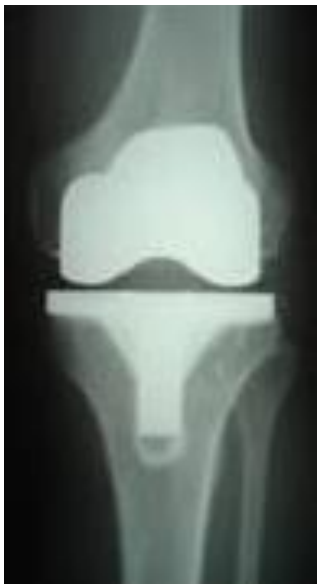
2.1.2.5 Total knee replacement

Primary total knee replacement involves resurfacing of the tibial plateau and femoral condyles (Images 7 and 8). The resurfacing components are usually made of metal but can be made from other materials such as oxinium (metal with a ceramic surface layer). The components may be attached to bone using a cemented or uncemented technique (similar to total hip arthroplasty). The

tibial and femoral components are separated by a polyethylene insert. This insert can be attached to the tibial component or mobile.

Image 7. Photograph of total knee replacements implants

Image 8. X-ray of total knee replacement



Total knee replacements are either unconstrained or constrained. Unconstrained implants rely on the native medial and lateral collateral ligaments of the knee for varus/valgus stability and can either be posterior-cruciate ligament (PCL) retaining or sacrificing. The main perceived advantage of PCL-retaining implants is improved range of motion due to preservation of the knee 'roll-back' mechanism. However, this sliding movement is also a disadvantage since it produces increased polyethylene wear. The main advantage of PCL-sacrificing implants is decreased polyethylene wear due to a higher degree of conformity between the implant components. As a disadvantage, the higher conformity results in a more restricted range of motion.

Constrained knee replacements are used when the medial and lateral collateral ligaments are non-functional. Constrained implants are either unhinged or hinged. The main disadvantage of these implants is greater constraint of movement that creates higher stress levels at the bone-implant interface. The components are therefore more prone to aseptic loosening and early failure. To overcome this problem long intramedullary stems are used on the femoral and tibial components. This spreads the stress at the bone-implant interface over a larger surface area.

Total knee arthroplasty is a very good procedure to improve pain[116] but total knee replacements do not feel like a 'normal' knee. This is due to irregular kinematics[117,118], abnormal patellar tracking[119] and a decreased range of

motion[120]. This is different to total hip arthroplasty where patients often report that the replaced hip feels the same as their native hip.

The National Joint Registry reports that the overall failure rate of cemented total knee replacements at 7 years following surgery is 3.81%. The failure rate of uncemented total knee replacements at 7 years following surgery is 4.75%[105].

2.1.2.6 Revision total knee replacement

Revision total knee replacement is necessary when a primary implant fails. The indications for revision surgery are similar to those for revision hip surgery. Polyethylene wear is more of a problem in TKR than THR since total knee replacements operate at the endurance level of polyethylene where as most total hip replacements operate well below it.

Similar to revision hip surgery, revision knee surgery is usually longer, requires a larger incision and incurs a greater blood loss than primary knee surgery. The aim is to preserve as much bone as possible and specialist equipment and implants are often needed. Bone graft and metal augments are often required. As with revision hip surgery, the risks of revision knee surgery are higher than primary knee arthroplasty[121].

2.2 Methods

2.2.1 General

This is a prospective longitudinal cohort study to evaluate the type and frequency of post-operative morbidity following lower limb arthroplasty using the Post-Operative Morbidity Survey (POMS). Ethical approval was gained from the Joint UCLH/UCL Committee on the Ethics of Human Research (reference number 01/0116) prior to commencement of the study. The requirement for patient consent was waived as collection of the POMS became a routine part of service evaluation within the institution.

2.2.2 Setting

The data in this study was collected from the Middlesex Hospital between March 1st 2004 and February 28th 2005. The Middlesex Hospital was part of the University College Hospital NHS Trust, London, UK. It was a central London teaching hospital. The Middlesex Hospital was closed in December 2005 when it merged with University College London Hospital and both hospitals moved into a new building.

2.2.3 Patients

All patients aged 18 or over undergoing elective lower limb arthroplasty were eligible for inclusion into this study. There were no exclusion criteria, ensuring a consecutive sample was taken. Elective lower limb arthroplasty procedures

included unicondylar knee replacement (UKR), total knee replacement (TKR), revision total knee replacement (RTKR), hip resurfacing (HR), total hip replacement (THR) and revision total hip replacement (RTHR).

2.2.4 Data collection

Data was collected by two study nurses. The age, sex, measures of pre-operative risk (ASA-PS score and variables needed to calculate the POSSUM score), surgical procedure, length of inpatient stay, mortality, post-operative destination (ward, High-Dependency Unit or Intensive Care Unit) and post-discharge destination (home, rehabilitation unit or another hospital / care institution) were recorded for each patient. This data was collected at the bedside and entered onto a standardised paper results form. A strong emphasis was placed on completing results forms in their entirety to ensure complete data sets were obtained. This data was entered into a Microsoft Access database (Microsoft Corporation, Redmond, WA, USA) in the Surgical Outcomes Research Centre in the Middlesex Hospital at a later date.

POMS data was collected on post-operative days (POD) 3, 5, 8 and 15 if the patient remained in hospital. Data was obtained from observation charts, medication charts, patient notes, routine blood test results and direct patient questioning, observation and examination. Further information was gained from the hospital clinical information system and consultation with staff looking after the patient. No additional investigations were required specifically for this study.

The staff looking after the patients (medical and nursing) were not aware of any study results i.e. they were blinded.

2.2.5 Data analysis

2.2.5.1 Patient characteristics

For characteristics with continuous variability, the mean and range are given. For other categories, the proportion of patients in each group is stated.

2.2.5.2 Pattern of POMS morbidity

For each type of arthroplasty, the proportion of the study population with post-operative morbidity as defined by the POMS is reported POD 3, 5, 8 and 15.

2.2.5.3 Statistical analysis

The aim of the analysis was to compare morbidity between different operation types. Not all comparisons between operations were of interest, with the focus on specific comparisons. UKR was compared to TKR and HR was compared to THR to assess if the newer 'less invasive' procedures (UKR and HR) were associated with less post-operative morbidity than total joint replacements (TKR and THR).

Primary joint replacement surgery is generally considered 'less invasive' than revision joint replacement surgery. To assess if revision surgery is associated

with increased post-operative morbidity, THR was compared to RTHR and TKR compared to RTKR.

To investigate if there is a difference in morbidity following hip and knee arthroplasty, analyses were performed between the TKR and THR groups and RTKR and RTHR groups.

The outcome variable in each comparative analysis was the occurrence of different types of morbidity. The occurrence of morbidity was a binary measure: either present or absent. Fisher's exact test was used for analyses.

At each time point (POD 3, 5, 8 and 15), there were six planned comparisons between the operation types. These multiple comparisons produce an increased risk of finding a difference due to chance alone. To allow for this issue a Bonferroni correction was applied to the p-values resulting from each statistical test. This involved multiplying the p-values by the number of tests performed (six in this case).

The difference in the POSSUM morbidity risk scores between the different operation types was also examined. These values were found to have a positively skewed distribution so the Kruskal-Wallis test was used for the analysis.

Further analyses examined the morbidity differences between operations after adjusting for the POSSUM risk score. These analyses were performed using logistic regression. Regression analyses were restricted only to comparisons between pairs of operations where an initial significant difference was observed. A Bonferroni adjustment was applied to the results of the regression analyses to allow for multiple testing.

2.2 Results

2.3.1 Study population characteristics

2.3.1.1 Knee arthroplasty patients

Table 8 shows a summary of the characteristics of the knee arthroplasty study population.

Table 8. Characteristics of 300 patients undergoing knee arthroplasty. All data expressed as percentage of patients for each procedure unless otherwise stated.

Characteristic		**Unicondylar knee replacement**	**Total knee replacement**	**Revision total knee replacement**
Number of patients in study		66	226	8
Mean age (range) / years		66.1 (45 – 87)	70.3 (23 – 90)	71.6 (46 – 88)
Sex (% Female)		54.5	60.6	75.0
ASA-PS score	I	24.2	13.3	25.0
	II	65.2	66.4	62.5
	III	7.6	18.6	0.0
	IV	0.0	0.0	0.0
	Missing	3.0	1.7	12.5
Mean POSSUM risk (range)	Morbidity	13.3 (6.5 - 39.7)	18.7 (9.3 – 53.0)	28.0 (10.7 – 71.5)
	Mortality	2.4 (1.3 – 5.5)	3.3 (1.7 – 9.8)	5.8 (1.4 – 19.2)
Post-op environment	ICU / HDU	3.0	1.3	12.5
	>1 day ICU	1.5	0.9	12.5
	Ward	97.0	98.7	87.5
Mean post-op length of stay (range) / days		5.5 (2 – 52)	5.6 (3 – 37)	22 (3 – 102)
Return to theatre		3.0	0.0	12.5
Return to ICU		1.5	0.0	0.0
Died in hospital		1.5	0.0	0.0
Discharge destination	Home	98.5	99.6	100.0
	Rehabilitation	0.0	0.4	0.0
	Other hospital	0.0	0.0	0.0
	Death	1.5	0.0	0.0

A total of 300 patients underwent knee arthroplasty procedures during the study period (66 UKRs, 226 TKRS and 8 RTKRs). The mean age of the study population was lowest in the UKR group (66.1), higher in the TKR group (70.3) and highest in the RTKR group (71.6). Sex distribution was roughly even in the UKR group. There was a higher proportion of females in both the TKR and RTKR groups (60.6% and 75.0% respectively). 1 UKR patient died after surgery. The cause of death was cardiac arrest on POD 2. Post-mortem examination showed the cardiac arrest was secondary to a pulmonary embolus. There were no deaths during the inpatient post-operative period in the TKR and RTKR groups.

The ASA-PS grading of patients varied throughout the three groups. There was a smaller proportion of ASA-PS Grade I patients in the primary TKR group (13.3%) compared to the UKR (24.2%) and RTKR (25.0%) groups. The proportion of ASA-PS Grade II patients was similar in the three groups (65.2% in the UKR group, 66.4% in the TKR group and 62.5% in the RTKR group). There was a higher proportion ASA-PS Grade III patients in the primary TKR group (18.6%) compared to the UKR (7.6%) and RTKR (0.0%) groups. No patients were ASA-PS Grade IV. The ASA-PS grading of 1 UKR patient, 2 TKR patients and 1 RTKR was not recorded.

The mean POSSUM risk scores for both morbidity and mortality were lowest in the UKR group (13.3 and 2.4 respectively), higher in the primary TKR group (18.7 and 3.3 respectively) and highest in the RTKR group (28.0 and 5.8

respectively). The proportion of patients admitted to the Intensive Care Unit (ICU) post-operatively was low in the UKR and TKR groups (3.0% and 0.0% respectively) but higher in the RTKR group (12.5%). There was only 1 patient admitted to ITU from the ward in the post-operative period. This was the UKR patient who had a cardiac arrest on POD 2 and subsequently died.

2 UKR patients returned to theatre. The first patient developed a wound infection and returned to theatre for an open knee washout and change of polyethylene insert. The second patient developed peritonitis secondary to a perforated gastric ulcer and was taken to theatre by the general surgical team for a laparotomy and ulcer repair. No TKR patients required a second surgical intervention. 1 RTKR patient required a second operation for infection. An open knee washout was performed together with exchange of polyethylene tibial insert. The indication for revision surgery was infection of the primary prosthesis.

Mean length of hospital stay was similar in the UKR and TKR groups (5.5 days and 5.6 days respectively). It was longer in the RTKR group (22 days). All UKR, RTKR and all but one TKR patients were discharged to their own home. The remaining TKR patient was discharged to a rehabilitation unit.

2.3.1.2 Hip arthroplasty patients

Table 9 shows the characteristics of the hip arthroplasty population.

Table 9. Characteristics of 229 patients undergoing hip arthroplasty. All data expressed as percentage of patients for each procedure unless otherwise stated.

Characteristic		Hip resurfacing	Total hip replacement	Revision total hip replacement
Number of patients in study		32	162	35
Mean age (range) / years		51.6 (22 – 70)	70.7 (21 – 91)	72.2 (26 – 88)
Sex (% Female)		50.0	60.5	62.9
ASA-PS score	I	53.1	11.1	8.6
	II	43.8	70.4	51.4
	III	3.1	15.4	37.1
	IV	0.0	1.2	0.0
	Missing	0.0	1.9	2.9
Mean POSSUM risk (range)	Morbidity	15.1 (7.6 – 60.6)	22.4 (9.3 – 65.7)	34.4 (9.3 – 90.5)
	Mortality	2.9 (1.4 – 14.1)	4.1 (1.7 – 16.5)	7.8 (1.7 – 42.1)
Post-op environment	ICU / HDU	0.0	4.9	14.3
	>1 day ICU	0.0	3.1	5.7
	Ward	100.0	95.1	85.7
Mean post-op length of stay (range) / days		6.4 (4 – 13)	10.2 (3 – 81)	20.6 (5 – 93)
Return to theatre		0.0	1.9	14.3
Return to ICU		0.0	3.1	2.9
Died in hospital		0.0	0.0	2.9
Discharge destination	Home	100.0	93.8	80.0
	Rehabilitation	0.0	6.2	17.1
	Other hospital	0.0	0.0	0.0
	Death	0.0	0.0	2.9

A total of 229 patients underwent hip arthroplasty procedures during the study period (32 HRs, 162 THRS and 35 RTHRs). The mean age of the study population was lowest in the HR group (51.6). The mean age of patients undergoing THR and RTHR was similar (70.3 and 71.6 respectively). Sex distribution was roughly even in the HR group but there was a higher proportion of females in both the THR and RTHR groups (60.5% and 62.9% respectively). There were no deaths in the immediate post-operative period in the HR and THR groups. There was one death in the RTHR group. This patient died on POD 3 following a cardiac arrest.

The ASA-PS grading of patients varied throughout the three groups. Over half the HR patients were ASA-PS Grade I (53.1%). This proportion was lower in the THR and RTHR groups (11.1% and 8.6% respectively). Most THR patients (70.4%) were ASA-PS Grade II. The proportion was lower in the RTHR (51.4%) and HR (43.8%) groups. 37.1% of RTHR patients were ASA-PS Grade III compared to 15.4% of THR patients and 3.1% of HR patients. 1.2% of THR patients were ASA-PS Grade IV. No HR or RTHR patients were ASA-PS Grade IV. The ASA-PS grading of 3 THR patients and 1 RTHR patient were not recorded.

The mean POSSUM risk scores for both morbidity and mortality were lowest in the HR group (15.1 and 2.9 respectively), higher in the primary THR group (22.4 and 4.1 respectively) and highest in the RTHR group (34.4 and 7.8 respectively). No HR patients were admitted to the Intensive Care Unit (ICU)

post-operatively. 4.9% of THR patients and 14.3% of RTHR patients were admitted to ICU following surgery. 5 THR patients were transferred to ICU from the ward. The reasons for transfer included cardiac monitoring for atrial fibrillation and supraventricular tachycardia, ventilatory support following a pulmonary embolus and ventilatory support following further surgery for a dislocated hip. 1 RTHR patient with generalised sepsis was transferred to ICU for cardiac monitoring.

No HR patients required a second surgical intervention during their inpatient stay. 3 THR patients required further surgery during their hospital stay: 2 patients underwent closed reduction of a dislocated hip and 1 patient had an open hip washout and exchange of polyethylene liner for infection. 5 RTHR patients required a second operation during their inpatient stay. 3 patients required open hip washouts with exchange of polyethylene liner for possible deep infection. The indication for revision hip surgery in these patients was infection of the primary prosthesis. The other 2 RTHR patients returned to theatre to address periprosthetic fractures sustained from falls: 1 patient underwent open reduction and internal fixation of the fracture and 1 patient had the femoral component re-revised.

Mean length of hospital stay was lowest in the HR group (6.4 days). This increased in the THR group (10.2 days) and was considerably longer in the RTHR group (20.6 days). All HR patients were discharged to their own home.

10 THR patients and 6 RTHR patients was discharged to a rehabilitation centre. The remaining patients were discharged to their own home.

2.3.2 Pattern of post-operative morbidity

2.3.2.1 Knee arthroplasty patients

Table 10 shows a summary of morbidity following knee arthroplasty according to the POMS.

Table 10. Post-operative morbidity (defined by the POMS) in 300 elective knee arthroplasty patients on days 3, 5, 8 and 15. Figures are a percentage of the total number of patients in each arthroplasty group.

	Unicondylar knee replacement (UKR)				Total knee replacement (TKR)				Revision total knee replacement (RTKR)			
	Day				Day				Day			
	3	5	8	15	3	5	8	15	3	5	8	15
Pulmonary	7.6	1.5	1.5	1.5	12.0	4.9	2.2	0.0	37.5	0.0	0.0	12.5
Infection	10.6	6.1	0.0	1.5	18.1	11.1	7.5	1.3	50.0	50.0	37.5	12.5
Renal	10.6	4.6	3.0	1.5	22.6	8.4	3.1	0.9	50.0	0.0	0.0	12.5
Gastrointestinal	9.1	3.0	1.5	0.0	16.8	8.4	2.7	0.4	0.0	0.0	0.0	12.5
Cardiovascular	0.0	1.5	1.5	0.0	3.5	4.4	4.0	0.9	12.5	12.5	12.5	0.0
Neurological	0.0	0.0	0.0	0.0	2.2	1.8	0.4	0.0	0.0	0.0	0.0	0.0
Wound	3.0	1.5	0.0	0.0	2.7	6.2	3.1	0.4	0.0	0.0	12.5	12.5
Haematological	0.0	1.5	0.0	0.0	2.7	0.4	0.4	0.0	0.0	0.0	0.0	0.0
Pain	4.6	1.5	1.5	0.0	7.1	0.0	0.0	0.0	25.0	0.0	0.0	0.0

2.3.2.1.1 Pulmonary morbidity

Pulmonary morbidity is defined as ‘a new need for oxygen or respiratory support’. No knee arthroplasty patient required respiratory support following surgery. 7.6% of UKR patients required supplementary oxygen on POD 3. This reduced to 1.5% (1/66) on POD 5, 8 and 15. This patient developed a

perforated gastric ulcer 2 days following knee surgery, which was repaired by the general surgical team. The patient needed supplementary oxygen for a total of 15 days following his abdominal surgery and remained an inpatient for a total of 22 days.

12.0% of TKR patients required supplementary oxygen on POD 3. This figure reduced to 4.9% by POD 5 and to 2.2% by POD 8. No TKR patients required oxygen on POD 15.

37.5% of RTKR patients required supplementary oxygen on POD 3. No RTKR patients required oxygen on POD 5 and 8. Supplementary oxygen was restarted for one patient on POD 13 following a second surgical procedure (knee washout and replacement of polyethylene liner) for an infected prosthesis.

2.3.2.1.2 Infectious morbidity

The POMS definition of infectious morbidity is a patient 'currently taking antibiotics' and/or has had 'a temperature of 38ºC or above in the last 24 hours'. On POD 3, 10.6% (7/66) of UKR patients had infectious morbidity. 3 of these patients had a temperature over 38ºC in the preceding 24 hours, 3 patients were started on antibiotics for a clinically suspected superficial surgical site infection and 1 patient was taking intravenous antibiotics following surgical repair of a perforated gastric ulcer.

On POD 5, the same 4 patients (6.1%) remained on antibiotics (3 for possible wound infection and 1 following rupture of gastric ulcer). By POD 8, no UKR patient had infectious morbidity. The patient who underwent repair of a perforated gastric ulcer was restarted on intravenous antibiotics on POD 11 and accounts for the 1 patient (1.5%) with infectious morbidity on POD 15.

On POD 3, 18.1% (41/226) of TKR patients had infectious morbidity. 14 patients had a temperature over 38ºC in the preceding 24 hours, 13 patients were started on antibiotics for possible superficial surgical site infection, 5 patients were prescribed antibiotics for a clinically diagnosed chest infection, 5 patients were taking antibiotics for urinary sepsis (based on clinical diagnosis and results from urine dipstick analysis) and 1 patient was started on antibiotics for a clinically diagnosed ear infection. 3 patients were taking antibiotics with no recorded reason.

By POD 5, the percentage of TKR patients with infectious morbidity had fallen to 11.1% (25/226). These patients all had infectious morbidity on POD 3. There were no 'new' patients with infection. On POD 8 the percentage of TKR patients with infectious morbidity fell to 7.5% (17/226). By POD 15, 1.3% (3/226) of TKR patients had infectious morbidity. There were no 'new' patients: 1 patient continued antibiotics for cellulitis surrounding the wound, 1 continued antibiotics for urinary sepsis and 1 for diarrhoea.

50.0% of RTKR patients had infectious morbidity on POD 3. This remained the same on POD 5 and reduced to 37.5% by POD 8 and to 12.5% by POD 15. These figures represent patients who had infected primary prostheses and were taking planned intravenous antibiotics following their revision surgery. Tissue samples were taken prior to surgery to ascertain organism sensitivity and plan an appropriate antibiotic regimen.

2.3.2.1.3 Renal morbidity

Renal morbidity is defined by the POMS as 'oliguria of less than 500ml/day, a raised creatinine level of over 30% compared to pre-operatively, or a urinary catheter in-situ'. On POD 3, 10.6% (7/66) of UKR patients had renal morbidity. 5 patients had a catheter in-situ and 2 had oliguria (although it was noted that both patients had incontinence so urine output may not have been recorded accurately). By POD 5, 4.6% (3/66) UKR patients had renal morbidity (all had a urinary catheters in-situ). By POD 8, 3.0% of patients had renal morbidity and this fell to 1.5% (1/66) by POD 15. This patient perforated a gastric ulcer following his knee surgery. His fluid balance was still being monitored on POD 15, necessitating a urinary catheter.

22.6% (51/226) of TKR patients had renal morbidity on POD 3. 1 patient had acute renal failure with a high creatinine level, 9 patients had low urine output and 41 patients had a catheter in-situ. The proportion of patients with renal morbidity fell to 8.4% on POD 5. 1 patient had on-going renal failure, 2 patients had oliguria and 16 had urinary catheters in-situ. The proportion of patients

with renal morbidity fell further to 3.1% (7/226) on POD 8. These patients had urinary catheters in-situ. The figure reduced further to 0.9% (2/226) by POD 15. One patient had slowly improving renal failure and another had a catheter in-situ.

50.0% of RTKR patients had renal morbidity on POD 3. These patients had a urinary catheter in-situ. No RTKR patients had renal morbidity on POD 5 and 8. By POD 15, 1 RTKR patient had developed 'new' renal morbidity. This patient developed generalised sepsis secondary to a deep knee infection with subsequent acute renal failure. This required urethral catheterisation to monitor fluid balance.

2.3.2.1.4 Gastro-intestinal morbidity

The POMS defines post-operative gastro-intestinal morbidity as 'unable to tolerate an enteral diet' or a patient experiencing 'nausea, vomiting or abdominal distension'. 9.1% of UKR patients had gastro-intestinal morbidity on POD 3. This reduced to 3.0% by POD 5 and to 1.5% by POD 8. No UKR patients had gastro-intestinal morbidity on POD 15.

16.8% of TKR patients had gastro-intestinal morbidity on POD 3. This reduced to 8.4% by POD 5, to 2.7% by POD 8 and to 0.4% (1/226) by POD 15. This patient had felt nauseous since surgery but was eating normally.

No RTKR patients had gastro-intestinal morbidity on POD 3, 5 or 8. 1 patient (12.5%) developed gastro-intestinal morbidity by POD 15. This patient had generalised sepsis secondary to an infected knee joint. He was vomiting regularly and was not able to tolerate an enteral diet.

2.3.2.1.5 Cardiovascular morbidity

The POMS defines cardiovascular morbidity as 'ischaemia or hypotension requiring drug therapy or fluid therapy of over 200ml per hour, atrial or ventricular arrhythmia, cardiogenic pulmonary oedema or new anticoagulation'. No UKR patient had cardiovascular morbidity on POD 3. By POD 5, 1 patient (1.5%) had developed cardiovascular morbidity: atrial fibrillation following perforation of a gastric ulcer. The same patient was still being treated for the arrhythmia on POD 8. By POD 15 no UKR patient had cardiovascular morbidity.

3.5% (8/226) of TKR patients had cardiovascular morbidity on POD 3. 6 patients were restarted on warfarin they had been taking pre-operatively. 2 patients were prescribed anticoagulation to treat DVTs. By POD 5, 4.4% (10/226) of TKR patients had cardiovascular morbidity. 4 of these patients were 'new'. 3 patients were prescribed anticoagulation for deep vein thrombosis and 1 patient had a myocardial infarction. The proportion of TKR patients with cardiovascular morbidity fell to 4.0% by POD 8 and fell further to 0.9% by POD 15.

No RTKR patients had cardiovascular morbidity on POD 3 and 5. By POD 8, 1 patient had developed cardiovascular morbidity. This patient restarted warfarin he had been taking preoperatively. He continued to take warfarin on POD 15.

2.3.2.1.6 Neurological morbidity

The POMS defines neurological morbidity as a 'new confusion/delirium, focal deficit or coma'. No UKR or RTKR patients developed neurological morbidity in the post-operative period. 2.2% (5/226) of TKR patients had neurological morbidity on POD 3. 3 patients were confused, 1 patient had a cerebro-vascular accident (CVA) and 1 patient developed acute psychosis. The proportion of TKR patients with neurological morbidity reduced to 1.8% by POD 5 and to 0.4% by POD 8. By POD 15 no TKR patient had neurological morbidity.

2.3.2.1.7 Wound morbidity

The POMS defines wound morbidity as a 'wound dehiscence requiring surgical exploration or drainage of pus from the operative wound with or without isolation of organisms'. 3.0% of UKR patients had wound morbidity on POD 3. This reduced to 1.5% by POD 5. No UKR patient had wound morbidity on POD 8 and 15.

2.7% of TKR patients had wound morbidity on POD 3. This rose to 6.2% by POD 5. This included 8 'new' patients who had not been diagnosed with wound morbidity on POD 3. By POD 8, 3.1% of TKR patients had wound morbidity and this figure fell to 0.4% by POD 15.

No RTKR patients had wound morbidity on POD 3 and 5. By POD 8, 1 patient had developed wound morbidity. The indication for revision surgery in this patient was infection of the primary prosthesis. The patient developed recurrent infection in the revision replacement necessitating further surgery. His wound morbidity remained on POD 15.

2.3.2.1.8 Haematological morbidity

The POMS defines haematological morbidity as a patient requiring 'red blood cells, platelets, fresh frozen plasma or cryoprecipitate within the last 24 hours'. On POD 3 no UKR patients had haematological morbidity. 1 patient (1.5%) had haematological morbidity on POD 5. This patient perforated a gastric ulcer and required a red cell transfusion following his abdominal surgery. No UKR patient had haematological morbidity on POD 8 and 15.

2.7% of TKR patients had haematological morbidity on POD 3. This reduced to 0.4% on POD 5 and 8. By POD 15, no TKR patient had haematological morbidity. All haematological morbidity referred to red cell transfusion. No knee arthroplasty patient required post-operative platelets, fresh frozen plasma or cryoprecipitate.

No RTKR patient had haematological morbidity on POD 3, 5, 8 or 15.

2.3.2.1.9 Pain morbidity

The POMS defines pain morbidity as ‘surgical wound pain significant enough to require parenteral opioids or regional anaesthesia’. 4.6% of UKR patients had pain morbidity on POD 3. This reduced to 1.5% by POD 5 and 8. By POD 15, no UKR patient had pain morbidity.

7.1% of TKR patients had pain morbidity on POD 3. No TKR patients had pain morbidity on POD 5, 8 or 15. 25.0% of RTKR patients had pain morbidity on POD 3. No RTKR patients had pain morbidity on POD 5, 8 or 15.

2.3.2.2 Hip arthroplasty patients

Table 11 shows a summary of morbidity following hip arthroplasty according to the POMS.

Table 11. Post-operative morbidity (defined by the POMS) in 229 elective hip arthroplasty patients on days 3, 5, 8 and 15. Figures are a percentage of the total number of patients in each arthroplasty group.

	Hip resurfacing (HR)				Total hip replacement (THR)				Revision total hip replacement (RTHR)			
	Day				Day				Day			
	3	5	8	15	3	5	8	15	3	5	8	15
Pulmonary	3.1	0.0	0.0	0.0	15.4	3.7	3.7	0.0	25.7	14.3	5.7	11.4
Infection	6.3	6.3	0.0	0.0	20.4	13.6	5.6	4.3	71.4	57.1	37.1	25.7
Renal	15.6	0.0	0.0	0.0	35.8	14.2	4.3	1.2	77.1	25.7	25.7	17.1
Gastrointestinal	6.3	3.1	0.0	0.0	11.7	6.8	2.5	1.2	17.1	17.1	8.6	11.4
Cardiovascular	0.0	0.0	0.0	0.0	4.3	4.9	4.3	2.5	3.0	5.7	11.4	5.7
Neurological	0.0	0.0	0.0	0.0	1.2	0.6	1.2	0.0	5.7	5.7	2.9	2.9
Wound	0.0	6.3	0.0	0.0	4.3	9.3	3.7	1.9	0.0	17.1	8.6	5.7
Haematological	6.3	0.0	0.0	0.0	4.3	2.5	1.2	0.0	11.4	8.6	8.6	2.9
Pain	3.1	3.1	0.0	0.0	8.6	1.2	0.0	0.0	28.6	5.7	2.9	5.7

2.3.2.2.1 Pulmonary morbidity

No hip arthroplasty patient required respiratory support following surgery. All figures represent patients requiring supplementary oxygen. 3.1% of HR patients required oxygen on POD 3. No HR patient had pulmonary morbidity on POD 5, 8 or 15.

15.4% of THR patients required supplementary oxygen on POD 3. This reduced to 3.7% on POD 5 and 8. By POD 15 no THR patient had pulmonary morbidity.

25.7% of RTHR patients had pulmonary morbidity on POD 3. This reduced to 14.3% by POD 5 and to 5.7% by POD 8. By POD 15 the figure increased to 11.4% (4/35). 1 patient had been on supplementary oxygen since surgery and 3 patients were 'new'. The 'new' patients had returned to theatre for further surgery due to infection. Supplementary oxygen was started following their second operative intervention.

2.3.2.2.2 Infectious morbidity

On POD 3, 6.3% (2/32) of HR patients had infectious morbidity. 1 patient had a temperature of over 38ºC in the preceding 24 hours and 1 patient was prescribed antibiotics for a clinically suspected superficial surgical site infection. On POD 5 the same 2 patients had infectious morbidity. No HR patient had infectious morbidity on POD 8 and 15.

On POD 3, 20.4% (33/162) of THR patients had infectious morbidity. 8 patients had a temperature exceeding 38ºC in the preceding 24 hours. The remaining 25 patients were prescribed antibiotics: 11 for a clinically suspected surgical site infection, 6 for clinically diagnosed chest infection, 4 for urinary tract infection, 1 for endocarditis and 1 for an infected intravenous catheter site. The reason for prescribing antibiotics was not recorded for 2 patients. By POD 5 the

percentage of THR patients with infectious morbidity had fallen to 13.6% (22/162). These patients all had infectious morbidity on POD 3. There were no 'new' patients. On POD 8 the number of THR patients with an infection had fallen to 5.6% (9 patients). There were no 'new' patients. By POD 15, 4.3% (7/162) of THR patients had infectious morbidity. 1 patient had a 'new' diagnosis of an infected intra-venous cannula site.

71.4% of RTHR patients had infectious morbidity on POD 3. This reduced to 57.1% by POD 5 and reduced further to 37.1% by POD 8. By POD 15, 25.7% of RTHR patients had infectious morbidity. These figures represent patients with known infected primary prostheses. The reason for revision hip surgery was to eradicate infection. The administration of appropriate antibiotics was planned pre-operatively.

2.3.2.2.3 Renal morbidity

On POD 3, 15.6% (5/32) of HR patients had renal morbidity. These patients had a urinary catheter in-situ. No HR patient had renal morbidity on POD 5, 8 and 15.

35.8% (58/162) of THR patients had renal morbidity on POD 3. 2 of these patients had a low urine output. The remaining patients had a urinary catheter in-situ. The percentage of THR patients with renal morbidity on POD 5 reduced to 14.2% (all had a urinary catheter in-situ) and reduced further to 4.3% by POD

8. By POD 15, 1.2% (2/162) of THR patients had renal morbidity. Both these patients had failed a trial without catheter.

77.1% of RTHR patients had renal morbidity on POD 3 (all patients had a urinary catheter in-situ). This figure reduced to 25.7% by POD 5 and 8 and reduced again to 17.1% by POD 15.

2.3.2.2.4 Gastro-intestinal morbidity

6.3% of HR patients had gastro-intestinal morbidity on POD 3. This reduced to 3.1% by POD 5. No HR patients had gastro-intestinal morbidity on POD 8 and 15.

11.7% of THR patients had gastro-intestinal morbidity on POD 3. This reduced to 6.8% by POD 5 and to 2.5% by POD 8. 1.2% (2/162) of THR patients had gastro-intestinal morbidity on POD 15. Both these patients felt nauseous but were tolerating an enteral diet.

17.1% (6/35) of RTHR patients had gastro-intestinal morbidity on POD 3. 4 patients were nauseous and unable to tolerate an enteral diet, 1 patient had haematemesis (endoscopy showed oesophagitis secondary to reflux) and 1 patient had melaena requiring a 4-unit red cell transfusion (colonoscopy, oesophago-gastro-duodenoscopy and abdominal imaging did not reveal the source of bleeding). The proportion of RTHR patients with gastro-intestinal morbidity remained at 17.1% on POD 5. By POD 8 the proportion of RTHR

patients with gastro-intestinal morbidity had reduced to 8.6%. The figure increased to 11.4% on POD 15. This increase represented one ‘new’ patient who had further surgery on POD 15 for an infected prosthesis. The patient was last on the operating list and did not eat all day to maintain an empty stomach for surgery.

2.3.2.2.5 Cardiovascular morbidity

No HR patient had cardiovascular morbidity on POD 3, 5, 8 or 15. 4.3% (7/162) of THR patients had cardiovascular morbidity on POD 3. 2 patients were re-started on warfarin they had been taking pre-operatively, 1 patient was prescribed warfarin for a deep vein thrombosis, 2 patients had myocardial infarctions (1 of these patients also had bilateral pulmonary emboli) and 2 patients developed supraventricular tachycardia. On POD 5 the proportion of THR patients with cardiovascular morbidity rose to 4.9%. There was one ‘new’ patient who was restarted on warfarin for atrial fibrillation. The patient had been taking warfarin pre-operatively. The proportion of THR patients with cardiovascular morbidity fell to 4.3% by POD 8 and fell further to 2.5% by POD 15.

3.0% (1/35) of RTHR patients had cardiovascular morbidity on POD 3. This patient had a supraventricular tachycardia. 5.7% (2/35) of RTHR patients had cardiovascular morbidity by POD 5. This included one ‘new’ patient who was prescribed anticoagulation for a pulmonary embolus. By POD 8 the proportion of RTHR with cardiovascular morbidity increased further to 11.4% (4/35).

The increase represented 2 ‘new’ patients: one was re-started on warfarin (which he had been taking pre-operatively for atrial fibrillation) and the other patient was prescribed warfarin for a deep vein thrombosis. By POD 15 the proportion of RTHR patients with cardiovascular morbidity had fallen to 5.7% (2/35). One of these patients had developed a pulmonary embolus by POD 5 and proceeded to develop endocarditis by day POD 15. This patient was transferred to the Heart Hospital for aortic valve replacement.

2.3.2.2.6 Neurological morbidity

No HR patient had neurological morbidity on POD 3, 5, 8 or 15. 1.2% (2/162) of THR patients had neurological morbidity on POD 3. Both patients were confused. The proportion of patients with neurological morbidity reduced to 0.6% by POD 5 and increased to 1.2% by POD 8. This increase represented one ‘new’ patient who developed right-sided weakness. Brain imaging revealed a small region of ischaemia. Symptoms resolved within a few days confirming a transient ischaemic attack (TIA). By POD 15 no THR patients had neurological morbidity.

On POD 3, 5.7% (2/35) of RTHR patients had neurological morbidity. 1 patient had acute confusion and 1 patient had right arm weakness. Brain imaging confirmed an ischaemic stroke. The figure remained the same on POD 5 and reduced to 2.9% by POD 8 and 15.

2.3.2.2.7 Wound morbidity

No HR patient had wound morbidity on POD 3. This increased to 6.3% (2/32) by POD 5. Both patients had purulent discharge from the surgical site. No HR patient had wound morbidity on POD 8 and 15.

4.3% of THR patients had wound morbidity on POD 3. This rose to 9.3% by POD 5. This represented 8 'new' patients who did not have wound morbidity on POD 3. By POD 8 the proportion of THR patients with wound morbidity had fallen to 3.7%. The proportion fell further to 1.9% by POD 15.

No RTHR patients had wound morbidity on POD 3. By POD 5, 17.1% (6/35) of RTHR patients had developed wound morbidity. The indication for revision surgery in these patients was infection of the primary prosthesis. They all developed a surgical site infection following revision surgery. The proportion of RTHR patients with wound morbidity reduced to 8.6% by POD 8 and to 5.7% by POD 15.

2.3.2.2.8 Haematological morbidity

On POD 3, 6.3% of HR patients had haematological morbidity. No HR patients had haematological morbidity on POD 5, 8 and 15. All haematological morbidity following hip arthroplasty represented red cell transfusion. No hip arthroplasty patient required platelets, fresh frozen plasma or cryoprecipitate.

4.3% of THR patients had haematological morbidity on POD 3. This reduced to 2.5% on POD 5 and to 1.2% on POD 8. By POD 15 no THR patient had haematological morbidity.

11.4% of RTHR patients had haematological morbidity on POD 3. This reduced to 8.6% on POD 5 and 8, and reduced further to 2.9% by POD 15.

2.3.2.2.9 Pain morbidity

3.1% of HR patients had pain morbidity on POD 3. This figure remained the same on POD 5. No HR patients had pain morbidity on POD 8 and 15. 8.6% of THR patients had pain morbidity on POD 3. This reduced to 1.2% by POD 5. No THR patients had pain morbidity on POD 8 and 15. 28.6% of RTHR patients had pain morbidity on POD 3. This reduced to 5.7% by POD 5, and further to 2.9% by POD 8. This figure increased to 5.7% by POD 15. This represented 1 'new' patient who had undergone further surgery on POD 13 for a periprosthetic fracture. It was 2 days after his second surgical procedure.

2.3.3 Comparison of post-operative morbidity between different types of arthroplasty

2.3.3.1 Pulmonary morbidity

Pulmonary morbidity following different types of hip and knee arthroplasty was

compared using Fisher's exact tests. Table 12 gives the Bonferroni adjusted p-values resulting from these tests.

Table 12. Bonferroni adjusted p-values from Fisher's exact tests comparing pulmonary morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	0.41	0.51	0.88	1.00	1.00
POD 5	1.00	1.00	1.00	0.17	1.00	1.00
POD 8	1.00	1.00	1.00	1.00	1.00	1.00
POD 15	1.00	0.21	1.00	**0.005**	1.00	1.00

The results suggest no statistically significant difference in pulmonary morbidity between operation types for POD 3, 5 and 8.

On POD 15 there is a significant difference in morbidity between the THR and RTHR groups. RTHR has a higher morbidity, with 11% of patients having pulmonary morbidity compared to no THR patients.

2.3.3.2 Infection morbidity

Infection following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 13 gives the Bonferroni adjusted p-values resulting from these tests.

Table 13. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative infection between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	0.28	0.46	<**0.001**	1.00	1.00
POD 5	1.00	0.06	1.00	<**0.001**	1.00	1.00
POD 8	0.10	0.14	1.00	<**0.001**	1.00	1.00
POD 15	1.00	0.78	1.00	**0.002**	0.61	1.00

The results suggest no significant difference between the majority of operations compared. The exception is between THR and RTHR, where a significant difference is observed on POD 3, 5, 8 and 15. The results suggest that on each of these days, infection is higher for RTHR than for THR. For example, on POD 3, 71% of RTHR patients had an infection, compared to 20% of THR patients.

Additionally, there was slight evidence that RTKR had a higher infection rate than TKR on POD 5, but this result was not quite statistically significant.

2.3.3.3 Renal morbidity

Renal morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 14 gives the Bonferroni adjusted p-values from these tests.

Table 14. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative renal morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	0.21	0.54	0.22	<**0.001**	**0.04**	1.00
POD 5	1.00	1.00	0.10	0.76	0.58	1.00
POD 8	1.00	1.00	1.00	**0.002**	1.00	1.00
POD 15	1.00	0.60	1.00	**0.003**	1.00	1.00

Statistically significant differences for renal morbidity between THR and RTHR were observed on POD 3, 8 and 15, although no difference was observed on POD 5. Where there was a difference, renal morbidity was higher for RTHR than for THR. For example, on POD 8, 26% of RTHR patients had renal morbidity but only 4% of THR patients.

A significant difference in renal morbidity between TKR and THR was observed on POD 3, with a higher occurrence of morbidity for THR than for TKR. No significant difference for renal morbidity between these two operation types was observed on subsequent days.

2.3.3.4 Gastrointestinal morbidity

Gastrointestinal morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 15 gives the Bonferroni adjusted p-values from these tests.

Table 15. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative gastrointestinal morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	1.00	1.00	1.00	1.00	1.00
POD 5	1.00	1.00	1.00	0.53	1.00	1.00
POD 8	1.00	1.00	1.00	0.65	1.00	1.00
POD 15	1.00	0.40	1.00	0.06	1.00	1.00

There was difference in gastrointestinal morbidity between any of the operations compared. There was slight evidence that RTHR patients have a higher occurrence of gastrointestinal morbidity than THR patients on POD 15 but this result was not quite statistically significant.

2.3.3.5 Cardiovascular morbidity

Cardiovascular morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 16 gives the Bonferroni

adjusted p-values from these tests.

Table 16. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative cardiovascular morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	1.00	1.00	1.00	1.00	1.00
POD 5	1.00	1.00	1.00	1.00	1.00	1.00
POD 8	1.00	1.00	1.00	0.66	1.00	1.00
POD 15	1.00	0.60	1.00	1.00	1.00	1.00

No statistically significant difference in post-operative cardiovascular morbidity was observed between operation types at any of the four time points.

2.3.3.6 Neurological morbidity

Neurological morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 17 gives the Bonferroni adjusted p-values from these tests.

Table 17. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative neurological morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	1.00	1.00	0.87	1.00	1.00
POD 5	1.00	1.00	1.00	0.49	1.00	1.00
POD 8	1.00	1.00	1.00	1.00	1.00	1.00
POD 15	1.00	1.00	1.00	1.00	1.00	1.00

The results suggest no significant difference in post-operative neurological morbidity between operation types.

2.3.3.7 Wound morbidity

Wound morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 18 gives the Bonferroni adjusted p-values from these tests.

Table 18. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative wound morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	1.00	1.00	1.00	1.00	1.00
POD 5	1.00	1.00	1.00	1.00	1.00	1.00
POD 8	1.00	1.00	1.00	1.00	1.00	1.00
POD 15	1.00	0.40	1.00	1.00	1.00	1.00

There was no statistically significant difference in post-operative wound morbidity between any of the observed groups.

2.3.3.8 Haematological morbidity

Haematological morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 19 gives the Bonferroni adjusted p-values from these tests.

Table 19. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative haematological morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	1.00	1.00	0.66	1.00	1.00
POD 5	1.00	1.00	1.00	0.65	0.99	1.00
POD 8	1.00	1.00	1.00	0.24	1.00	1.00
POD 15	1.00	1.00	1.00	1.00	1.00	1.00

There was no statistically significant difference in post-operative haematological morbidity between any of the observed groups.

2.3.3.9 Pain morbidity

Pain morbidity following different types of hip and knee arthroplasty was compared using Fisher's exact tests. Table 20 gives the Bonferroni adjusted p-values from these tests.

Table 20. Bonferroni adjusted p-values from Fisher's exact tests comparing post-operative pain morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	1.00	0.71	1.00	**0.02**	1.00	1.00
POD 5	1.00	1.00	1.00	0.87	1.00	1.00
POD 8	1.00	1.00	1.00	1.00	1.00	1.00
POD 15	1.00	1.00	1.00	0.18	1.00	1.00

The results suggest a significant difference in pain morbidity between the THR and RTHR groups on POD 3. Pain was present in 29% of RTHR patients at this time point but in only 9% of THR patients. No other differences in post-operative pain morbidity between operations were observed.

2.3.3.10 The presence of any morbidity

The proportion of patients with the presence of any type of morbidity as defined by the POMS following different types of hip and knee arthroplasty is shown in Table 21.

Table 21. The proportion of patients with any type of morbidity on post-operative days 3, 5, 8 and 15. Figures are a percentage of the total number of patients in each arthroplasty group.

Time	UKR (n=66)	TKR (n=226)	RTKR (n=8)	HR (n=32)	THR (n=162)	RTHR (n=35)
POD 3	26.7%	49.1%	75.0%	28.1%	54.9%	91.4%
POD 5	10.5%	27.0%	50.0%	15.6%	37.0%	71.4%
POD 8	4.5%	16.4%	62.5%	0.0%	17.3%	57.1%
POD 15	1.5%	3.1%	12.5%	0.0%	8.0%	40.0%

The presence of morbidity in each arthroplasty group was compared using Fisher's exact tests. Table 22 gives the Bonferroni adjusted p-values from these tests.

Table 22. Bonferroni adjusted p-values from Fisher's exact tests comparing the presence of any post-operative morbidity between different types of arthroplasty

Time	UKR vs TKR	TKR vs RTKR	HR vs THR	THR vs RTHR	TKR vs THR	RTKR vs RTHR
POD 3	**0.008**	1.00	0.04	<**0.001**	1.00	1.00
POD 5	**0.03**	1.00	0.14	**0.002**	0.27	1.00
POD 8	0.08	**0.03**	**0.03**	<**0.001**	1.00	1.00
POD 15	1.00	1.00	0.79	<**0.005**	0.22	1.00

The results suggest significant increased morbidity following RTHR compared to THR at all four time points following surgery. For example, at POD 3, over 90% of RTHR patients had morbidity, compared to just over half of THR patients.

There was a difference between UKR and TKR on POD 3 and 5, and also a slight evidence of a difference on POD 8. Morbidity was higher for TKR than for UKR at all these time points.

There were further differences in morbidity on POD 8 between TKR and RTKR, and also between HR and THR. On this day RTKR had a higher morbidity than TKR, whilst THR had a higher rate of morbidity than HR. Differences between

these operation types were not statistically significant on POD 3, 5 and 15.

2.3.4 Comparison of post-operative morbidity between different types of arthroplasty with POSSUM risk adjustment

2.3.4.1 POSSUM morbidity scores in the arthroplasty groups

A possible confounding variable in a comparison of operation types is the difference in risk of morbidity between the operation groups. An analysis was performed to compare the POSSUM morbidity scores between the 6 types of arthroplasty. The POSSUM values were found to have a positively skewed distribution so the Kruskal-Wallis test was used for the analysis. The results are summarised in Table 23. The median and inter-quartile ranges are reported for each operation type, together with a p-value indicating the overall difference between the six operations.

Table 23. POSSUM morbidity scores in the six arthroplasty groups

Operation	POSSUM morbidity score Median (IQR)	P-value
UKR	11.7 (8.8, 14.2)	**<0.001**
TKR	18.5 (16.2, 21.1)	
RTKR	22.5 (16.4, 31.9)	
HR	11.4 (9.6, 14.8)	
THR	18.5 (14.2, 26.9)	
RTHR	31.2 (18.5, 42.3)	

The results suggest a highly significant difference in POSSUM morbidity scores between the arthroplasty groups. Scores were lowest for the UKR and HR groups, with a median value of around 11 in both. The highest scores were for RTHR, which had a median score of 31.

2.3.4.2 Comparison of post-operative morbidity between different types of arthroplasty with adjustment for POSSUM scores

2.3.4.1.1 THR vs RTHR

The initial analyses comparing morbidity between the arthroplasty groups with no adjustment suggested some differences in morbidity between THR and RTHR, with morbidity higher for RTHR when a difference was found. The POSSUM scores shown in Table 23 show that the risk for morbidity was

higher in the RTHR groups than the THR group. Therefore, any differences in morbidity could be attributable to a difference in risk score rather than a difference between operation types.

For this reason further regression analyses were performed specifically to re-visit the significant differences between operations, adjusting for differences in POSSUM score between operation types. The majority of differences observed were between THR and RTHR. The results of the unadjusted and POSSUM-adjusted differences between these two hip procedures are summarised in Table 24. The difference in morbidity between operations is reported as an odds ratio, together with corresponding confidence intervals. The odds ratio is the odds of morbidity for RTHR relative to the odds for THR. Bonferroni adjusted p-values indicating the significance of the results are also reported.

Table 24. Unadjusted and POSSUM-adjusted differences between post-operative morbidity following THR and RTHR

Morbidity	Time	Unadjusted		Adjusted	
		OR (95% CI) (*)	P-value	OR (95% CI) (*)	P-value
Pulmonary	POD 15	(†)		(†)	
Infection	POD 3	9.77 (4.27, 22.3)	**<0.001**	8.18 (3.45, 19.4)	**<0.001**
	POD 5	8.48 (3.79, 19.0)	**<0.001**	8.92 (3.71, 21.5)	**<0.001**
	POD 8	10.0 (3.85, 26.2)	**<0.001**	9.82 (3.50, 27.6)	**<0.001**
	POD 15	7.66 (2.63, 22.4)	**<0.001**	4.92 (1.52, 16.0)	**0.05**
Renal	POD 3	6.05 (2.58, 14.2)	**<0.001**	4.59 (1.90, 11.1)	**0.006**
	POD 8	7.66 (2.63, 22.4)	**<0.001**	3.77 (1.12, 12.7)	0.19
	POD 15	16.6 (3.18, 86.1)	**0.006**	9.71 (1.63, 57.7)	0.07
Pain	POD 3	4.23 (1.69, 10.6)	**0.01**	2.31 (0.82, 6.53)	0.69
Any	POD 3	8.75 (2.57, 29.7)	**0.006**	6.10 (1.74, 21.4)	**0.03**
	POD 5	4.25 (1.91, 9.46)	**<0.001**	3.38 (1.47, 7.81)	**0.02**
	POD 8	6.38 (2.91, 14.0)	**<0.001**	4.14 (1.77, 9.68)	**0.006**
	POD 15	7.64 (3.16, 18.5)	**<0.001**	4.45 (1.68, 11.8)	**0.02**

(*) Odds Ratios given as odds of morbidity for RTHR relative to THR
(†) No occurrences of morbidity for THR, thus logistic regression unable to be performed

The initial analyses found a difference in the presence of pulmonary morbidity between THR and RTHR on POD 15. There were no occurrences of morbidity for THR so logistic regression could not be performed in this instance.

The results suggest that differences in post-operative infection between THR and RTHR remain after adjusting for the POSSUM morbidity score. The result at POD 15 is of borderline statistical significance.

The difference in renal morbidity following RTHR compared to THR remained on POD 3 after adjusting for the POSSUM risk score. Renal morbidity was still higher for RTHR than THR at POD 8 and 15 but the difference was no longer statistically significant.

There was no longer a statistically significant difference in pain at POD 3 between the two operations after adjusting for the POSSUM score.

A difference in the presence of any type of morbidity following RTHR as compared to THR remained on POD 3, 5, 8 and 15 after adjusting for the POSSUM score. The size of difference in morbidity between operations was slightly reduced after adjustment for POSSUM.

2.3.4.1.2 THR vs TKR

Logistic regression was also used to compare TKR and THR where a previous difference was observed. The results are summarised in Table 25.

Table 25. Unadjusted and POSSUM-adjusted differences between post-operative morbidity for THR and TKR

Morbidity	Time	Unadjusted		Adjusted	
		OR (95% CI) (*)	P-value	OR (95% CI) (*)	P-value
Renal	POD 3	1.91 (1.22, 2.99)	**0.02**	1.70 (1.08, 2.70)	0.14

(*) Odds Ratios given as odds of morbidity for THR relative to TKR

The results suggest that the difference in renal morbidity is still raised for THR relative to TKR after adjusting for the POSUM score but this difference is no longer of statistical significance.

2.3.4.1.3 TKR vs UKR

The next comparison was made between the TKR and UKR groups where a previous difference was observed on POD 3 and 5. The results are summarised in Table 26.

Table 26. Unadjusted and POSSUM-adjusted differences between post-operative morbidity for TKR and UKR

Morbidity	Time	Unadjusted		Adjusted	
		OR (95% CI) (*)	P-value	OR (95% CI) (*)	P-value
Any	POD 3	2.63 (1.44, 4.79)	**0.01**	1.91 (1.00, 3.62)	0.29
	POD 5	3.17 (1.37, 7.31)	**0.04**	2.18 (0.91, 5.22)	0.48

(*) Odds Ratios given as odds of morbidity for TKR relative to UKR

The results suggest that after adjusting for the POSSUM score, there was no longer a statistically significant difference between the two types of knee arthroplasty on either POD 3 or 5.

2.3.4.1.4 RTKR vs TKR

The next comparison was made between the RTKR and TKR groups where a previous difference was observed on POD 8. The results are summarised in Table 27.

Table 27. Unadjusted and POSSUM-adjusted differences between post-operative morbidity for RTKR and TKR

Morbidity	Time	Unadjusted		Adjusted	
		OR (95% CI) (*)	P-value	OR (95% CI) (*)	P-value
Any	POD 8	8.51 (1.94, 37.2)	**0.02**	6.40 (1.31, 31.2)	0.13

(*) Odds Ratios given as odds of morbidity for RTKR relative to TKR

The results suggest that, after adjustment, the difference between the two types of knee arthroplasty for the presence of any morbidity at POD 8 is no longer statistically significant.

The initial analyses also found a difference in the presence of any morbidity between HR and THR on post-operative day eight. There were no occurrences of morbidity for HR, and thus logistic regression could not be performed in this instance.

2.4 Discussion

2.4.1 Summary of post-operative morbidity following lower limb arthroplasty and suggestions for improvement

This is the first time the POMS has been used to record and compare morbidity following different types of lower limb arthroplasty. Overall, the most prevalent types of morbidity are ‘infection’ and ‘renal’. There are lower levels of

‘pulmonary’, ‘pain’ and ‘gastro-intestinal’ morbidity. There is little ‘cardiovascular’, ‘neurological’, ‘wound’ or ‘haematological’ morbidity. The study shows that many patients remain in hospital with no morbidity as defined by the POMS.

Medical staff treating arthroplasty patients should be aware of the prevalence of post-operative morbidity in all organ systems. This will allow prompt diagnosis and treatment of morbidity, minimising its impact. Knowing the prevalence of morbidity following arthroplasty allows appropriate pre-operative patient education and counselling. This enables patients to have realistic expectations regarding the post-operative period.

To reduce post-operative morbidity, the prevalence and type of morbidity must first be known. This will allow targets to be set for the future. Morbidity with the highest prevalence or greatest long-term impact can be targeted first. The following paragraphs summarise the causes of each category of morbidity and suggest ways to reduce morbidity rates.

All pulmonary morbidity in this study referred to the use of supplementary oxygen. The administration of supplementary oxygen could be audited to assess if it was required in all cases. If inappropriate use was found, guidelines for correct prescribing could be produced and the audit cycle repeated. In this way, pulmonary morbidity due to the inappropriate use of oxygen could be reduced.

Morbidity due to infection included wound infections, chest infections and urinary tract infections. Surgical wounds can become infected either in theatre or on the ward. Many measures are already in place to keep surgical site infection rates to a minimum but further methods could be trialled such as the use of antibiotic-coated sutures in theatre and the use of 'ring-fenced' beds on the ward to ensure that arthroplasty patients do not come into contact with infected patients. To reduce the rate of chest infections, greater input from chest physiotherapists could be trialled. To reduce the rate of urinary tract infections, earlier removal of catheters could be trialled.

Renal morbidity was mainly due to the presence of a urinary catheter. Some catheters may remain in-situ longer than necessary. A trial of earlier removal of urinary catheters could be conducted in an effort to reduce renal morbidity and speed up patient discharge. The incidence of urinary retention would need to be closely monitored. If the incidence increased, the trial would need to be modified or abandoned.

Most gastro-intestinal morbidity was due to nausea or the patient unable to tolerate an enteral diet. Wider use of anti-emetics could reduce these problems. Routine prescription of anti-emetics in the first 3 days following surgery could be trialled in an attempt to reduce gastro-intestinal morbidity. Adverse effects of the anti-emetic medication would need to be monitored. If adverse effects outweighed the benefits of a reduction in nausea, the trial should be abandoned.

Cardiovascular morbidity was mainly due to the re-introduction of warfarin in patients who were taking anticoagulation pre-operatively. Restarting anticoagulation can be done as an outpatient. It may be prudent to remove this criterion from the definition of cardiovascular morbidity since it is not an adequate reason for a patient to remain in hospital. Even patients who require subcutaneous injections of anticoagulant until the correct serum warfarin level is achieved can be treated on an outpatient basis.

Another cause of cardiovascular morbidity was the diagnosis of deep vein thrombosis (DVT) and pulmonary embolism (PE). The rate of DVT was 0% following UKR, RTKR and HR. Rate of DVT following TKR was 2.2% (5/226), following THR was 1.2% (2/162) and following RTHR was 5.7% (2/35). Rate of PE was 0% following UKR, TKR, RTKR and HR. Rate of PE following THR was 0.6% (1/162) and following RTHR was 2.9% (1/35). These rates are comparable or better than those published in the literature[122]. To improve rates further, earlier post-operative physiotherapy could be introduced. Further audit would be needed to assess the impact.

Another cause of cardiovascular morbidity was myocardial infarction (MI). The rate of MI following UKR, RTKR, HR and RTHR was 0%. The rate of MI following TKR was 0.4% (1/226) and following THR was 1.2% (2/162). A further cause of cardiovascular morbidity was development of a new cardiac arrhythmia. No TKR, RTKR or HR patient developed a new arrhythmia.

1.5% (1/66) of UKR patients developed a cardiac arrhythmia, 0.8% (2/226) of THR patients and 2.9% (1/35) of RTHR patients. The rate of MI and new onset cardiac arrhythmia is low in all groups. The rate should be audited regularly to ensure it does not increase.

Neurological morbidity was very infrequent. 1 TKR patient (0.4%) and 1 THR patient (0.6%) suffered a CVA post-operatively. 1.3% (3/226) of TKR patients and 1.2% (2/162) of THR patients had acute post-operative confusion. 1 patient (0.4%) developed acute psychosis following TKR. These rates should be audited regularly to ensure they remain low.

Wound morbidity is mainly due to superficial surgical site infection. As already discussed, methods to reduce surgical site infection rate such as the use of 'ring-fenced' beds, minimal dressing changes and alternate wound closure methods could be trialled to see if any improvement could be made.

Haematological morbidity was entirely due to the requirement for red cell transfusion. Attempts could be made to reduce this by the more judicious use of a 'cell saver' during surgery, by ensuring good haemostasis prior to wound closure and by trying different pressure dressings to reduce post-operative bleeding.

Pain morbidity could be reduced by using larger volumes of local anaesthetic at the time of surgery and ensuring that all patients receive regular analgesia in

the early post-operative period. Patients could be discharged with a short course of low-dose opioids. The on-going use of opioids could be removed from the definition of pain morbidity since it is not an adequate reason for a patient to remain in hospital.

2.4.2 Pattern of POMS morbidity over time

In general, with increasing time after surgery, there was a decrease in the proportion of patients with each type of morbidity. However, there were exceptions to this trend as described below.

- Pulmonary morbidity increased between POD 8 and 15 in both the RTKR and RTHR groups. This increase represented 4 patients who required supplementary oxygen following second surgical procedures to treat infected prostheses.
- Infectious morbidity increased in the UKR group between POD 8 and 15 due to 1 patient restarting antibiotics following gastric ulcer repair. Infectious morbidity increased in the THR group between POD 8 and 15 due to one patient developing an infection at the site of a peripheral intra-venous cannula.
- Renal morbidity increased in the RTKR group between POD 8 and 15. This was due to 1 patient with generalised sepsis secondary to an infected prosthesis who had a urinary catheter inserted to monitor fluid balance.

- Gastro-intestinal morbidity increased between POD 8 and 15 in the RTKR group due to 1 patient vomiting secondary to generalised sepsis.
- Cardiovascular morbidity increased in the UKR group between POD 3 and 5 due to a patient developing atrial fibrillation (AF) following a perforated gastric ulcer. Cardiovascular morbidity increased in the TKR group between POD 3 and 5 due to 3 patients diagnosed with DVT and 1 patient diagnosed with MI. Cardiovascular morbidity increased in the RTHR group between POD 3 and 5, and again between POD 5 and 8. These increases were due to 1 patient in each time interval being diagnosed with a PE. Patients restarted on anticoagulation they were taking pre-operatively have not been included here.
- Neurological morbidity increased in the THR group between POD 5 and 8 due 1 patient having a CVA.
- Wound morbidity increased in all groups except UKR. 8 TKR patients were diagnosed with wound morbidity between POD 3 and 5. 1 RTKR patient was diagnosed with wound morbidity between POD 5 and 8. 2 HR patients were diagnosed with wound morbidity between POD 3 and 5. 8 THR patients were diagnosed with wound morbidity between POD 3 and 5. 6 RTHR patients were diagnosed with wound morbidity between POD 3 and 5.
- Haematological morbidity increased in the UKR group between POD 3 and 5 due to 1 patient requiring a red cell transfusion following abdominal surgery to repair a perforated gastric ulcer.

- Pain morbidity increased in the RTHR group between POD 8 and 15 due to 1 patient having pain following a second surgical procedure to reduce and fix a periprosthetic fracture.

2.4.3 Differences in post-operative morbidity between types of lower limb arthroplasty

The only risk-adjusted statistical difference in post-operative morbidity was found between primary and revision hip arthroplasty. The presence of any type of morbidity was increased following RTHR compared to THR on POD 3, 5, 8 and 15. There was more infectious morbidity following RTHR compared to THR on POD 3, 5, 8 and 15. There was more renal morbidity following RTHR compared to THR on POD 3.

This result is expected since revision hip arthroplasty usually takes longer to perform than primary arthroplasty, it usually involves a larger incision with greater dissection of tissues and there is often greater blood loss. Therefore it seems intuitive that patients will have more morbidity following revision hip arthroplasty than primary arthroplasty.

The indication for many revision arthroplasty procedures is infection of the primary prosthesis. For this reason a higher rate of post-operative infection is expected in the revision arthroplasty group since infection is present pre-operatively. The higher rate of renal morbidity in the revision arthroplasty group

was mainly due to an increased number of patients with a urinary catheter in-situ. Revision arthroplasty usually involves greater blood loss than primary arthroplasty and anaesthetic times are usually longer. For these reasons more revision arthroplasty patients require monitoring of their fluid balance in the early post-operative period (necessitating the placement of a urinary catheter), which could account for the difference in renal morbidity levels.

No difference in post-operative morbidity was seen between primary and revision knee arthroplasty. The most likely reason for this is the low number of patients in the RTKR group. If greater numbers had been included in the study, a difference may have been found. This area requires further study.

After adjusting for POSSUM morbidity scores, no difference in post-operative morbidity was found between UKR and TKR, and HR and THR. UKR and HR preserve more bone stock than total total joint replacements making revision surgery potentially easier. For this reason UKR and HR are often described as 'less invasive' than total joint replacements. This study shows that UKR and HR produce the same levels of post-operative morbidity as total joint arthroplasty. However, the study populations in the UKR and HR groups are small and this result may represent a type 2 error. Further investigation with larger study numbers is needed for further clarification.

After adjusting for POSSUM morbidity scores, no difference in post-operative morbidity was found between primary hip and knee arthroplasty. Similarly,

no difference was found between revision hip and knee arthroplasty. Arthroplasty procedures on the hip inflict the same levels of post-operative morbidity as arthroplasty procedures on the knee.

2.4.4 Comparison with other morbidity estimates

Post-operative morbidity following primary and revision lower limb arthroplasty has not previously been compared using the POMS method. There are few studies in the orthopaedic literature that directly compare morbidity following primary and revision arthroplasty. There are many case series reporting the outcome of various types of arthroplasty. Direct comparisons between series are difficult due to the reporting of different measures of morbidity over varying timescales. Furthermore, most studies do not adjust for patient and surgical factors when reporting morbidity.

Most case series and studies comparing alternate prostheses report the incidence of limb-related morbidity only. Commonly reported morbidity includes deep vein thrombosis[123-125], periprosthetic fracture[126-128] and deep surgical site infection[129-131]. Our study collected morbidity relating to complete organ systems rather than individual diagnoses. There are very few studies in the literature that explore all aspects of morbidity following arthroplasty but there are some case series that report different aspects of non-limb related post-operative morbidity. These papers will be discussed below in relation to the findings from this study.

2.4.4.1 Pulmonary morbidity

Wallace et al[132] reported on a series of 31,510 primary hip arthroplasty patients and 32,303 primary knee arthroplasty patients. Data was collected from GP databases and was analysed retrospectively. Morbidity was reported for six months following surgery. Respiratory tract infection rate was 0.55% in the hip group and 0.59% in the knee group.

Mantilla et al[133] retrospectively examined the medical records of 10,244 primary hip and knee arthroplasty patients and reported morbidity for 30 days following surgery. The rate of pulmonary embolus was 0.7%.

Parvizi et al [134] conducted a prospective analysis of post-operative morbidity in 1,636 patients for 6 weeks following primary hip and knee replacement. They reported 25 pulmonary emboli (1.53%) and 4 cases of pneumonia (0.24%).

The POMS definition of pulmonary morbidity is ‘a new need for oxygen or respiratory support’. Cardiac pathologies, as well as primary pulmonary pathologies, often require the administration of oxygen. This could partially explain the higher rate of pulmonary morbidity reported in this study (12.0% on POD 3 in the TKR group and 15.4% on POD 3 in the THR group) compared to the studies described above. A further explanation could be that this study was conducted prospectively where as the first two studies described above were

performed retrospectively. Prospective studies will always capture more events than retrospective studies. A further explanation could be that oxygen was given unnecessarily to patients in this study.

2.4.4.2 Infection

Wallace et al[132] reported a respiratory infection rate of 0.55% following hip arthroplasty and 0.59% following knee arthroplasty. The wound infection rate was 2.23% in the hip group and 3.41% in the knee group. The urinary tract infection rate was 2.45% in the hip group and 2.23% in the knee group. This gives an overall infection rate of 5.23% in the hip arthroplasty group and 6.23% in the knee arthroplasty group (although this summation may not be accurate since more than one type of infection may have occurred in the same patients).

Parvizi et al [134] reported an overall urinary tract infection rate of 3.06% following hip and knee arthroplasty. The rate of pneumonia was 0.24% and the rate of wound infection was 0.67%. This gives an overall combined rate of infection of 3.97% (although again, this summation may be inaccurate for the same reasons as described above).

Pulido et al[135] reported on prospectively collected data from 9,245 patients following primary hip and knee arthroplasty. The rate of deep infection was 0.7%. Mean time of follow-up was 43 months. Other types of infection were not reported.

The rate of infectious morbidity in this study (18.1% for TKR on POD 3 and 20.4% for THR on POD 3) is higher than in the studies described above. There are several possible explanations for this. The most obvious explanation is a difference in the definition of infectious morbidity. The POMS definition is a patient 'currently taking antibiotics' and/or has had 'a temperature of 38ºC or above in the last 24 hours'. There are many indications for antibiotics including prophylaxis and the treatment of infection at any site (e.g. upper respiratory tract infection, ear infection, phlebitis). Patients taking antibiotics for these reasons would not have been captured in the studies described above. Also, a temperature over 38ºC does not always indicate an infection. Many hip and knee replacements are performed in patients with rheumatoid arthritis. These patients can have a rise in body temperature due the inflammatory nature of their underlying condition rather than an active infection. Therefore, our infectious morbidity rate may be an over-estimate. The fact this study was conducted prospectively rather than retrospectively could also partially account for the higher rates of infection.

2.4.4.3 Renal morbidity

Parvizi et al [134] reported an acute renal failure rate of 0.86%. The rate of renal morbidity was much higher in our study (22.6% for TKR on POD 3 and 35.8% for THR on POD 3). The most likely reason for this higher rate of morbidity is the POMS definition: 'oliguria of less than 500ml/day, a raised creatinine level of over 30% compared to pre-operatively, or a urinary catheter in-situ'. Most renal morbidity in our study was due to the presence of a urinary catheter. Out of

the 529 patients in our study, only 1 patient developed acute renal failure (0.19%). This rate of renal failure is comparable with that reported by Parvizi.

2.4.4.4 Gastro-intestinal morbidity

Parvizi et al [134] reported gastric immobility in 0.73% of patients and nausea and vomiting in 0.61% of patients following arthroplasty. Rates of gastrointestinal morbidity were higher in our study with the highest rates seen on POD 3 (16.8% in TKR group and 11.7% in THR group). POMS defines gastrointestinal morbidity as 'unable to tolerate an enteral diet' or the patient experiencing 'nausea, vomiting or abdominal distension'. Parvizi would not have identified patients unable to tolerate an enteral diet as having morbidity, which could largely explain the difference in the figures between the 2 studies.

2.4.4.5 Cardiovascular morbidity

Wallace et al[132] reported a combined rate of pulmonary embolus and deep vein thrombosis of 2.86% following hip arthroplasty and 2.76% following knee arthroplasty. The rate of myocardial infarction in the hip group was 0.40% and in the knee group 0.41%.

Mantilla et al[133] reported a 0.4% rate of myocardial infarction, 0.7% rate of pulmonary embolus and 1.5% rate of deep vein thrombosis following hip and knee arthroplasty.

Lalmohammed et al[136] investigated 95,227 patients for 6 weeks following primary THR and TKR. The data was collected retrospectively using Danish health registries. The rate of myocardial infarction in the THR group was 0.51% and in the TKR group was 0.21%.

Pedersen et al[137] used retrospective information from Danish health registries to report on the rate of cardiovascular events in 83,756 patients following primary hip and knee arthroplasty. The combined rate of myocardial infarction and stroke was reported as 0.5%, the rate of symptomatic venous thromboembolism was 1.3% and the rate of a major bleed was 0.6%.

Parvizi et al [134] reported a 0.06% rate of cardiac arrest, 5.07% rate of arrhythmias, 0.61% rate of congestive cardiac failure, 0.37% rate of myocardial infarction, 0.24% rate of hypotensive crisis, 0.37% rate of pulmonary embolus and 1.53% rate of deep vein thrombosis. This gives a combined value of 8.25% (but this summation may be invalid since individual patients may have experienced more than one type of cardiovascular morbidity).

The rate of cardiovascular morbidity in our study (4.4% of TKR patients and 4.9% of THR patients on POD 5) was similar to the rate reported by Parvizi. The definition of POMS cardiovascular morbidity is ‘ischaemia or hypotension requiring drug therapy or fluid therapy of over 200ml per hour, atrial or ventricular arrhythmia, cardiogenic pulmonary oedema or new anticoagulation’.

Therefore our study and Parvizi's study collected similar data allowing a meaningful comparison to be made.

If the cardiovascular morbidity domain of POMS is broken down into its components, the rate of MI in our study (0.4% following TKR and 1.2% following THR) was similar to that reported by Wallace, Mantilla, Lalmohammed and Parvizi. The rate of DVT (2.2% following TKR and 1.2% following THR) and PE (0% following TKR and 0.6% following THR) in our study was also comparable or lower than that reported by Wallace, Mantilla and Parvizi.

2.4.4.6 Neurological morbidity

Wallace et al[132] reported a 0.50% rate of stroke following primary hip arthroplasty and a 0.44% rate of stroke following knee arthroplasty. Lalmohamed et al[138] retrospectively looked at 66,583 patients using Danish health registries for 6 weeks following primary THR. An overall stroke rate of 1.14% was reported (0.22% ischaemic, 0.56% haemorrhagic, 0.36% unspecified).

Parvizi et al [134] reported a stroke rate of 0.37% following hip and knee arthroplasty. Mortazavi[139] et al reported a 0.2% rate of stroke at a mean of 62 months following primary hip and knee arthroplasty.

The rate of neurological morbidity in our study (2.2% in TKR group at post-operative day three and 1.2 % in THR group on post-operative day three) is

higher than in the studies described above. The most likely reason for this is the wording of POMS definition of neurological morbidity: 'new confusion/ delirium, focal deficit or coma'. The studies described above only report the rate of stroke. This study reports the combined rate of stroke, transient ischaemic attack, confusion and reduced consciousness. The rate of stroke in our study was 1/226 (0.4%) in the TKR group and 0/162 (0%) in the THR group, which is comparable or better than the rates in the studies described above.

2.4.4.7 Wound morbidity

Wallace et al[132] reported a wound infection rate of 2.23% following hip arthroplasty and 3.41% following knee arthroplasty. Parvizi et al [134] reported a 4.52% rate of persistent wound discharge and a 0.67% rate of wound infection.

Pulido et al[135] reported a deep infection rate of 0.7% in 9,245 patients following primary hip and knee arthroplasty. Data was collected prospectively. Mean time of follow-up was 43 months.

The definition of POMS morbidity is 'wound dehiscence requiring surgical exploration or drainage of pus from the operative wound with or without isolation of organisms'. The rate of wound morbidity was higher in our study (6.2% in TKR patients and 9.3% in THR patients on POD 5) than in the studies described above. Possible reasons for this discrepancy include the use of different definitions of wound morbidity and the use of different methods of data collection (prospective data collected at the patient's bedside vs

retrospective data collected from medical notes). However, these explanations cannot be assumed and a real difference between this study and previous studies may exist. Surgical site infection is a major risk of arthroplasty surgery, with potentially catastrophic consequences. For these reasons, I will investigate the wound domain of POMS further in Chapter 5 of this thesis.

2.4.4.8 Haematological morbidity

Verlicchi et al[140] reported a 69% (185/268) red cell transfusion rate for HR, 53% (263/497) transfusion rate for THR, 72% (71/98) for RTHR, 43% (233/541) for TKR and 41% (9/22) for RTKR. Evans et al[141] reported a red cell transfusion rate of 0.9 % (1/102) following TKR, 13.5% (7/52) following THR, 27.3% (3/11) following RTKR and 56.3% (9/16) following RTHR. Gombotz et al[142] reported a 30% (367/1223) red cell transfusion rate following THR and 25% (307/1227) transfusion rate following TKR.

The POMS definition of haematological morbidity is 'transfusion of red blood cells, platelets, fresh frozen plasma or cryoprecipitate within the last 24 hours'. The only blood product given to any arthroplasty patient following surgery was red cells. The rates of transfusion in this study are much lower than those reported in the studies above. The most likely reason for this is the fact that data was only collected on POD 3, 5, 8 and 15. Red cells transfused in the preceding 24 hours were recorded. Therefore, transfusions that were given on other days (e.g. POD 1 and 2) would not have been captured in this study. The rate of transfusion also depends on other factors such as the patient's pre-

operative haemoglobin level, the volume of operative blood loss, use of an intra-operative 'cell saver', the quality of surgical haemostasis and wound closure, and the threshold haemoglobin level for transfusion. Any of these factors could have contributed to the lower transfusion level in our study.

2.4.4.9 Pain

There are many studies examining pain following hip and knee arthroplasty[143-145]. It is difficult to make meaningful comparisons between studies since different tools are used to assess pain severity, pain is reported at different time points and different analgesia regimens are used. Petre et al[146] reported on a series of 352 patients following hip and knee arthroplasty. Patients had a mean pain score of between 4 and 5 (on a VAS scale with a maximum score of 10) in the first five post-operative days.

The POMS definition of pain morbidity is 'surgical wound pain significant enough to require parenteral opioids or regional anaesthesia'. A low proportion of patients were reported as having pain morbidity (7.1% of TKR patients on POD 3 with none thereafter, 6.3% of THR patients POD 5 with none thereafter). These figures appear favourable to the study mentioned above. However, since different measures of pain have been used, direct comparisons may not be valid.

2.4.5 Strengths of the study

This study has several strengths. Firstly, there is a large population sample size. Secondly, specially trained staff collected complete prospective data sets. The same clearly defined data was collected on all patients allowing meaningful comparisons between arthroplasty groups.

2.4.6 Limitations of the study

There are several limitations to this study. Firstly, this study was conducted at a single site and therefore the results may not be transferable to other centres. Secondly, the number of revision arthroplasty patients was lower than the primary arthroplasty groups. Therefore, results regarding morbidity following revision arthroplasty may be less accurate than those regarding morbidity following primary arthroplasty. This is particularly pertinent for the revision knee arthroplasty group, which only contained 8 patients.

2.5 Summary

1. Morbidity following arthroplasty has previously been poorly recorded.
2. This is the first time the POMS has been used to record morbidity following hip and knee arthroplasty.
3. The most common types of morbidity following hip and knee arthroplasty are infection and renal morbidity. Pulmonary, pain and gastro-intestinal morbidity are less common. Cardiovascular, wound, neurological and haematological morbidity are the least common.

4. Most post-operative morbidity decreases with time after surgery.
5. Medical staff treating arthroplasty patients should be aware of the prevalence of post-operative morbidity in all organ systems. This allows prompt diagnosis and treatment of morbidity, minimising its impact.
6. Knowing the prevalence of morbidity following arthroplasty allows appropriate pre-operative patient education and counselling. This enables patients to have realistic expectations regarding the post-operative period.
7. There is a statistically significant risk-adjusted difference in the presence of any type of post-operative morbidity following primary and revision hip arthroplasty on POD 3, 5, 8 and 15. Morbidity is higher in the RTHR group.
8. There are statistically significant risk-adjusted higher levels of infection following RTHR than THR on POD 3, 5, 8 and 15. On POD 3, there are higher levels of renal morbidity following RTHR than THR.
9. There is no difference in post-operative morbidity following primary and revision knee arthroplasty. The most likely reason for this is the low number of patients in the RTKR group.
10. After adjusting for POSSUM morbidity scores, no difference in post-operative morbidity was found between UKR and TKR, and HR and THR. This could be a result of a type 2 error due to a small study population in the UKR and HR groups. Further investigation with larger study numbers is required.

11. After adjusting for POSSUM morbidity scores, no difference in post-operative morbidity was found between hip and knee procedures.
12. Most levels of post-operative morbidity in this study are higher than in previously published studies. The use of different definitions of morbidity could partially account for the discrepancy. The definitions of POMS morbidity in different organ systems are broad and will therefore capture more events than narrowly defined measures of morbidity. Different methods of data collection could also account for the discrepancy; this study collected data prospectively, which will always capture more events than retrospectively collected data.
13. This study provides baseline data against which future audits can be compared. Strategies can be implemented to reduce morbidity levels and the audit cycle repeated to evaluate their impact.

Chapter 3: POMS as a bed utilisation tool

3.1 Introduction

This chapter assesses the use of the POMS as a bed utilisation tool and its ability to identify inappropriate bed occupancy following hip and knee arthroplasty. I will describe the methods used to collect and analyse the POMS data and I will present the results for both hip and knee arthroplasty procedures.

The number of inappropriate bed occupancy days for each type of arthroplasty will be calculated together with the potential cost saving if patients with no morbidity were discharged at the earliest opportunity. I will describe the reasons why patients remained in hospital with no identifiable morbidity. I will identify patients who developed morbidity following a period with no morbidity and report the number of patients re-admitted to the same hospital within one year of surgery. I will conclude with a discussion of the results including the strengths and weaknesses of this study. I will compare the results to other bed utilisation studies. Finally I will present a summary of key findings.

Appropriately timed discharge of patients following surgery is essential for optimal patient care and efficient hospital functioning. A patient discharged early is at risk of under-diagnosis of medical complications with consequent adverse outcome. A patient whose discharge is delayed is at risk of

developing a hospital-associated complication (e.g. hospital-acquired infection) and incurs an unnecessary cost to the health-care provider. Post-operative patients should be discharged at the earliest safe opportunity to reduce the rate of hospital-associated complications and the cost of each inpatient episode. Appropriate discharge timing should increase patient throughput and reduce waiting times.

Historically, hospitals in the UK were paid according to contracts with no financial incentive to treat increased numbers of patients. This changed in 2000 when the NHS Plan[147] announced that hospital income would be directly linked to activity. Payment by Results[6] began in 2003 and now every healthcare provider is paid a sum (tariff) for each procedure undertaken. In the UK many patients remain in hospital with no medical indication[18]. One study showed that 31% of post-operative patients remained in hospital inappropriately [85]. Payment by Results aims to reduce this figure by rewarding efficiency and encouraging increased activity.

In order to improve efficiency, hospitals must first recognise inappropriate bed occupancy. The Post-Operative Morbidity Survey (POMS)[20] is the only validated prospective method of measuring short-term post-operative morbidity in the literature.

In the US over 98% of post-operative inpatients had morbidity defined by the POMS[20]. This implies that patients with a POMS score of zero are fit for

discharge. Therefore, as well as providing useful clinical research and audit data, the POMS may have utility for assessing and improving hospital bed utilisation.

The aim is this chapter is to assess the utility of the POMS as a bed utilisation tool and its ability to identify inappropriate bed occupancy following hip and knee arthroplasty. The reasons patients remain in hospital with no identifiable morbidity will be reported and particular attention will be given to 2 groups of patients: those who developed morbidity following a period with no morbidity and patients re-admitted to the same hospital within one year of surgery. Inappropriate bed utilisation rates will compared to other bed utilisation studies.

3.2 Methods

The methods used to collect the POMS data are described fully in Chapter 2. The patient demographics are also described in detail. POMS data was collected on post-operative days (POD) 3, 5, 8 and 15 if the patient remained in hospital. Presence of post-operative morbidity was defined as occurring in any patient meeting POMS criteria for morbidity in one or more domain of the survey on the day of data collection. The reason for patients remaining in hospital without morbidity was recorded on POD 8 and 15. The use of mobility aids on these days was also recorded.

The number and percentage of patients with no identifiable morbidity according to the POMS was calculated for POD 3, 5, 8 and 15. The number of days a

patient remained in hospital with no morbidity was calculated by subtracting the number of the POD on which the patient first had a POMS score of zero from their total length of stay. An overall estimated cost saving was calculated by multiplying this figure by the average cost for one inpatient night on an orthopaedic ward.

The number of patients with post-operative morbidity, subsequent to having a period free of morbidity, was recorded. The number of readmissions to the same hospital in the first year following discharge was also recorded.

3.3 Results

Data collection was completed on 529 patients. Characteristics of the study population are shown in tables 8 and 9 in Chapter 2. The mean age of the study population was 68.9 years, the median ASA grade was 2 and 62% of patients were female. The median length of stay was 7 days and the overall inpatient mortality rate was 0.4%.

3.3.1 Hip arthroplasty patients

The location of hip arthroplasty patients on POD 3, 5, 8 and 15 is shown in table 28.

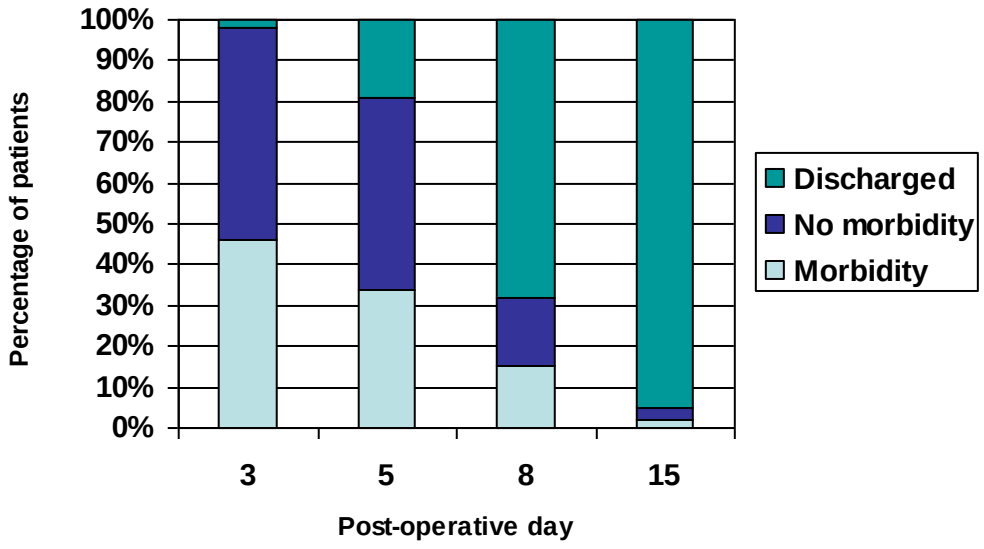
Table 28. Location of 229 hip arthroplasty patients on POD 3, 5, 8 and 15

			POD 3	POD 5	POD 8	POD 15
Procedure	**HR**	Patients discharged	0/32 (0%)	2/32 (6%)	27/32 (84%)	32/32 (100%)
		Inpatients POMS >0	8/32 (25%)	5/32 (16%)	0/32 (0%)	0/32 (0%)
		Inpatients POMS = 0	24/32 (75%)	25/32 (78%)	5/32 (16%)	0/32 (0%)
	THR	Patients discharged	0/162 (0%)	13/162 (8%)	78/162 (48%)	138/162 (85%)
		Inpatients POMS >0	87/162 (54%)	62/162 (38%)	29/162 (18%)	13/162 (8%)
		Inpatients POMS = 0	75/162 (46%)	87/162 (54%)	55/162 (34%)	11/162 (7%)
	RTHR	Patients discharged	0/35 (0%)	0/35 (0%)	3/35 (9%)	20/35 (57%)
		Inpatients POMS >0	31/35 (89%)	25/35 (71%)	21/35 (60%)	14/35 (40%)
		Inpatients POMS = 0	4/35 (11%)	10/35 (29%)	11/35 (31%)	1/35 (3%)
	TOTAL	Patients discharged	0/229 (0%)	16/229 (7%)	109/229 (47%)	191/229 (83%)
		Inpatients POMS >0	127/229 (55%)	92/229 (40%)	50/229 (22%)	27/229 (12%)
		Inpatients POMS = 0	103/229 (45%)	122/229 (53%)	71/229 (31%)	12/229 (5%)

A significant proportion of hip resurfacing patients remained in hospital with no morbidity on POD 3 (75%) and POD 5 (78%). This reduced to 16% by POD 8. All hip resurfacing patients were discharged by POD 15. Similarly, a significant proportion of total hip replacement patients remained in hospital with no identifiable morbidity on POD 3 (46%) and POD 5 (54%). This reduced to 34% by POD 8 and to 7% by POD 15. The proportion of revision total hip replacement patients remaining in hospital with no identifiable morbidity was lower than the other hip arthroplasty groups (11% on post-POD 3, 29% on POD 5, 31% on POD 8 and 3% on POD 15).

Discharge status and prevalence of morbidity for all hip arthroplasty patients combined are presented in figure 1.

Figure 1 Discharge status and prevalence of morbidity following hip arthroplasty



3.3.2 Knee arthroplasty patients

The location of patients on POD 3, 5, 8 and 15 is shown in table 29.

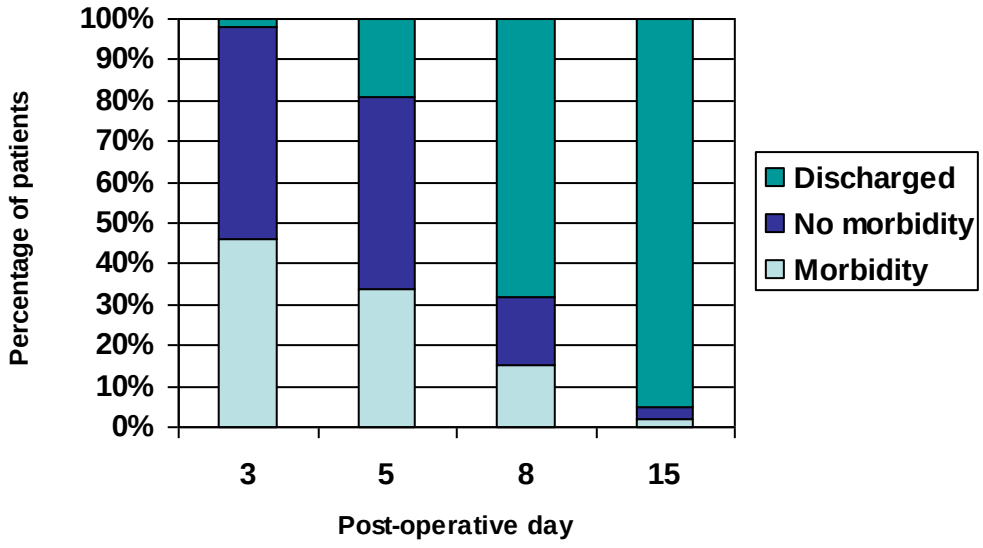
Table 29. Location of 300 knee arthroplasty patients on POD 3, 5, 8 and 15

			POD 3	POD 5	POD 8	POD 15
Procedure	**UKR**	Patients discharged	7/66 (11%)	33/66 (50%)	59/66 (89%)	65/66 (98%)
		Inpatients POMS >0	17/66 (26%)	7/66 (11%)	3/66 (5%)	1/66 (2%)
		Inpatients POMS = 0	42/66 (63%)	26/66 (39%)	4/66 (6%)	0/66 (0%)
	TKR	Patients discharged	0/226 (0%)	22/226 (10%)	145/226 (64%)	211/226 (93%)
		Inpatients POMS >0	114/226 (50%)	90/226 (40%)	38/226 (17%)	7/22 (3%)
		Inpatients POMS = 0	112/226 (50%)	114/226 (50%)	43/226 (19%)	8/226 (4%)
	RTKR	Patients discharged	0/8 (0%)	1/8 (13%)	1/8 (13%)	6/8 (75%)
		Inpatients POMS >0	6/8 (75%)	4/8 (50%)	5/8 (62%)	1/8 (12.5%)
		Inpatients POMS = 0	2/8 (25%)	3/8 (37%)	2/8 (25%)	1/8 (12.5%)
	TOTAL	Patients discharged	7/300 (2%)	56/300 (19%)	205/300 (68%)	282/300 (94%)
		Inpatients POMS >0	137/300 (46%)	101/300 (34%)	46/300 (15%)	9/300 (3%)
		Inpatients POMS = 0	156/300 (52%)	143/300 (47%)	49/300 (17%)	9/300 (3%)

A significant proportion of UKR patients remained in hospital with no morbidity on POD 3 (63%) and POD 5 (39%). This reduced to 6% by POD 8. All UKR patients were discharged by POD 15. Similarly, a significant proportion of TKR patients remained in hospital with no identifiable morbidity on POD 3 (50%) and POD 5 (50%). This reduced to 19% by POD 8 and to 4% by POD 15. The proportion of RTKR patients remaining in hospital with no identifiable morbidity was similar to the primary knee arthroplasty group (52% on POD 3, 47% on POD 5, 17% on POD 8 and 3% on POD 15).

Discharge status and prevalence of morbidity for all knee arthroplasty patients combined are presented in figure 2.

Figure 2 Discharge status and prevalence of morbidity following all types of knee arthroplasty



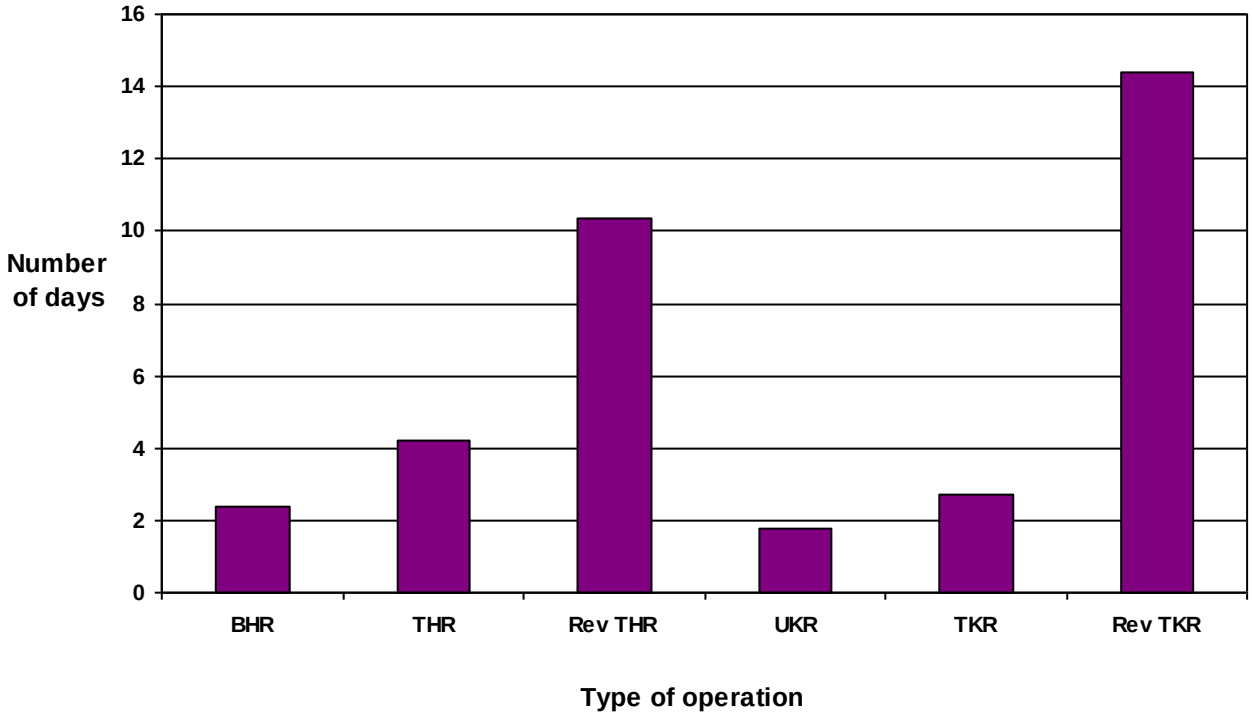
3.3.3 Overall inappropriate bed occupancy days

Table 30 and figure 3 show the average number of days that patients remained in hospital with no identifiable morbidity. HR patients stayed an average of 2.36 days, THR patients 4.19 days and revision THR patients 10.37 days. UKR patients stayed an average of 1.76 days with no identifiable morbidity, TKR patients 2.73 days, and revision TKR patients 14.38 days.

Table 30. Number of inappropriate inpatient days classified by type of arthroplasty

	Total number of patients	Total number of inappropriate inpatient days	Average number of inappropriate inpatient days per patient
HR	32	78	2.43
THR	162	678	4.19
RTHR	35	363	10.37
UKR	66	111	1.68
TKR	226	620	2.74
RTKR	8	115	14.38
Total	529	1965	3.71

Figure 3. Average number of inappropriate inpatient days classified by type of arthroplasty



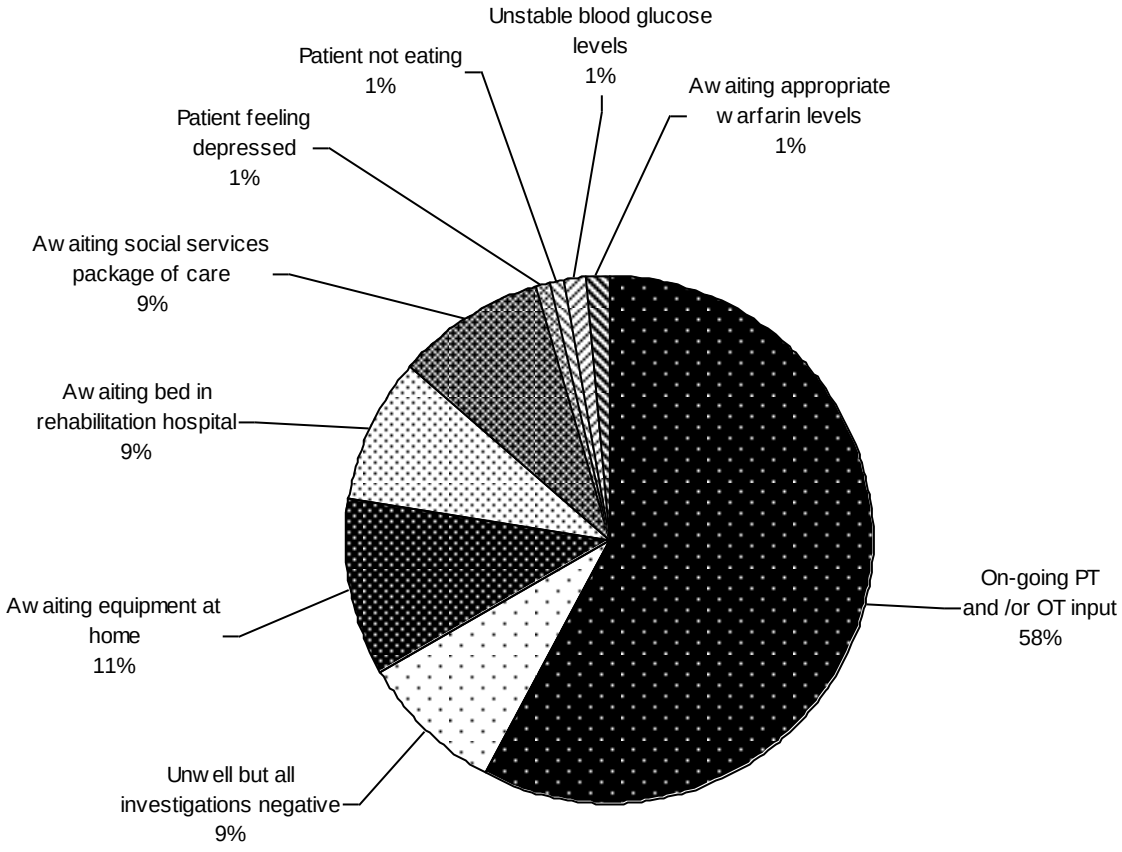
3.3.4 Cost of inappropriate bed occupancy days

529 patients were included in this study. These patients remained in hospital for a total of 1965 days with no morbidity as defined by the POMS. A surgical inpatient bed costs up to £400 per night[148]. If these patients had been discharged when their POMS score was zero a saving of up to £786,000 could have been made.

3.3.5 Reasons for patients with no morbidity remaining in hospital

Of the 529 patients participating in this study, 120 remained in hospital with no morbidity defined by the POMS on POD 8. By POD 15, there were 20 patients in hospital with no identifiable morbidity. The reasons given for non-discharge are shown in figure 4.

Figure 4. Reasons hip and knee arthroplasty patients remained in hospital with no morbidity on post-operative days 8 and 15



The most common reason for patients remaining in hospital with no morbidity was on-going physiotherapy and occupational therapy input. Other reasons included waiting for home equipment, waiting for a rehabilitation bed, waiting for a social services package of care and patients feeling unwell with negative investigations.

Of the patients remaining in hospital with no morbidity identified by the POMS, 24% were mobilising with a Zimmer frame, 55% were mobilising with two crutches, 14% with a single crutch and 7% were mobilising unaided. This study did not record how far each patient could mobilise or if they were able to climb stairs.

3.3.6 New morbidity and readmission

Several patients in the study developed morbidity as an inpatient following a period with no morbidity. Re-starting anticoagulation medication that the patient was taking pre-operatively was not counted as morbidity in this evaluation.

Some patients had a second surgical procedure during their inpatient stay. They developed morbidity after the second surgical procedure following a period without morbidity between the two procedures. This occurred in the following situations:

- 1 UKR patient developed a perforated duodenal ulcer 2 days after his knee surgery. This patient developed infectious, cardiovascular and haematological morbidity following his laparotomy.

- 1 RTKR patient had no morbidity on POD 8. He subsequently developed a deep knee infection and returned to theatre on POD 12 for knee washout and change of liner. On post-operative day 15 he had pulmonary, renal and gastro-intestinal morbidity.
- 3 RTHR patients had no morbidity on POD 8. 2 of these patients developed deep infection and returned to theatre on POD 9 and POD 12 respectively for hip washout and change of liner. They both had pulmonary morbidity and 1 had gastro-intestinal morbidity on POD 15. The other patient fell and sustained a periprosthetic fracture. He returned to theatre for surgical fixation. On POD 15 he had pulmonary and pain morbidity.

Some patients developed morbidity after a period with no morbidity without a further surgical procedure. This occurred in the following cases:

- 4 TKR patients had no morbidity on POD 3 and developed cardiovascular morbidity by POD 5. 3 patients had deep vein thromboses and 1 patient had a myocardial infarction.
- 8 TKR patients had no morbidity on POD 3 and developed wound morbidity on POD 5.
- 1 RTKR had no morbidity on POD 5 and developed wound morbidity by POD 8. The indication for revision surgery was infection of the primary prosthesis.
- 2 HR patients had no morbidity on POD 3 and developed wound morbidity by POD 5.

- 8 THR patients had no morbidity on POD3 and developed wound morbidity by POD 5.
- 1 THR patient had no morbidity on POD 5 and developed neurological morbidity by POD 8 due to a CVA.
- 1 THR patient had no morbidity on POD 8 and developed infectious morbidity by POD 15 due to an infected peripheral intra-venous cannula site. This could have been avoided if the cannula had been removed and the patient discharged on POD 8.
- 2 RTHR patients had no morbidity and then developed cardiovascular morbidity (on POD 5 and 8 respectively) due to the commencement of anticoagulation for pulmonary emboli.
- 6 RTHR patients had no morbidity on POD 3 and developed wound morbidity by POD 5. The indication for revision surgery in these patients was infection of the primary prosthesis.

In summary, 33 out of 501 patients who underwent a single surgical procedure developed morbidity as defined by the POMS after a period of time with no morbidity. 25 of these patients developed wound morbidity, 6 developed cardiovascular morbidity, 1 patient developed neurological morbidity and 1 patient developed infectious morbidity.

Of the 25 patients who developed wound morbidity, 24 of them had no morbidity on POD 3 and developed wound morbidity by POD 5. 1 RTKR patient had no morbidity on POD 5 and developed wound morbidity by POD 8.

Of the 6 patients that developed cardiovascular morbidity following a period with no morbidity, 5 of them had no morbidity on POD 3 and developed morbidity by POD 5. 3 patients had a DVT, 1 had a PE and 1 had an MI. 1 RTHR patient had no morbidity on POD 5 and developed cardiovascular morbidity on POD 8. This patient had a PE.

1 THR patient had no morbidity on POD 5 and had a CVA by POD 8. This is a rare event (1/529). 1 patient had no morbidity on POD 8 and developed infectious morbidity (an infected peripheral cannula site) by POD 15.

No patient in this study was readmitted to the same hospital within one year of discharge for any reason relating to his or her surgery.

3.4 Discussion

3.4.1 Summary

This study identifies that many patients remain in hospital with no identifiable morbidity following hip and knee arthroplasty in a UK teaching hospital. The rate of inappropriate bed occupancy varies with the type of arthroplasty and time after surgery. Two general trends are seen. Firstly, the more 'invasive' the surgery, the longer a patient remains in hospital with no identifiable morbidity. Thus revision arthroplasty patients remain in hospital with no morbidity longer

than primary total knee/hip replacement patients who remain longer than unicondylar knee/hip resurfacing patients. Secondly, the proportion of patients remaining in hospital with no morbidity decreases with time following surgery. This trend is seen in the UKR, TKR, HR and THR groups. In the revision arthroplasty groups there is a rise in patients with no morbidity between POD 3 and 5. The proportion then decreases.

The most common reason for patients remaining in hospital with no identifiable morbidity was on-going physiotherapy and occupational therapy. This suggests that improving both pre- and post-operative therapy planning could reduce inappropriate bed occupancy. Prior to surgery, patients could be taught their post-operative physiotherapy exercises in group classes. Occupational therapists could assess each patient's home environment and ensure necessary modifications are made prior to surgery. In the post-operative period 'fast-track' pathways could be used to ensure maximum therapy input at the earliest possible opportunity. Some physiotherapy and occupational therapy could be provided post-operatively at the patient's home rather than as an inpatient. This would require safety and cost evaluation prior to implementation.

3 of the top 5 reasons for patients remaining in hospital with no identifiable morbidity relate to 'social' issues (awaiting home equipment, awaiting a rehabilitation bed, awaiting a package of care from social services). Pre-operative clinics could identify and address these problems prior to admission. Such clinics could also be used to manage patient expectation so that each

patient is made aware of the difficulties they may encounter in the post-operative period and the expected timing of discharge.

This is the first time the POMS has been used as a bed utilisation tool. It has not been validated for this purpose but has previously been used to identify patients in hospital without morbidity[20,21]. In the US over 98% of inpatients had morbidity defined by the POMS[20] suggesting that those with a POMS score of zero were rapidly discharged. In a previous UK study, 63% of orthopaedic patients remained in hospital with no morbidity on POD 3 and 42% on POD 5 suggesting that discharge efficiency was lower in the UK institution.

Use of the POMS as a bed utilisation tool relies on the assumption that it captures all reasons for remaining in hospital. In this study, the main reason for remaining in hospital with no identifiable morbidity was 'on-going physiotherapy and occupational therapy input'. A specific concern in this patient group is that these patients may not be sufficiently mobile to be discharged safely. Including a specific domain for mobility may improve the sensitivity of the POMS for morbidity requiring hospitalisation following orthopaedic surgery. Criteria for a positive result could include an inability to mobilise 10 metres or climb a single flight of stairs. Whilst this domain could be especially relevant for orthopaedic patients, this requires further investigation.

Use of the POMS as a "fitness for discharge" tool in hospital rests on the assumption that patients do not develop new morbidity after they have

become free from morbidity, either in hospital or following discharge. 33 of 442 (7%) patients remaining in hospital with no morbidity subsequently developed 'new' morbidity without undergoing any further surgery. Following discharge, no patients were readmitted to the study hospital in the first post-operative year for complications linked to surgery.

Of the 33 patients who developed 'new' morbidity, 25 developed wound morbidity. A proportion of wound infections may have been due to hospital acquired infection. If these patients had been discharged when first free of morbidity, they may not have developed infection. However, this cannot be assumed. Of the 25 patients who developed wound morbidity, 24 of them had no morbidity on POD 3 and developed wound morbidity by POD 5. Thus, if an arthroplasty patient is discharged before POD 5, regular wound review should be performed by a medical practitioner up until this day. 1 RTKR patient had no morbidity on POD 5 and developed wound morbidity by POD 8. Thus, if a revision arthroplasty patient is discharged before POD 8, regular wound review should be performed by a medical practitioner up until this day. A doctor or nurse, in either a primary or secondary care setting, could perform the review. If this protocol is followed, there should be prompt diagnosis and treatment of surgical site infection, thus minimising its impact.

1 patient developed infectious morbidity (an infected peripheral cannula site) after a period with no morbidity. If the patient had been discharged when first

free of morbidity, the cannula would have been removed and this infection would have been prevented.

Of the 6 patients who developed cardiovascular morbidity following a period with no morbidity, 3 had a DVT, 2 had a PE and 1 had an MI. 1 THR patient had a CVA after a period with no morbidity. These results stress the importance of patient education regarding the symptoms of DVT, PE, MI and CVA prior to discharge. Patients should be made aware that these morbidities could occur following a period of feeling 'well' with no apparent morbidity. Patients should be given clear written instructions about what to do if they suspect one of these complications is occurring. As long as these precautionary measures are put in place, POMS has potential as a bed utilisation tool.

3.4.2 Strengths of study

This study has several strengths. A large consecutive dataset was collected prospectively using a validated methodology for measuring post-operative morbidity. This is the first published study to prospectively evaluate the appropriateness of discharge following hip and knee arthroplasty surgery specifically.

3.4.3 Limitations of study

The weaknesses of this study are that it was conducted in a single centre, the POMS is not validated as a bed utilisation tool, there was not daily recording of data so the calculation of excess days are an approximation, and patient

mobility was not fully assessed. Data was collected regarding mobility aids but the distance each patient could mobilise was not recorded.

3.4.4 Comparison with other countries

The most commonly used tool to assess appropriate bed utilisation in the literature is the Appropriateness Evaluation Protocol (AEP)[86]. The AEP is a retrospective tool that relies on data from the inpatient record. It has been shown to be valid and reliable in some studies[86] but not in others[87]. The POMS is a prospective tool that could be used in real time to assist with appropriate patient discharge. The AEP is a retrospective tool that can only be used to evaluate past events. Data for the POMS is collected directly from contemporary data sources whilst the patient is in hospital; the AEP relies solely on past patient records and is therefore dependent on completeness and accuracy of record keeping for reliable functioning.

The AEP has been used in several European countries to examine bed utilization. In Portugal 50% of inpatient days were deemed inappropriate[88], in Italy 37.3%[18], in Germany 28%[89], in Switzerland 8-15%[90] and in France 7%[91]. This study indicates bed utilisation in the UK is comparable to that seen in Portugal and Italy but such a direct comparison may have limited validity since different bed utilisation tools have been used.

The finding that many fewer patients remain in hospital with no morbidity (as defined by the POMS) in the US[20] when compared with the UK suggests

that bed utilisation in the US is superior to that seen in the UK. The implementation of 'payment by results' in the UK aims to improve appropriate bed occupancy to optimise patient care and improve efficiency. If the patients in this study had been discharged when they first had no morbidity defined by the POMS, a saving of over £750,000 could have been made in one year (based on a cost of £400 per bed-day).

3.5 Summary

1. Many patients in the UK remain in hospital following lower limb arthroplasty with no identifiable morbidity.
2. The number of patients remaining in hospital with no morbidity varies with type of arthroplasty and time following surgery.
3. Two general trends are seen. Firstly, patients undergoing more 'invasive' surgery remain in hospital with no morbidity longer than patients undergoing less 'invasive' surgery. Secondly, with increasing time after surgery, fewer patients with no identifiable morbidity remain in hospital.
4. The most common reason for non-discharge of patients with no morbidity is on-going physiotherapy and occupational therapy input.
5. For the POMS to be used as a bed utilisation tool, it must capture all morbidity. Mobility is an important factor when assessing if an arthroplasty patient is fit for discharge. The POMS does not assess this. Addition of a mobility domain may make the POMS more reliable as a bed utilisation tool for orthopaedic patients.

6. Use of the POMS as a bed utilisation tool assumes that patients do not develop morbidity after a period without morbidity (when they would be discharged). A small proportion of inpatients developed wound, cardiovascular and neurological morbidity following a period without morbidity. To address this, certain precautionary measures need to be taken before and after patient discharge.
7. Primary arthroplasty patients should have wound reviews till POD 5 and revision arthroplasty patients should have wound reviews till POD 8.
8. Patients should be aware they could develop wound infection, deep vein thrombosis, pulmonary embolus, myocardial infarction, and stroke after hospital discharge. They should receive clear written instructions about the symptoms of these conditions and what to do if they occur.
9. Bed utilisation in the US is superior to that in the UK using the POMS as a measure. Savings could be made if bed utilisation was improved in the UK. POMS could be used to identify patients remaining in hospital without clinically significant morbidity and could be used prospectively as a bed utilisation tool.

Chapter 4: Can short-term post-operative morbidity predict longer-term outcome?

4.1 Introduction

Patient-reported outcome measures (PROMs) are commonly used to report the success or otherwise of an orthopaedic intervention[149,150]. One disadvantage of PROMs is the time taken to collect and analyse data. PROMs can be assessed at any point in time, from a few weeks post-surgery up to several years post-surgery. If an earlier surrogate marker for long-term PROMs could be found, operative and peri-operative procedures could be assessed more quickly and outliers identified. Thus both positive and negative results could be circulated earlier and appropriate action taken. This should ultimately lead to better patient care.

There is evidence that short-term post-operative morbidity is associated with a risk of premature death[151,152], but there is no evidence that morbidity is associated with longer-term patient reported outcome. The POMS is a short-term measure of morbidity. Data is collected whilst the patient is in hospital following surgery. This chapter will investigate whether there is any association between POMS results in the early post-operative period and PROM scores at 18 months post-surgery.

Traditionally, length of hospital stay has been used as an early marker of the success of a procedure. I will examine whether there is any association between length of hospital stay and PROM scores to justify its use for this purpose. I will also investigate if there are any patient or operation characteristics that are associated with better PROM scores at 18 months post-surgery. I will concentrate on patient age, sex, ASA score and length of operation.

In this chapter I will first provide background information about PROMs, the scoring properties of PROMs and POMS, length of inpatient stay as an outcome measure, patient factors that may have an association with long-term outcome and the association between length of operating time and outcome. I will then describe the methods used to collect the data for this chapter including patient demographics, operation information, POMS scores and PROM scores. I will describe the statistical methods used to analyse the data.

I will then present the results, first for hip arthroplasty patients and then for knee arthroplasty patients. I will report the association between short-term measures of outcome (POMS and length of hospital stay) and longer-term measures of outcome (PROMs at 18 months post-surgery). I will report the association between patient factors and PROMs, and operation time and PROMs.

Following this I will discuss conclusions that can be drawn from the results. I

will comment on the strengths and limitations of the study and I will end the chapter with a summary of the key points.

The aim of this chapter is to assess whether the POMS can act as an early surrogate marker for patient–reported outcome measures (PROMs). If an earlier surrogate marker for long-term PROMs could be found, procedures could be assessed more quickly ultimately leading to better patient care.

4.1.1 Patient-reported outcome measures

The outcome of orthopaedic procedures was traditionally evaluated by the operating surgeon. Outcomes included examination findings (e.g. joint range of motion), x-ray appearance and walking distance. It became apparent that such evaluations were often influenced by the opinion of the surgeon and did not truly reflect the function of the patient[68]. This lead to the development of patient-reported outcome measures (PROMs). PROMs have multiple uses including monitoring the effectiveness of interventions, for internal audit purposes and in clinical trials to compare different surgical procedures as well as peri-operative interventions. PROMs can be used by patients to guide them in their choice of surgeon and hospital. Health-care providers can use PROMs to guide the allocation of resources.

As described in Chapter 1, PROMs used to assess orthopaedic procedures can be divided into 3 broad categories: generic, disease-specific and joint-specific. Outcome measures in each category evaluate different aspects of a

patient's health and for this study I chose to evaluate one measure from each group: SF-36 (generic), WOMAC (disease-specific) and the Oxford Hip/Knee Score (joint-specific).

4.1.2 Scoring properties of POMS and PROMs

Both PROMs and the POMS are composed of multiple items. PROMs for orthopaedic procedures are reported as an overall score. This score is calculated from the sum, or in some cases the weighted sum, of the item scores. When the POMS was developed it was hypothesised it could measure a single underlying construct (i.e. systemic morbidity associated with surgery). When the internal consistency of the POMS was tested, it was discovered this is not the case[153]. This lack of homogeneity means that the POMS does not have the scaling properties necessary to generate an overall score. The POMS can only be used to define the presence or absence of morbidity following surgery i.e. the POMS provides a dichotomous 'yes/no' result.

4.1.3 Length of stay as a measure of outcome

In the past, length of hospital stay has been used as a surrogate for clinical outcome. Length of stay is accurately recorded by most hospitals so is a readily available piece of data. If length of hospital stay is to be used as a measure of outcome, two discharge criteria must be fulfilled. Firstly, all patients must be discharged at the same level of 'wellness' e.g. when patients are able to walk a certain distance. If one hospital discharges patients who can walk 10 metres and another hospital discharges patients who can walk 25 metres, the

second hospital will have longer inpatient times and ‘worse’ outcome results. This would be unfair and not a true reflection of patient health status. Secondly, all patients must be discharged as soon as they are ‘well’. It is known that length of hospital stay is influenced by many factors other than the health status of the patient[17,18]. In the UK it is common for ‘well’ patients to remain in hospital for reasons such as awaiting placement in a rehabilitation centre or care home.

4.1.4 Patient factors and long-term outcome

Evidence exists that patients with higher ASA scores (ASA 3 or 4) have worse functional outcome following hip and knee arthroplasty than those with lower ASA scores (1 or 2)[154]. There is conflicting evidence whether there is a difference in outcome between ASA grade 1 and 2 patients. A recent analysis of the New Zealand Joint Registry showed a difference in outcome between ASA grade 1 and ASA grade 2 hip arthroplasty patients but no difference in outcome between ASA grade 1 and ASA grade 2 knee arthroplasty patients[154].

There is good evidence that increasing age is associated with decreasing functional outcome scores following hip and knee arthroplasty[155,156]. However, relative scores rather than absolute scores show that older patients generally show a significant improvement in functional outcome following surgery. Thus, the fact that they attain on average a lower level of function than younger patients should not be cited as a reason for refusing surgery.

It has recently been shown that overall patient satisfaction is determined by many factors other than PROMs including meeting pre-operative expectations and adequate pain relief. When looking at overall patient satisfaction, age is not a determining factor[157].

There is conflicting evidence as to whether there is a difference in outcome between males and females following hip arthroplasty[158,159]. It is thought that the modular design of total hip replacements allows for sufficient variation in the implant to accommodate differences in anatomy between the two sexes[160]. There is some evidence that males have a superior functional outcome than females following knee arthroplasty[161]. This led to the development of knee replacements specifically for females. However, most studies show that there is no difference in outcome become gender-neutral and gender-specific replacement in females[162].

4.1.5 Operation time and long-term outcome

It is unknown whether the time taken to perform a arthroplasty procedure has any effect on long-term outcome. There are comparative studies investigating different methods of performing arthroplasty e.g. standard surgery versus minimally invasive surgery[163,164] and standard surgery versus computer-assisted surgery[165,166]. Minimally invasive surgery and computer-assisted surgery generally take longer to perform than standard surgery. However, there are several variables between study groups so outcome cannot simply be attributed to operation time alone. It is unknown if the time taken to perform the

same standardised arthroplasty procedure has any effect on long-term outcome.

In general, less experienced surgeons take longer to perform an operation than more experienced surgeons[167]. It is known that 'high volume' experienced surgeons have better outcomes than 'low volume' less experienced surgeons[167]. Therefore it seems logical that shorter operation times could be associated with better long-term outcome.

Furthermore, arthroplasty for complex joint problems (e.g. significant joint deformity, bone erosion or soft tissue disease) takes longer to perform than standard arthroplasty. Patients with complex joint problems will have lower pre-operative function than patients with simple osteoarthritis undergoing standard arthroplasty. Pre-operative function is known to be associated with post-operative function. This is further reason to suggest that patients with longer operation times may have poorer long-term outcome.

4.2 Methods

The methods used to collect the POMS data are fully described in Chapter 2. POMS data was collected from March 1st 2004 till February 28th 2005 at the Middlesex Hospital. Results were recorded on post-operative days (POD) 3, 5, 8 and 15 if the patient remained in hospital.

Operation time was taken from the anaesthetic record. It was calculated as the time the patient entered the operating room to the time the patient left the operating room.

Patients who underwent TKR and THR in the first 6 months of the POMS study (March 1st 2004 – August 31st 2004) were contacted by post 18 months after discharge requesting them to complete SF-36, WOMAC and Oxford hip/knee questionnaires and return them in pre-paid envelopes. If questionnaires were not returned within 1 month, each patient was contacted by telephone. If they agreed to participate in the study, a second set of questionnaires was posted to them. All surveys and questionnaires were scored according to standard protocols.

Examination of the SF-36, WOMAC and Oxford scores revealed that they were approximately normally distributed. Therefore all analyses were performed using linear regression. Analysis was performed separately for the knee arthroplasty and hip arthroplasty groups.

Analysis was performed in two stages. Firstly the effect of each variable (presence or absence of POMS on POD 3, 5, 8 and 15, length of stay and patient demographics) upon the long-term outcome (SF-36, WOMAC and Oxford scores) was assessed individually in a series of univariable analyses. Secondly, the joint effect of each of the variables upon each outcome was assessed in a multivariable analysis. Due to the fairly small sample sizes,

these analyses were restricted to variables showing some evidence of an association with the outcome in the univariable analyses ($p<0.2$). A backwards selection procedure was used to retain only the statistically significant variables in the final model.

4.3 Results

111 patients underwent primary TKR and 88 patients underwent primary THR during the study period. All patients participated in the initial POMS phase of the study. Of these 199 patients, 123 participated in the second part of the study by completing quality of life questionnaires at 18 months post-surgery. The 76 patients who did not participate in the second part of the study were contacted by post and telephone but both were unsuccessful.

Of the 123 patients who fully participated in the study, 76 underwent total hip replacement and 47 underwent total knee replacement. The average age of the patients undergoing total hip replacement was 64.2. 26 (34%) were male. 17 (22%) were graded as ASA 1, 43 (57%) ASA 2 and 16 (21%) ASA 3. The average length of surgery was 199 minutes. The average age of the patients undergoing total knee replacement was 69.7 and 15 (32%) were male. 10 (21%) were graded as ASA 1, 27 (58%) ASA 2 and 10 (21%) ASA 3. The average length of surgery was 162 minutes.

4.3.1 Hip arthroplasty patients

4.3.1.1 Association between POMS and PROMS

4.3.1.1.1 SF-36

Initially the separate effect of each variable (presence of morbidity on POD 3, 5, 8 and 15) upon this outcome was examined. A summary of results is given in table 31. The first column of figures provides the mean and standard deviation SF-36 score in each category. The next column gives the regression coefficients with corresponding confidence intervals. This coefficient represents the difference in outcome between each category and a reference category. P-values indicating the significance of the results are also reported.

Table 31. Association between the presence of morbidity on post-operative days 3, 5, 8 and 15 and SF-36 scores at 18 months post hip arthroplasty

Variable	Category	Mean (SD)	Coefficient (95% CI)	P-value
POD 3	POMS –ve	47.6 (14.1)	0	0.82
	POMS +ve	44.1 (19.7)	3.5 (-10.2, 12.5)	
POD 5	POMS –ve	48.8 (15.6)	0	0.38
	POMS +ve	44.5 (20.7)	4.3 (-3.0, 13.1)	
POD 8	POMS –ve	50.4 (18.2)	0	0.43
	POMS +ve	44.4 (16.9)	6.0 (-5.8, 14.7)	
POD 15	POMS –ve	48.9 (17.0)	0	0.79
	POMS +ve	46.4 (11.1)	2.5 (-11.4, 12.6)	

These results show that there is no association between the presence of morbidity in the first 15 days following hip arthroplasty and SF-36 scores at 18 months post-surgery.

4.3.1.1.2 WOMAC

The separate effect of each variable (presence of morbidity on POD 3, 5, 8 and 15) upon WOMAC scores was examined. A summary of results is given in table 32.

Table 32. Association between the presence of morbidity on post-operative days 3, 5, 8 and 15 and WOMAC scores at 18 months post hip arthroplasty

Variable	Category	Mean (SD)	Coefficient (95% CI)	P-value
POD 3	POMS –ve	27.1 (21.8)	0	0.53
	POMS +ve	30.6 (21.8)	3.6 (-7.6, 14.7)	
POD 5	POMS –ve	27.3 (20.3)	0	0.42
	POMS +ve	31.8 (23.4)	4.0 (-5.9, 14.0)	
POD 8	POMS –ve	28.1 (22.6)	0	0.33
	POMS +ve	33.6 (19.4)	5.5 (-5.6, 16.6)	
POD 15	POMS –ve	27.5 (20.6)	0	0.007
	POMS +ve	50.6 (22.5)	23.1 (6.7, 39.6)	

These results show that there is no association between the presence of morbidity on POD 3, 5 and 8 and WOMAC scores at 18 months post-surgery.

There is a statistically significant association between the presence of morbidity on POD 15 and poorer WOMAC scores at 18 months post-surgery. On POD 15, hip arthroplasty patients with morbidity had WOMAC scores that were on average 23 units higher than patients without morbidity.

4.3.1.1.3 Oxford Hip Score

The separate effect of each variable (presence of morbidity POD 3, 5, 8 and 15) upon Oxford Hip Scores was examined. A summary of results is given in table 33.

Table 33. Association between the presence of morbidity on post-operative days 3, 5, 8 and 15 and Oxford Hip Scores at 18 months post hip arthroplasty

Variable	Category	Mean (SD)	Coefficient (95% CI)	P-value
POD 3	POMS –ve	24.6 (10.4)	0	0.63
	POMS +ve	26.0 (11.6)	1.4 (-4.4, 7.2)	
POD 5	POMS –ve	24.3 (9.5)	0	0.30
	POMS +ve	27.0 (12.9)	2.7 (-2.4, 7.9)	
POD 8	POMS –ve	24.3 (11.1)	0	0.11
	POMS +ve	28.9 (11.3)	4.6 (-1.1, 10.3)	
POD 15	POMS –ve	24.7 (10.8)	0	0.02
	POMS +ve	34.7 (12.8)	10.0 (1.4, 18.7)	

These results show no association between the presence of morbidity on POD

3, 5 and 8 and Oxford Hip Scores at 18 months post-surgery. There is a statistically significant association between the presence of morbidity on POD 15 and poorer Oxford Hip Scores at 18 months surgery. On POD 15, patients with morbidity had Oxford Hip Scores that were on average 10 units higher than patients without morbidity.

4.3.1.2 Association between length of stay and long-term outcome

The effect of length of stay on SF-36, WOMAC and Oxford Hip scores is shown in Table 34. The regression coefficients are reported for a 5-day increase in length of stay.

Table 34. Association between length of hospital stay and PROMs at 18 months post hip arthroplasty

	Coefficient for a 5-day increase in length of stay (95% CI)	**P-value**
SF-36	2.8 (-1.8, 7.9)	0.12
WOMAC	4.9 (1.7, 8.1)	0.003
Oxford Hip Score	3.2 (1.6, 4.8)	<0.001

These results show there is no association between length of hospital stay and SF-36 scores at 18 months post-surgery. Length of stay is significantly associated with WOMAC scores. A 5-day increase in the length of hospital stay was associated with a 5-unit increase in WOMAC score. There is a highly significant association between length of stay and Oxford Hip Scores. A 5-day

increase in length of stay was associated with a 3-unit increase in Oxford Hip Score.

4.3.1.3 The effect of patient factors on long-term outcome

4.3.1.3.1 Patient Age

The effect of patient age on SF-36, WOMAC and Oxford Hip scores is shown in Table 35. The regression coefficients are reported for a 10-year increase in age.

Table 35. Association between patient age and PROMs at 18 months post hip arthroplasty

	Coefficient (95% CI)	**P-value**
SF-36	1.7 (-1.4, 3.2)	0.29
WOMAC	1.3 (-1.7, 4.0)	0.38
Oxford Hip Score	1.1 (-0.5, 2.7)	0.17

These results show no statistical association between patient age and patient-reported outcome measures at 18 months post-surgery.

4.3.1.3.2 Patient Sex

The effect of sex on SF-36, WOMAC and Oxford Hip scores is shown in Table 36. The coefficient represents the difference in outcome between the two categories (male and female).

Table 36. Association between sex and PROMs at 18 months post hip arthroplasty

Outcome measure	Sex	Mean (SD)	Coefficient (95% CI)	P-value
SF-36	Female	49.1 (18.3)	0	0.64
	Male	46.4 (23.0)	2.7 (-11.8, 16.9)	
WOMAC	Female	24.4 (19.9)	0	0.14
	Male	32.2 (22.3)	7.8 (-2.6, 18.3)	
Oxford Hip Score	Female	23.6 (12.0)	0	0.28
	Male	26.6 (10.9)	3.0 (-2.5, 8.4)	

These results show no statistical association between the sex of a patient and patient-reported outcome measures at 18 months post-surgery.

4.3.1.3.3 ASA grade

The effect of ASA grade on SF-36, WOMAC and Oxford Hip scores is shown in Table 37. The coefficient represents the difference in outcome between each category and the reference category (ASA grade 1).

Table 37. Association between ASA grade and PROMs at 18 months post hip arthroplasty

Outcome	ASA grade	Mean (SD)	Coefficient (95% CI)	P-value
SF-36	1	57.5 (18.1)	0	0.21
	2	49.8 (17.6)	7.7 (-2.2, 17.2)	
	3	46.7 (13.8)	10.6 (0.2, 21.8)	
WOMAC	1	19.5 (16.3)	0	0.05
	2	30.7 (22.5)	11.3 (-0.8, 22.3)	
	3	37.7 (21.8)	18.2 (3.5, 32.9)	
Oxford Hip Score	1	21.5 (8.3)	0	0.14
	2	25.8 (11.5)	4.3 (-2.1, 10.6)	
	3	29.3 (12.5)	7.8 (0.1, 15.5)	

These results show that although there is a tendency towards patients with higher ASA scores having worse long-term SF-36 and Oxford Hip Scores, this is not statistically significant. There is some stronger evidence ($p = 0.05$) that ASA grade is associated with WOMAC scores. ASA grade 3 patients scored 18 units higher than ASA grade 1 patients.

4.3.1.4 Association between length of operation and outcome

The effect of length of operation on SF-36, WOMAC and Oxford Hip scores is shown in Table 38. The regression coefficients are reported for a one-hour increase in length of time of operation.

Table 38. Association between length of operation and PROMs at 18 months post hip arthroplasty

	Coefficient (95% CI)	P-value
SF-36	2.8 (-3.4, 8.0)	0.15
WOMAC	4.0 (-1.0, 8.9)	0.11
Oxford Hip Score	3.0 (0.5, 5.5)	0.02

These results show that although there is a tendency towards patients with longer operations having worse long-term SF-36 and WOMAC Scores, this is not statistically significant. There is stronger evidence (p = 0.02) that a longer operation time is associated with worse Oxford Hip scores. A one-hour increase in operation time is associated with a 3-unit increase in Oxford Hip Score.

4.3.1.5 Multivariable analysis of results

4.3.1.5.1 SF-36

None of the studied variables (POMS scores, length of patient stay, patient factors and length of operation) had any association with SF-36 scores so no multivariable analysis was performed.

4.3.1.5.2 WOMAC

The multivariable analysis considers variables showing some association with outcome. The presence of POMS morbidity on POD 15, ASA grade and length of stay were associated with WOMAC scores using univariable analysis.

On multivariable analysis, only length of stay was significantly associated with WOMAC scores at 18 months post-surgery. There was no additional effect of POMS score on POD 15 or ASA grade, having adjusted for this variable. As this variable was the only variable in the final model, the size effects were equivalent to those found in the univariable analysis.

4.3.1.5.3 Oxford Hip Score

A multivariable analysis again suggested that only post-operative length of stay was statistically significant. There were no additionally significant variables after accounting for this factor.

4.3.2 Knee arthroplasty patients

4.3.2.1 Association between POMS and PROMS

4.3.2.1.1 SF-36

Initially the separate effect of each variable (presence of morbidity POD 3, 5, 8 and 15) upon this outcome was examined. A summary of results is given in table 39. The coefficient represents the difference in outcome between each category and a reference category.

Table 39. Association between the presence of morbidity on post-operative days 3, 5, 8 and 15 and SF-36 scores at 18 months post knee arthroplasty

Variable	Category	Mean (SD)	Coefficient (95% CI)	P-value
POD 3	POMS –ve	54.3 (17.2)	0	0.77
	POMS +ve	50.1 (23.7)	3.2 (-15.5, 22.0)	
POD 5	POMS –ve	52.9 (20.5)	0	0.59
	POMS +ve	48.1 (18.3)	4.8 (-14.9, 23.9)	
POD 8	POMS –ve	52.0 (18.7)	0	0.49
	POMS +ve	49.1 (15.8)	2.9 (-9.9, 15.4)	
POD 15	POMS –ve	49.9 (24.0)	0	0.83
	POMS +ve	46.4 (8.7)	3.5 (-19.6, 26.8)	

These results show no association between the presence of morbidity in the first 15 days following knee arthroplasty and SF-36 scores 18 months after surgery.

4.3.2.1.2 WOMAC

The separate effect of each variable (presence of morbidity on POD 3, 5, 8 and 15) upon WOMAC scores was examined. A summary of results is given in table 40.

Table 40. Association between the presence of morbidity on post-operative days 3, 5, 8 and 15 and WOMAC scores at 18 months post knee arthroplasty

Variable	Category	Mean (SD)	Coefficient (95% CI)	P-value
POD 3	POMS –ve	23.3 (15.8)	0	0.06
	POMS +ve	33.8 (21.0)	10.6 (-0.3, 21.5)	
POD 5	POMS –ve	25.2 (16.5)	0	0.11
	POMS +ve	34.7 (22.6)	9.5 (-2.1, 21.1)	
POD 8	POMS –ve	28.5 (18.9)	0	0.96
	POMS +ve	28.1 (21.5)	-0.3 (-15.4, 14.7)	
POD 15	POMS –ve	28.2 (19.5)	0	0.67
	POMS +ve	34.0 (5.7)	5.8 (-22.2, 33.9)	

These results show no association between the presence of morbidity on POD 3, 5, 8 and 15 and WOMAC scores at 18 months post-surgery. There is a tendency towards an association between the presence of morbidity on POD 3 and POD 5 and poorer WOMAC scores at 18 months surgery but this is not statistically significant.

4.3.2.1.3 Oxford Knee Score

The separate effect of each variable (presence of morbidity on POD 3, 5, 8 and 15) upon Oxford Knee Scores was examined. A summary of results is given in table 41.

Table 41. Association between the presence of morbidity on post-operative days 3, 5, 8 and 15 and Oxford Knee Scores at 18 months post knee arthroplasty

Variable	Category	Mean (SD)	Coefficient (95% CI)	P-value
POD 3	POMS –ve	25.5 (10.2)	0	0.34
	POMS +ve	28.3 (10.0)	2.8 (-3.1, 8.8)	
POD 5	POMS –ve	25.9 (10.1)	0	0.39
	POMS +ve	28.6 (10.3)	2.7 (-3.6, 9.0)	
POD 8	POMS –ve	27.0 (10.3)	0	0.80
	POMS +ve	26.0 (9.3)	-1.1 (-9.0, 7.0)	
POD 15	POMS –ve	26.6 (10.2)	0	0.42
	POMS +ve	32.5 (3.5)	5.9 (-8.8, 20.6)	

These results show no association between the presence of morbidity on POD 3, 5, 8 and 15 and Oxford Knee Scores at 18 months post-surgery.

4.3.2.2 Association between length of stay and long-term outcome

The effect of length of stay on SF-36, WOMAC and Oxford Knee scores is shown in Table 42. The regression coefficients are reported for a 5-day increase in length of stay.

Table 42. Association between length of hospital stay and PROMs at 18 months post knee arthroplasty

	Coefficient (95% CI)	P-value
SF-36	4.1 (-2.2, 10.3)	0.20
WOMAC	9.6 (0.7, 18.5)	0.04
Oxford Knee Score	3.6 (-1.2, 8.4)	0.14

These results show no association between length of hospital stay and SF-36 and Oxford Knee Scores at 18 months post knee arthroplasty. Length of stay is significantly associated with WOMAC scores. A 5-day increase in hospital stay was associated with a 10-unit increase in the WOMAC score.

4.3.2.3 Effect of patient factors on long-term outcome

4.3.2.3.1 Patient Age

The effect of patient age on SF-36, WOMAC and Oxford Knee scores is shown in Table 43. The regression coefficients are reported for a 10-year increase in age.

Table 43. Association between patient age and PROMs at 18 months post knee arthroplasty

	Coefficient (95% CI)	P-value
SF-36	0.8 (-4.8, 6.7)	0.68
WOMAC	-1.9 (-6.4, 2.6)	0.40
Oxford Knee Score	-1.0 (-3.4, 1.4)	0.39

These results show no statistical association between patient age and patient-reported outcome measures at 18 months post-surgery.

4.3.2.3.2 Patient Sex

The effect of sex on SF-36, WOMAC and Oxford Knee scores is shown in Table 44. The coefficient represents the difference in outcome between the two categories (male and female).

Table 44. Association between sex and PROMs at 18 months post knee arthroplasty

Outcome measure	Sex	Mean (SD)	Coefficient (95% CI)	P-value
SF-36	Female	51.7 (19.8)	0	0.77
	Male	49.7 (21.5)	2.0 (-14.1, 16.4)	
WOMAC	Female	27.8 (20.0)	0	0.89
	Male	28.7 (19.0)	0.9 (-11.3, 13.0)	
Oxford Knee Score	Female	27.7 (12.1)	0	0.68
	Male	26.4 (9.2)	-1.3 (-7.7, 5.1)	

These results show no statistical association between sex and patient-reported outcome measures at 18 months post knee arthroplasty.

4.3.2.3.3 ASA grade

The effect of ASA grade on SF-36, WOMAC and Oxford Knee scores is shown in Table 45. The coefficient represents the difference in outcome between each category and the reference category (ASA grade 1).

Table 45. Association between ASA grade and PROMs at 18 months post knee arthroplasty

Outcome	ASA grade	Mean (SD)	Coefficient (95% CI)	P-value
SF-36	1	54.3 (16.2)	0	0.32
	2	51.0 (18.1)	3.3 (-7.9, 14.6)	
	3	46.6 (13.8)	7.7 (-5.0, 20.9)	
WOMAC	1	24.3 (11.8)	0	0.42
	2	27.4 (20.4)	3.1 (-11.2, 17.4)	
	3	35.2 (21.3)	10.9 (-6.4, 28.2)	
Oxford Knee Score	1	23.2 (6.1)	0	0.05
	2	25.8 (10.2)	2.6 (-4.6, 9.7)	
	3	33.4 (10.8)	10.2 (1.5, 18.9)	

These results show there is no statistical association between ASA grade and SF-36 and WOMAC scores following knee arthroplasty. There is some evidence of an association between ASA grade and Oxford Knee Score although this is of borderline statistical significance ($p = 0.05$). There was a relatively small difference between patients with an ASA grade of 1 and 2. However, ASA grade 3 patients had an average Oxford Knee Score ten units higher than ASA grade 1 patients.

4.3.2.4 Association between length of operation and outcome

The effect of length of operation on SF-36, WOMAC and Oxford Knee scores is shown in Table 46. The regression coefficients are reported for a one-hour

increase in length of operation.

Table 46. Association between length of operation and PROMs at 18 months post knee arthroplasty

	Coefficient (95% CI)	P-value
SF-36	2.6 (-7.9, 12.8)	0.65
WOMAC	1.8 (-10.9, 14.4)	0.78
Oxford Knee Score	0.3 (-6.4, 7.0)	0.93

These results show no statistical association between length of knee surgery and PROMs.

4.3.2.5 Multivariable analysis of results

4.3.2.5.1 SF-36

None of the studied variables (POMS scores, length of patient stay, patient factors and length of operation) had any association with SF-36 scores so no multivariable analysis was performed.

4.3.2.5.2 WOMAC

A multivariable analysis was performed using factors showing some association with the WOMAC scores on univariable analysis. These variables included POMS score on POD 3 and 5, and length of stay. Multivariable analysis suggested that only post-operative length of stay was significantly associated with WOMAC scores at 18 months post-surgery. There was no additional effect

of the POMS scores on POD 3 and 5 once length of stay had been adjusted for. As length of hospital stay was the only variable in the final model, the size effects are equivalent to those found in the univariable analysis.

4.3.1.5.3 Oxford Knee Score

A multivariable analysis suggested that only ASA grade was significantly associated with the Oxford Knee Score at 18 months post-surgery. There were no additional significant variables after accounting for this factor.

4.4 Discussion

4.3.1 Summary

This chapter investigated the association between morbidity on POD 3, 5, 8 and 15 and PROMs (SF-36, WOMAC and Oxford Hip/Knee Scores) at 18 months post surgery. For hip arthroplasty patients, univariable analysis indicated an association between morbidity on POD 15 and WOMAC and Oxford Hip Scores. However, multivariable analysis did not support this. For knee arthroplasty patients, no association was found between post-operative morbidity and PROMs. Since POMS has no association with longer-term PROMs, POMS should not be used as an early surrogate marker of surgical outcome.

The relationship between length of hospital stay and PROMs was also investigated. In the hip arthroplasty group, univariable analysis revealed an

association in length of stay and WOMAC and Oxford Hip Scores. In the knee arthroplasty group, univariable analysis revealed an association between length of stay and WOMAC scores. Multivariable analysis confirmed these associations.

The relationship between patient factors (age, sex, ASA) and PROMs was assessed. There was no association between age and PROMs, or sex and PROMs. Univariable analysis revealed an association between ASA score and WOMAC scores in the hip arthroplasty group. Multivariable analysis confirmed this was not statistically significant with other variables taken into account. Univariable analysis revealed an association between ASA score and Oxford Knee Score in the knee arthroplasty group. Multivariable analysis confirmed this was statistically significant.

The relationship between operating time and PROMs was assessed. In the hip arthroplasty group univariable analysis revealed an association between operating time and Oxford Hip Scores. Multivariable analysis confirmed this was not statistically significant when other variables were taken into account. In the knee arthroplasty group no association was found between operation length and PROMs.

4.3.2 Strengths of study

This study involved collecting a large data set. A dedicated research team recorded all POMS data, patient data, length of stay and length of operation.

The data was accurate and complete. A second dedicated research team collected the patient-reported outcome scores. Again, this ensured that full and accurate data sets were obtained.

4.3.3 Limitations of study

The main limitation of this study is the small sample size. This may have lead to false negative results. If a larger sample size had been used, more significant associations between post-operative morbidity and long-term patient-reported outcome may have been found. This is an area for future research. A further limitation of the study is the fact it was performed at a single centre. Thus the results may not be transferable to other centres.

4.3.4 Comparison to other studies

The association between the presence of any type of post-operative morbidity and longer term PROMs has not previously been investigated. Thus it is not possible directly to compare this part of the study to previous studies.

Patients with post-operative morbidity could be expected to recover less well than patients with no post-operative morbidity, and thus have poorer long-term PROM scores. This was not shown to be the case in this study. In the hip arthroplasty group, univariable analysis showed an association between morbidity on POD 15 and poorer WOMAC and Oxford Hip Scores. This was statistically insignificant on multivariable analysis. If a larger population had been used, an association between morbidity on POD 15 and longer-term

PROMs may have been found, even using multivariable analysis. This requires further investigation.

Many studies use length of stay as an early surrogate marker of outcome. Our study confirms that this is appropriate. In the hip arthroplasty group, an association was found between length of stay and WOMAC and Oxford Hip Scores. In the knee arthroplasty group, an association was found between length of stay and WOMAC scores.

Our study did not find any association between patient age and PROMs. In the knee arthroplasty group, there was a statistically insignificant negative correlation between patient age and PROMs. Most studies report worse outcome with increasing age[155,156]. Based on our results, arthroplasty surgery provides equal long-term benefit to patients of all ages.

The orthopaedic literature is mixed regarding whether sex is associated with outcome following arthroplasty[158,159]. Our study does not support any difference between the sexes. Several studies report poorer outcome in patients with higher ASA grades[154]. Our study generally supports this. A statistically significant association was found between ASA grade and WOMAC scores on univariable analysis in the hip arthroplasty group. The association was not found to be statistically significant on multivariable analysis but this may be due to the small sample size. A statistically significant association was found

between ASA grade and Oxford Knee Scores on both univariable and multivariable analysis in the knee arthroplasty group.

There are no previous published studies regarding the association between length of arthroplasty surgery and PROMs. In the hip arthroplasty group, univariable analysis found an association between operating time and Oxford Hip Scores, with longer operating time resulting in worse Oxford Hip Scores. This association was statistically insignificant on multivariable analysis but this may be due to a small sample size. Further investigation using a larger study population may prove a true association.

4.5 Summary

1. There is no association between post-operative morbidity defined by the POMS and long-term PROMs. This study does not support the use of POMS as an early surrogate marker of long-term patient outcome.
2. There was a tendency towards patients with morbidity on POD 15 having poorer PROMs but this was not statistically significant.
3. Length of hospital stay was associated with PROMs in both the hip and knee arthroplasty groups. This justifies the use of length of hospital stay as an early marker of outcome.
4. There was no association between patient age and PROMs.
5. There was no association between patient sex and PROMs.
6. Higher ASA grade was associated with worse PROMs in the knee

arthroplasty group. The association was not statistically significant in the hip group.

7. There is no association between length of operation and PROMs.

Chapter 5 How reliable is the 'wound' item in the POMS

5.1 Introduction

Surgical site infection (SSI) following arthroplasty is a potentially devastating complication. It can require revision surgery, and in the worst-case scenario, can lead to limb amputation. To avoid these complications, every effort is made to keep infection rates to a minimum. In order to monitor infection rates, a reliable and reproducible method of diagnosing infection must be used.

This chapter will examine the accuracy of the wound domain of the POMS. The POMS definition of wound morbidity is 'wound dehiscence requiring surgical exploration or drainage of pus from operative wound with or without isolation of organisms'. The POMS aims to identify morbidity that warrants inpatient hospital care. Therefore, the wound domain of the POMS would be expected to identify SSI that requires intravenous antibiotics or surgical treatment. The POMS would not be expected to identify mild superficial wound infections that can be treated with oral antibiotics and outpatient monitoring.

In order to test accuracy, the wound domain of POMS needs to be compared to a 'gold standard'. There are several definitions of wound infection, as discussed in chapter one. These include American Centres for Disease Control

(CDC) definition, the English Nosocomial Infection National Surveillance Scheme (NINSS) definition and the English ASEPSIS definition (Additional treatment, Serous discharge, Erythema, Purulent exudate, Separation of deep tissues, Isolation of bacteria and Stay as inpatient prolonged over fourteen days), all of which have undergone some psychometric analysis. Substantial gaps exist in the literature regarding the validity of these definitions. The CDC definition is known to be weak since 3 out of the 4 defining criteria are subjective. CDC is also known to be unreliable[35]. NINSS evolved in an attempt to increase the reliability of CDC but evaluation has confirmed that reproducibility remains low[36]. The original ASEPSIS method has been shown to be both objective and repeatable[38] for sternal wounds but a more recent revised version has not been psychometrically evaluated. A recent systematic review concluded that ASEPSIS is the 'gold standard' for scoring surgical site infections[168]. It was therefore decided to compare the wound domain of POMS to ASEPSIS to test its accuracy.

Patient follow-up is essential to record accurately SSI rates since half of infections present after hospital discharge[33]. Therefore SSI rates cannot be simply defined as a 'short-term' outcome measure. ASEPSIS defines wound infection as occurring up to 2 months post-surgery. This presents a problem when comparing it to the wound domain of the POMS. The POMS only identifies wound morbidity during the inpatient episode, where as the ASEPSIS considers both the inpatient episode and outpatient follow-up. Furthermore, the POMS only identifies wound morbidity that is sufficiently serious to warrant

inpatient care. ASEPSIS identifies mild wound morbidity that could be treated as an outpatient as well more serious morbidity requiring inpatient care. For these two reasons, the incidence of wound morbidity defined by the POMS could be expected to be lower than the incidence defined by ASEPSIS.

In an attempt to overcome this problem, and make a fair comparison, the wound domain of the POMS was compared to two different ASEPSIS scores: the inpatient ASEPSIS score (based on information from the inpatient episode only) and the overall ASEPSIS score. In this way, the wound domain of POMS could be directly compared to another inpatient assessment of wound morbidity.

At the time of the POMS study, ASEPSIS scores were routinely collected on all patients who remained in the hospital for at least 2 days. Therefore ASEPSIS data was available for all patients in the POMS study.

5.2 Methods

5.2.1 General methodology

The methods used to collect the POMS data are described fully in Chapter 2. The patient demographics are also described. POMS data was collected on post-operative days (POD) 3, 5, 8 and 15 if the patient remained in hospital.

ASEPSIS data was collected on the same patients. A member of a specialist wound surveillance team, made up of 4 nurses and a health care assistant,

assessed each wound. The sole role of these five members of staff was to collect and record wound infection data, and all received specialist training in the diagnosis of surgical site infection.

Each patient was reviewed on 3 separate occasions: once pre-operatively and twice post-operatively. The same standardized data collection sheet was completed for each patient. Details collected pre-operatively included patient age, height, weight, clinical team and consultant in charge. Operative information and microbiology results were recorded from a direct interface with hospital computer databases. Microbiology tests, such as wound swabs or tissue cultures, were performed according to clinical judgment. No specific microbiology tests were requested for study purposes alone.

Surgical wounds were inspected on POD 2 or 3, and again on POD 4 or 5 (if the patient remained in hospital). The proportion of each wound exhibiting erythema, serous discharge, purulent discharge or dehiscence was recorded. Wounds were directly inspected by surveillance staff if undressed, but if a dressing was present, the relevant information was gained by questioning nursing staff. This was done to avoid an unnecessary increase in the risk of infection. Nurses were encouraged to fill out the data collection sheet at the time of dressing change.

At each post-operative visit, patient notes and drug charts were inspected. The

prescription of therapeutic antibiotics and the opening of a wound or drainage of an abscess were recorded.

At the time of discharge patients were given a simple ‘yes/no’ questionnaire regarding their wound. They were asked to complete and return the questionnaire in a pre-paid envelope 2 months post-surgery. Patients were contacted by telephone if no postal questionnaire was returned. The questionnaire was used to ascertain if a wound infection had been diagnosed since discharge, if antibiotics had been prescribed for the wound, if any further surgery had been necessary, and if the hospital stay had been longer than 14 days.

Data was stored on a modified Access 97® database which was only accessible to surveillance team members. A single patient episode was defined as an operation with post-operative follow-up of either 3 months or until a further operation was performed, whichever was shorter. SSI resulting in readmission at any time was recorded in the database.

5.2.2 Calculation of the ASEPSIS score

ASEPSIS is a quantitative wound scoring method. The score is calculated using objective criteria based both on visual characteristics of the wound and the consequences of infection.

The wound is examined on 2 separate occasions post-operatively and given a score of between 0 and 30. This score is calculated according to the proportion of the wound affected by serous exudate, erythema, purulent exudate and separation of deep tissues. The greater the proportion of wound affected, the higher the score. There is a maximum score of 5 points for serous exudate and erythema, and a maximum score of 10 points for purulent exudate and separation of deep tissues. A detailed breakdown of the point scale for the ASEPSIS wound inspection score is provided in Table 5, Page 31, Chapter 1. Only the higher of the 2 scores contributes to the overall ASEPSIS score.

The second part of the ASEPSIS score is derived from data concerning the consequences of infection. 10 points are given for the prescription of antibiotics for the wound, 5 points for drainage of pus under local anaesthetic, 10 points for the drainage of pus under general anaesthetic, 10 points for the isolation of bacteria and 5 points for a stay in hospital over 14 days. Full details of this points scale are provided in Chapter 1, Page 32, Table 6. This part of the ASEPSIS score is calculated partly from inpatient data and partly from questionnaire data collected 2 months post-surgery.

To calculate a final ASEPSIS score, the higher of the two wound scores (calculated from data gathered during the inpatient stay) is added to the score gained from the consequences of infection (calculated from data gathered during the inpatient stay together with data from the questionnaire completed 2 months post-surgery).

5.2.3 Data analysis

In order to calculate the accuracy of the wound domain of the POMS, it was compared to the inpatient ASEPSIS score and the overall ASEPSIS score. The sensitivity, specificity, positive predictive value, negative predictive value and accuracy of the wound domain of POMS were calculated relative to both the inpatient ASEPSIS score and the overall ASEPSIS score.

5.3 Results

The same 529 patients described in Chapter 2 were included in this part of the study. Table 47 shows SSI rates as defined by the wound domain of the POMS and the inpatient ASEPSIS score. Table 48 shows SSI rates as defined by the wound domain of the POMS and the total ASEPSIS score. Out of 529 patients, 497 (94%) completed the ASEPSIS questionnaire 2 months following surgery. For the 32 patients who could not be contacted, only inpatient data was available to calculate the second part of the ASEPSIS score (evaluating the clinical consequences of infection). Therefore, the inpatient AEPSIS score and overall ASEPSIS score was the same for these 32 patients and the true rate of infection may be under-estimated in the overall ASEPSIS score.

Table 47. Presence of wound infection according to the wound domain of POMS and the inpatient ASEPSIS score

		Wound infection according to POMS		
		Yes	**No**	**Total**
Wound infection according to inpatient ASEPSIS score	**Yes**	10	2	12
	No	49	468	517
	Total	59	470	529

Table 48. Presence of wound infection according to the wound domain of POMS and the total ASEPSIS score

		Wound infection according to POMS		
		Yes	**No**	**Total**
Wound infection according to total ASEPSIS score	**Yes**	11	8	19
	No	48	462	510
	Total	59	470	529

The sensitivity, specificity, positive predictive value, negative predictive value and overall accuracy of the wound domain of the POMS were calculated and compared to (i) the inpatient ASEPSIS score and (ii) the total ASEPSIS score. A summary of the estimated values is given in table 49 and 50 respectively.

Table 49. Characteristics of the wound domain of POMS compared to the inpatient ASEPSIS score

Statistic	Number	Estimate (95% CI)
Sensitivity	10/12	0.83 (0.52, 0.98)
Specificity	468/517	0.91 (0.88, 0.93)
Positive predictive value	10/59	0.17 (0.08, 0.29)
Negative predictive value	468/470	0.996 (0.985, 0.999)
Overall accuracy	478/529	0.90 (0.88, 0.93)

Table 50. Characteristics of the wound domain of POMS compared to the total ASEPSIS score

Statistic	Number	Estimate (95% CI)
Sensitivity	11/19	0.58 (0.33, 0.80)
Specificity	462/510	0.91 (0.88, 0.93)
Positive predictive value	11/59	0.18 (0.09, 0.30)
Negative predictive value	462/470	0.98 (0.97, 0.99)
Overall accuracy	473/529	0.89 (0.86, 0.92)

These results show the wound domain of POMS is reasonably sensitive compared to the inpatient ASEPSIS score with a value of 0.83. Sensitivity of the wound domain of POMS is lower (0.58) when compared to the overall ASEPSIS score. Specificity of the wound domain of POMS is high (0.91) when compared to both the inpatient and overall ASEPSIS scores.

There is a low positive predictive value for the wound domain of the POMS compared to both the inpatient ASEPSIS score (0.17) and the total ASEPSIS score (0.18). This suggests that less than one-fifth of patients with a positive result according to POMS will have a positive result according to either the inpatient or total score of ASEPSIS. This stems from the wound domain of POMS over-predicting the occurrence of a positive result. The negative predictive value is very high with almost all patients with a negative POMS result also having a negative ASEPSIS result.

The overall accuracy of the wound domain of POMS is about 90% when compared to both the inpatient ASEPSIS and total ASEPSIS scores. This high value is partially due to the majority of patients being negative.

5.4 Discussion

5.4.1 General conclusions

The wound domain of the POMS only assesses wounds during the inpatient stay. For this reason, comparisons between the wound domain of POMS and the inpatient ASEPSIS score are more meaningful than comparisons with the total ASEPSIS score. It is known that up to half of wound infections present after hospital discharge[33]. Thus, the wound domain of POMS is not appropriate for monitoring SSI rates. However, satisfactory wound healing is an important factor in determining whether a patient is ready for hospital discharge. It is

important that the wound domain of POMS is highly sensitive and is able to recognise wound infections that would be identified by an established definition of SSI. It is important that wound infections receive appropriate treatment to avoid potentially catastrophic consequences.

When compared to the inpatient ASEPSIS score, the wound domain of POMS has a reasonable sensitivity (0.83), good specificity (0.91), a high negative predictive value (0.996) but a poor positive predictive value (0.17). Overall accuracy is 0.90. In other words, the wound domain of POMS is reasonably good at identifying infected wounds as defined by ASEPSIS, but over four-fifths of the wounds identified as having morbidity by POMS are not infected according to ASEPSIS.

The low positive predictive value of the wound domain of POMS in comparison to the inpatient ASEPSIS score is surprising considering the two definitions. The wound domain of POMS is designed to identify severe wound problems (either a dehisced wound requiring surgery or a wound draining pus). ASEPSIS is designed to identify mild, moderate and severe wound infections. Therefore, it follows that ASEPSIS would be expected to identify more wound morbidity than the wound domain of POMS, since ASEPSIS would identify mild wound infections that the wound domain of POMS would miss. The study shows the opposite is true.

The definition of wound morbidity according to POMS includes 'wound dehiscence requiring surgical exploration...with or without isolation of organisms'. The wording of this definition could partially explain the low positive predictive value of POMS compared to the inpatient ASEPSIS score. Most wound dehiscence is due to infection. However, a wound can dehisce for other reasons such as trauma, haematoma formation or poor surgical technique. Non-infected dehiscence can be repaired on the ward using steristrips or interrupted sutures under local anaesthetic. Thus, these cases would be positive for wound morbidity defined by POMS but not ASEPSIS.

The personnel used for data collection could also explain the difference in wound morbidity according to the POMS and ASEPSIS definitions. ASEPSIS data was collected by study personnel with specific training in the diagnosis of wound infection. POMS data was collected by study personnel who did not receive training on the evaluation of surgical wounds. Thus the POMS data may be less reliable.

The 2 criteria used to diagnose POMS wound morbidity are very similar to 2 of the 4 criteria used to diagnose SSI according to the CDC definition (Table 4, Page 31, Chapter 1). One criterion in the CDC definition is 'purulent discharge from the incision'. This is the same as one of the POMS criteria. Another CDC criterion is 'spontaneous dehiscence or deliberate opening of a deep incision, following fever or pain or tenderness around the wound (unless cultures are negative)'. This is similar to the other POMS criterion ('wound dehiscence

requiring surgical exploration'). CDC has been shown to be unreliable[35]. Thus, the wound domain of POMS may be similarly unreliable and this may partially explain the different results between the wound domain of POMS and the wound score of ASEPSIS.

The wound domain of POMS identifies over 80% of wounds classified as infected by the inpatient ASEPSIS score. This leaves almost 20% of infected wounds undetected by the wound domain of POMS. This is not necessarily a problem since not all wound infections require inpatient care. As long as patients receive appropriate outpatient follow-up and are given information regarding wound care and wound infection, it is safe to discharge patients with minor wound problems. There is no method of diagnosing wound infection that is 100% sensitive and almost half of SSI presents after hospital discharge. Therefore, it is very important that surgical wounds are inspected at least once after hospital discharge. At present, a district nurse or GP practice nurse usually does this between 10 and 14 days after surgery. The wound is usually inspected again by a surgeon 6 weeks post surgery in the outpatient clinic.

The wound domain of POMS over-diagnoses wound infection compared to the wound score of ASEPSIS. If the POMS is used to identify morbidity that requires inpatient care, patients could be kept in hospital unnecessarily due to the mis-diagnosis of wound morbidity. This would result in higher costs for the healthcare provider and greater exposure of patients to hospital-acquired infections.

It is now mandatory in the UK for all hospitals to collect SSI data. No definition of SSI has been fully validated but some psychometric analysis is available for ASEPSIS. It may be more accurate to use the inpatient ASEPSIS score to identify POMS wound morbidity than the definition that exists at present. This data would be readily available if ASEPSIS is used for wound surveillance purposes.

5.4.2 Strengths of study

This study has several strengths. It was performed prospectively and contains a large study population. Standardised data collection sheets were completed for both the POMS and ASEPSIS studies. This ensured complete data sets.

5.4.3 Weaknesses of study

One weakness of this study is that the staff collecting data for the wound domain of POMS did not receive specific training with regards the examination of surgical wounds and the diagnosis of wound morbidity. A second weakness is that 32/529 patients could not be contacted to complete the second part of the ASEPSIS scoring system. Therefore, the overall ASEPSIS score may be under-reported for these patients. However, this study mainly looked at agreement between the wound domain of POMS and the inpatient ASEPSIS score. Therefore, the fact that a proportion of patients did not complete the ASEPSIS questionnaire 2 months after surgery will not affect the comparison of inpatient diagnoses of wound morbidity.

5.4.4 Comparisons with other studies

There are no previous studies examining the accuracy of the wound domain of the POMS so no comparisons can be made.

5.5 Summary

1. The wound domain of POMS has a high specificity and reasonable sensitivity when compared to the inpatient ASEPSIS score.
2. The wound domain of POMS has a high negative predictive value but a low positive predictive value when compared to the inpatient ASEPSIS score.
3. The wording of the definition of POMS wound morbidity may account in part for the low positive predictive value.
4. The wound domain of the POMS could be replaced by an established method of diagnosing SSI. It is mandatory for all hospitals in the UK to collect SSI data so this information is readily available.

Chapter 6: How should surgical site infection be measured?

6.1 Introduction

'High Quality Care For All[7]' is a publication from The Department of Health in the UK. It states that all healthcare providers working for or on behalf of the NHS must publish 'Quality Accounts'. These 'accounts' cover three aspects of patient care: safety, patient experience and patient outcome. As part of the safety aspect, surgical site infection prevalence data must be collected and is readily available for the general public to review on the 'NHS Choices' website.

In the 'High Quality Care For All' report, seven steps were described to improve quality:

1) Set standards of high quality
2) Measure quality
3) Publish quality performance
4) Recognise and reward quality
5) Raise standards
6) Safeguard quality
7) Stay ahead by encouraging innovation

The report stated that surgical site infection (SSI) data will be available on the NHS Choices website (step 3) but did not state what a 'good' SSI rate is (step1)

nor did the report describe how SSI rates will be measured (step 2). There is a common misconception that SSIs are easy to define and diagnose. There are several definitions of SSI and diagnosis of infection varies between surgeons.

SSIs were traditionally diagnosed using the examination findings of pain, redness, heat, swelling and impairment of function. More reliable and reproducible methods of diagnosing SSI are now available. Three SSI definitions commonly used today are the US Centers for Disease Control (CDC) definition, the English Nosocomial Infection National Surveillance Scheme (NINSS) definition and the English ASEPSIS definition. These definitions are described in detail in Chapter 1.

In Chapter 5, I looked at the accuracy of the wound domain of POMS compared to ASEPSIS. This led me to consider how reliable and comparable other methods of diagnosing wound infection are. The POMS measures morbidity during the inpatient stay only. A significant proportion of SSIs present following hospital discharge and the commonly used definitions of SSI take account of this. The CDC and NINSS definitions define SSI as occurring within 1 year of surgery if an implant is present. ASEPSIS defines SSI as occurring within 2 months of surgery.

In the last chapter, I compared the wound domain of POMS to inpatient ASEPSIS data to ensure that a fair and meaningful comparison was made. A

comparison of the wound domain of POMS to overall ASEPSIS scores (which is based on inpatient and out-patient data) would not be meaningful.

For this chapter I will use data collected from inpatient and outpatient episodes. Approximately half of SSIs presents after hospital discharge so it important that this information is captured when reporting overall SSI rates. If outpatient data was not included, infection rates would be under-reported.

The purpose of this chapter is to assess if three commonly used methods of diagnosing SSI (CDC, NINSS and ASEPSIS) report similar rates of infection in the same series of patients. SSI rates are used as a performance indicator and are reported for every NHS hospital in England and Wales on the NHS Choices website. Different institutions use different methods to assess SSI. If published SSI rates are to be meaningful, different definitions of SSI must give similar values. This ensures that comparisons between surgeons and hospitals are fair and made against the same benchmark. Alternatively, the same diagnostic method must be adopted by all.

6.2 Methods

6.2.1 General methodology

A wound surveillance program was started in the Department of Trauma and Orthopaedics at University College London Hospital in May 2000. For the first two years of the program, due to funding constraints, data was only

collected for 6 months of each year, from May till October. This represented 35% of total Orthopaedic admissions. From 2002 onwards, data collection became continuous. This study is part of a hospital audit programme so ethics committee approval was not required.

Criteria for inclusion in the wound surveillance program included all trauma and elective orthopaedic patients with a minimum 2-night stay in hospital and an operation involving the incision of tissue. Traumatic wounds were not included in the study, only incisions made at the time of surgery.

Each patient was reviewed on 3 separate occasions: once pre-operatively and twice post-operatively. Each review was performed by a member of a specialist wound surveillance team. The same standardized data collection sheet was completed for each patient. The methods used to collect the wound data are described fully in Chapter 5. In addition to this, the diagnosis of a wound infection by a medical practitioner was noted.

Data was stored on a modified Access 97® database which was only accessible to surveillance team members. A single patient episode was defined as an operation with post-operative follow-up of either 3 months or until a further operation was performed, whichever was shorter. At any time point, SSI resulting in readmission was recorded in the database. Sufficient information was gathered to allow each wound to be diagnosed according to the CDC, NINSS and ASEPSIS definitions of infection.

6.2.2 Calculation of SSI rates

The ASEPSIS definition of SSI gives every wound a score. The way in which this score is calculated is described in detail in Chapter 5. A score of 21 or over indicates SSI. Both CDC and NINSS have certain criteria that must be fulfilled to diagnose SSI. There is no score, just a simple dichotomous yes/no result.

6.2.3 Statistical analysis

Crude infection rates were calculated for each definition of infection and 95% confidence intervals calculated. The agreement between crude infection rates was calculated using the kappa statistic. Kendall's tau b value was used to assess correlation between the subdivisions of each definition.

6.3 Results

7448 orthopaedic wounds in 7299 patients were assessed between May 2000 and October 2008. The follow-up rate of patients two months following surgery was 91%. Details of patient demographics are shown in table 51. The distribution of age and ASA grade within the study population is shown in greater detail in figure 5 and figure 6 respectively.

Table 51. Patient demographics of 7299 trauma and orthopaedic patients

Gender M:F	 44:56
Mean age in years (range)	56.4 (0 – 99)
ASA I II III IV	2109 (28.3%) 3639 (48.9%) 1595 (21.4%) 105 (1.4%)
Type of surgery Elective : Emergency	 86:14

Figure 5. Age distribution of study population

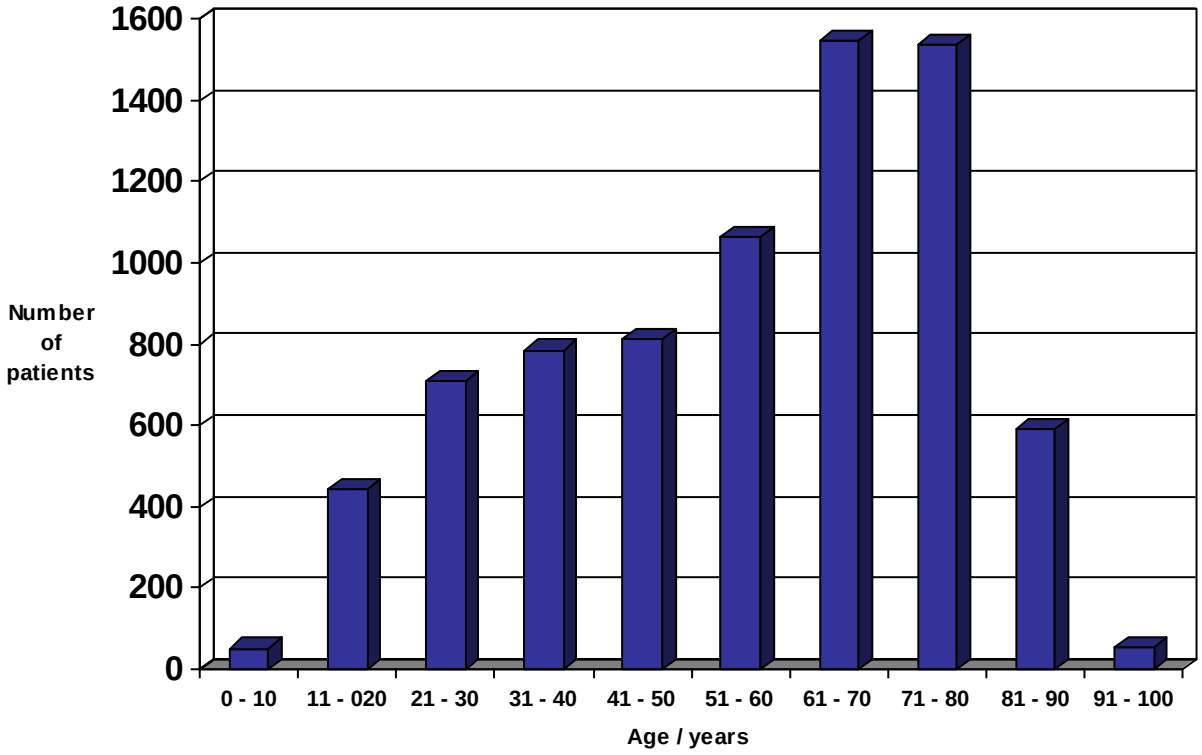
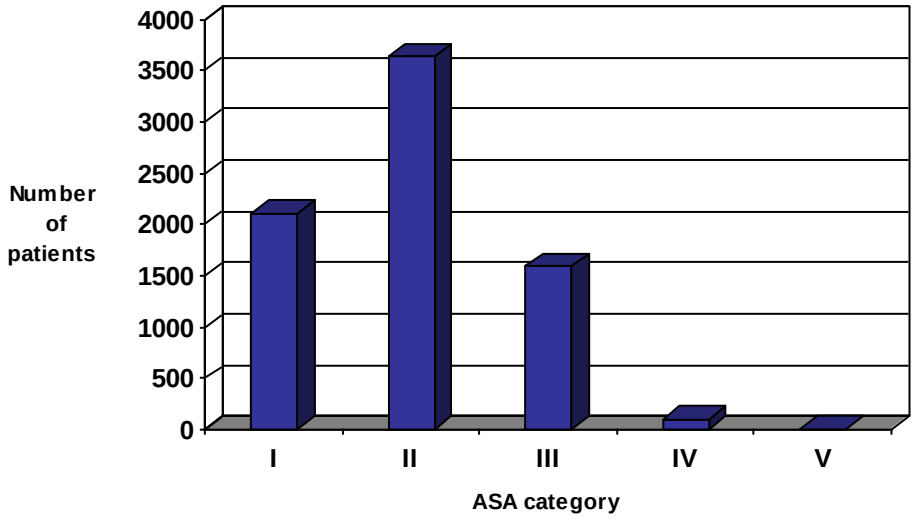


Figure 6. ASA distribution of study population

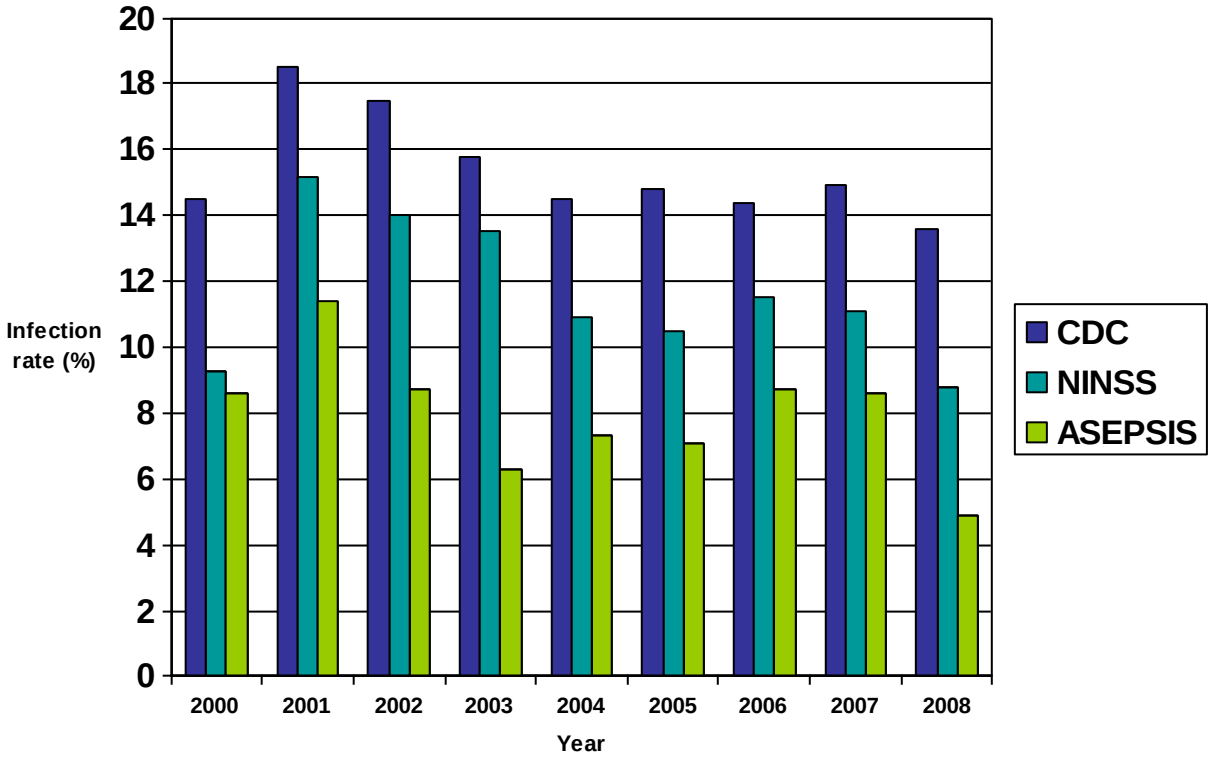


The crude infection rates (together with 95% confidence intervals) according to CDC, NINSS and ASEPSIS are shown in table 52. The incidence of infection according to the three definitions over time is shown in figure 7.

Table 52. Infection rates according to CDC, NINSS and ASEPSIS in the same series of 7299 orthopaedic patients

Definition	Number of uninfected wounds	Number of infected wounds	95% CI for infection rate
CDC	6297 (84.55%)	1151 (15.45%)	14.63 – 16.27%
NINSS	6605 (88.68%)	843 (11.32%)	10.60 – 12.04%
ASEPSIS	6793 (91.21%)	655 (8.79%)	8.15 – 9.44%

Figure 7. Incidence of infection according to CDC, NINSS and ASEPSIS from 2000 to 2008



CDC and NINSS can be divided into 'no infection', 'superficial incisional infection' and 'deep incisional infection' (table 53). ASEPSIS can be divided into 'no infection', 'disturbance of healing', 'mild infection', 'moderate infection' and 'severe infection' (table 54).

Table 53. 'Superficial' and 'deep' incisional infection rates according to CDC and NINSS

Definition	Number of wound with no infection	Number of wounds with superficial incisional infection	Number of wounds with deep incisional infection
CDC	6297 (84.5%)	689 (9.3%)	462 (6.2%)
NINSS	6605 (88.7%)	663 (8.9%)	180 (2.4%)

Table 54. Grade of infection according to ASEPSIS

ASEPSIS grade of infection	Number of wounds
No infection	6110 (82.0%)
Disturbance in healing	683 (9.2%)
Minor infection	297 (4.0%)
Moderate infection	140 (1.9%)
Severe infection	218 (2.9%)

Tables 55, 56 and 57 show the level of agreement of crude infection rates for the three definitions. The agreement between CDC and ASEPSIS is 88.94% (kappa statistic 0.4861, $p<0.0001$). The agreement between NINSS and ASEPSIS is 89.61% (kappa statistic 0.4266, $p<0.0001$). The agreement between CDC and NINSS is 95.27% (kappa statistic 0.7969, $p<0.0001$).

Table 55. Agreement between CDC and ASEPSIS infection rates

		CDC		Total
		No infection	**Infection**	
ASEPSIS	**No infection**	6133 (82.3%)	660 (8.9%)	6793 (91.2%)
	Infection	164 (2.2%)	491 (6.6%)	655 (8.8%)
Total		6297 (84.5%)	1151 (15.5%)	7448 (100.00%)

Table 56. Agreement between NINSS and ASEPSIS infection rates

		NINSS		Total
		No infection	**Infection**	
ASEPSIS	**No infection**	6312 (84.7%)	481 (6.5%)	6793 (91.2%)
	Infection	293 (3.9%)	362 (4.9%)	655 (8.8%)
Total		6605 (88.6%)	843 (11.4%)	7448 (100.00%)

Table 57. Agreement between NINSS and CDC infection rates

		NINSS		Total
		No infection	**Infection**	
CDC	**No infection**	6275 (84.2%)	22 (0.3%)	6297 (84.5%)
	Infection	330 (4.5%)	821 (11.0%)	1151 (15.5%)
Total		6605 (88.7%)	843 (11.3%)	7448 (100.00%)

Tables 58, 59 and 60 show the level of agreement between the subgroups of the three definitions of SSI. Kendall's tau b correlation coefficient between ASEPSIS and CDC is 0.5932, between ASEPSIS and NINSS is 0.4493, and between NINSS and CDC is 0.7870 (all $p<0.0001$).

Table 58. Agreement between CDC superficial and deep incisional infection rates and ASEPSIS scores

		CDC definition			Total
		No infection	Superficial infection	Deep infection	
ASEPSIS score	0-10	5770 (77.5%)	304 (4.1%)	36 (0.5%)	6110 (82.1%)
	11-20	363 (4.9%)	226 (3.0%)	94 (1.3%)	683 (9.2%)
	21-30	89 (1.2%)	98 (1.3%)	110 (1.5%)	297 (4.0%)
	31-40	18 (0.2%)	50 (0.7%)	72 (1.0%)	140 (1.9%)
	>40	57 (0.8%)	11 (0.1%)	150 (2.0%)	218 (2.9%)
Total		6297 (84.6%)	689 (9.2%)	462 (6.2%)	7448 (100.0%)

Table 59. Agreement between NINSS superficial and deep incisional infection rates and ASEPSIS scores

		NINSS Definition			Total
		No infection	**Superficial infection**	**Deep infection**	
ASEPSIS definition	**0-10**	5814 (78.1%)	281 (3.8%)	15 (0.2%)	6110 (82.0%)
	11-20	498 (6.7%)	166 (2.2%)	19 (0.3%)	683 (9.2%)
	21-30	172 (2.3%)	101 (1.4%)	24 (0.3%)	297 (4.0%)
	31-40	33 (0.4%)	86 (1.1%)	21 (0.3%)	140 (1.9%)
	>40	88 (1.2%)	29 (0.4%)	101 (1.4%)	218 (2.9%)
Total		6605 (88.7%)	663 (8.9%)	180 (2.4%)	7448 (100.00%)

Table 60. Agreement between NINSS and CDC superficial and deep incisional infection rates

		NINSS classification			Total
		No infection	Superficial infection	Deep infection	
CDC classification	No infection	6275 (84.3%)	22 (0.3%)	0 (0.00%)	6297 (84.6%)
	Superficial infection	121 (1.6%)	568 (7.6%)	0 (0.00%)	689 (9.2%)
	Deep infection	209 (2.8%)	73 (1.0%)	180 (2.4%)	462 (6.2%)
Total		6605 (88.7%)	663 (8.9%)	180 (2.4%)	7448 (100.00%)

6.4 Discussion

6.4.1 Summary

Wound surveillance on the same series of patients gives varying rates of SSI depending on the definition used. On this series of 7448 patients, the SSI rate according to CDC was 15.45%, according to NINSS was 11.32% and according to ASEPSIS was 8.79%.

Further breakdown of ASEPSIS scores reveals that 9.17% of patients had disturbance of healing (score 11-20). A score of 11-20 is known to include some infected wounds, but is thought to indicate wound breakdown from another cause in most cases[2]. If these patients were regarded as having 'true' infections, the SSI rate according to ASEPSIS would be similar to the rate according to CDC and NINSS.

A positive correlation was found between an increasing ASEPSIS score and deep incisional infection according to CDC and NINSS. Therefore, the higher the ASEPSIS score, the greater the likelihood of a deep incisional SSI.

Wound surveillance in Orthopaedic Surgery became mandatory in the NHS in England in 2004. Reported SSI rates depend on the method used for diagnosis, case mix, the thoroughness of surveillance and documentation, and the length of patient follow-up. Post-discharge follow-up is essential in any wound surveillance program since more than half of SSIs present after hospital discharge[33].

Wound surveillance has several advantages. Firstly, the feedback of infection rates to surgeons has been shown to decrease infection rates[2]. Secondly, outliers can be identified so that good practice can be recognized and propagated and, where there is poor practice, remedial assistance can be given. Thirdly, SSI rates can be used as a performance indictor to compare surgeons, departments, hospitals and countries. However, a dedicated

programme with full outpatient follow up can result in apparently high rates of infection being reported. Comparison should only be made within the institution or with hospitals operating to the same standards.

This study illustrates the necessity for one accurate and reproducible definition of SSI to allow its use as a performance indicator. The 3 definitions investigated in this study are not directly comparable and there is no good evidence at present as to which definition is best to use. CDC is used worldwide but this popularity may be unwarranted since the definition is subjective and has been shown to be unreliable[35]. The UK NINSS version of CDC was devised in an attempt to make CDC more objective, but again NINSS is not reproducible[35]. The difference in reported infection rates according to CDC and NINSS in this study shows how a small alteration in the definition of CDC has a marked effect on the reported outcome. Wound scoring methods are postulated to be superior to CDC and NINSS since they provide more information and are objective. The original ASEPSIS scoring method is known to be repeatable and related to outcome[12] but ASEPSIS has now been revised and psychometric analysis is awaited.

This study emphasizes the importance of choosing one reliable and reproducible method for diagnosing SSI. This will help to ensure that published results are both accurate and meaningful. At present, institutions choose which method they use to assess SSI. Therefore comparisons are invalid and misleading.

6.4.2 Strengths of study

This study has many strengths. It is prospective and has a large study population. A dedicated wound surveillance team with specialist training collected the data. This ensured accurate data collection and recording.

6.4.3 Limitations of study

One weakness of this study is that it was performed at a single centre so results may not be applicable to other institutions. However, the study included patients following a wide variety of orthopaedic procedures and cared for by over 20 different orthopaedic consultants so most sources of variation should be represented. A further weakness is that some wound data was recorded by ward nurses. This may have lead to a decrease in inter-observer reliability. To minimize this problem, each nurse completed a standardized simple results table with tick-boxes. Wound observations were also verified by a member of the surveillance team by asking ward nurses specific questions. Finally, some assumptions were used when diagnosing wound infection according to the different definitions. For example, a surgeon's diagnosis of wound infection was assumed when either antibiotics were prescribed or surgical wound debridement was performed. However, since the diagnosis of wound infection is known to be poorly recorded in the patient record, we considered this assumption to be more accurate than consulting the patient record alone.

6.4.4 Comparisons to other studies

ASEPSIS has been compared previously to other definitions of SSI[36] in surgical patients. In the series of mixed surgical patients there was a wider discrepancy between SSI rates according to CDC and ASEPSIS (19.2% and 6.8% respectively) than in this series of orthopaedic patients. The study of mixed surgical patients found a significant difference in the SSI rate according to CDC and NINSS, despite there being only a slight difference in the two definitions. This study of orthopaedic patients alone confirmed the finding. There have been no previous studies comparing ASEPSIS to other scoring methods in orthopaedic patients.

6.5 Summary

1. On the same series of orthopaedic patients, surgical site infection rates vary widely depending on the definition used.
2. On this series of 7448 patients, the SSI rate according to CDC was 15.45%, according to NINSS was 11.32% and according to ASEPSIS was 8.79%.
3. All institutions should use the same definition of surgical site infection to make comparisons of infection rates meaningful and fair.

Chapter 7 Conclusions and further work

7.1 Summary of thesis

Outcome measures are used to evaluate the benefits and risks of surgery. They are used for both research and audit purposes. The routine collection and reporting of outcome measures started in the UK in 2000 and has gradually evolved over time. Outcomes are now published on the 'NHS Choices' website to guide patients in their choice of institution and surgeon.

Post-operative outcome measures include mortality rates, morbidity rates and patient-reported outcome measures. The Post-Operative Morbidity Survey (POMS) is the only validated measure of post-operative morbidity in the UK. It recognises morbidity of a severity that prevents discharge from hospital.

Morbidity following hip and knee arthroplasty has previously been poorly recorded. This is the first time the POMS has been used for this purpose. Infection and renal complications are the most common types of morbidity following hip and knee arthroplasty. Pulmonary, pain and gastro-intestinal morbidity are less common. Cardiovascular, wound, neurological and haematological morbidity are the least common. Most types of morbidity decrease with time after surgery.

It is important that medical staff involved with the care of arthroplasty patients

are aware of the prevalence of post-operative morbidity in all organ systems. This allows prompt diagnosis and treatment of morbidity, minimising its impact. Knowing the prevalence of morbidity also allows appropriate pre-operative patient education and counselling. This gives patients realistic expectations and goals regarding the recovery period.

There is a statistically significant risk-adjusted difference in the presence of any type of post-operative morbidity following primary and revision hip arthroplasty on POD 3, 5, 8 and 15. Morbidity is higher in the RTHR group. There is a statistically significant risk-adjusted higher level of infection following RTHR than THR on POD 3, 5, 8 and 15. Following RTHR, there are higher levels of renal morbidity on POD 3 than following THR.

There is no difference in post-operative morbidity following primary and revision knee arthroplasty. The most likely reason for this is the low number of patients in the RTKR group in this study. After adjusting for POSSUM morbidity scores, no difference in post-operative morbidity was found between UKR and TKR, and HR and THR. Thus, newer 'bone preserving' arthroplasty procedures do not result in less post-operative morbidity than total joint replacements. After adjusting for POSSUM morbidity scores, no difference was found between post-operative morbidity following hip and knee arthroplasty procedures.

Most levels of post-operative morbidity in this study are higher than in previously published studies. The most likely reason for this discrepancy is

the use of different definitions of morbidity. The POMS definitions of morbidity in different organ systems are broad and capture more events than narrowly defined measures of morbidity. Different methods of data collection could also account for the discrepancy; this study collected data prospectively, which will always capture more events than retrospectively collected data. This study sets the benchmark for future audit. Strategies can be implemented to reduce morbidity levels and the audit cycle repeated to evaluate their impact.

Many arthroplasty patients remain in hospital with no identifiable morbidity. Two general trends are seen. Firstly, patients undergoing more 'invasive' surgery remain in hospital longer with no morbidity than patients undergoing less 'invasive' surgery. Secondly, with increasing time after surgery, fewer patients with no identifiable morbidity remain in hospital. The most common reason for non-discharge of patients with no apparent morbidity is on-going physiotherapy and occupational therapy input.

The use of the POMS as a bed utilisation tool relies on the fact it captures all morbidity. Mobility is an important factor in determining when a patient is ready for discharge following arthroplasty surgery. Addition of a 'mobility' domain to the POMS may make it more reliable as a bed utilisation tool for orthopaedic patients. The definition of mobility morbidity could be 'unable to mobilise 10m with or without walking aids and/or unable to ascend and descend a single flight of stairs'.

The use of POMS as a bed utilisation tool also assumes that patients do not develop morbidity after a period without morbidity (when they would be discharged). This was generally true, but a small proportion of patients developed wound, cardiovascular and neurological morbidity following a period without morbidity. This is not a reason to discount the POMS as a bed utilisation tool, but certain precautionary measures need to be taken. Firstly, primary arthroplasty patients should receive regular wound reviews till POD 5 and revision arthroplasty patients should receive regular wound reviews till POD 8. Patients should be aware they could develop wound infection, deep vein thrombosis, pulmonary embolus, myocardial infarction or stroke after hospital discharge. They should be given clear written instructions about the symptoms of these conditions and what to do if they develop.

Significant savings could be made if bed utilisation were to be improved in the UK. POMS could be used to identify patients remaining in hospital without clinically significant morbidity and could be used prospectively as a bed utilisation tool.

This study showed no association between post-operative morbidity defined by the POMS and longer-term patient-reported outcome measures (PROMs). Thus, this study does not support the use of the POMS as an early surrogate marker of long-term patient outcome. There was a tendency towards patients with morbidity on POD 15 having poorer PROMs, but this was not statistically significant.

Longer post-operative stay was associated with poorer PROMs in both the hip and knee arthroplasty groups. This justifies the use of length of hospital stay as an early marker of longer-term outcome. There was no association between patient age and PROMs and no association between patient sex and PROMs. Higher ASA grades were associated with poorer longer-term PROMs in the knee arthroplasty group. The association was not statistically significant in the hip group. There is no association between length of operation and PROMs.

The wound domain of the POMS has a high specificity, reasonable sensitivity, high negative predictive value and low positive predictive value when compared to the inpatient ASEPSIS score. The wording of the definition of POMS wound morbidity could account in part for the low positive predictive value. The wound domain of POMS has not been validated. It may be prudent to replace it with a psychometrically evaluated and widely used definition of wound infection such as ASEPSIS. It is mandatory for wound infection data to be collected by all hospitals so this data should be readily available.

Assessing the accuracy of the wound domain of POMS compared to the inpatients ASEPSIS score lead me to consider the level of agreement between other definitions of surgical site infection. CDC, NINSS and ASEPSIS are all definitions of surgical site infection commonly used in the UK. On the same series of orthopaedic patients, surgical site infection rates varied widely depending on the definition used. The SSI rate according to CDC was 15.45%, according to NINSS was 11.32% and according to ASEPSIS was 8.79%.

This highlights the need for all institutions to use the same definition to make comparisons meaningful and fair.

7.2 Outstanding questions

7.2.1 Use of the POMS as an audit tool

Morbidity following hip and knee arthroplasty has now been fully recorded for the first time using a validated measure of post-operative morbidity. This data should be presented together with suggestions to improve morbidity rates in each of the POMS domains. Once changes have been implemented, repeat POMS data should be collected. This is an area for further research to assess whether the collection and presentation of POMS data ultimately leads to an improvement in patient outcome. Ways by which the morbidity rates in each of the POMS domains could be improved are presented in detail in Chapter 2.

7.2.2 Use of the POMS as a bed utilisation tool

This thesis has shown that the POMS can be used as a prospective bed utilisation tool. Several ways were suggested in chapter 3 to improve bed utilisation. These include the provision of physiotherapy at home rather than as an inpatient, the delivery of home equipment prior to surgery and the arrangement of rehabilitation provision prior to surgery. Each of these suggestions requires further investigation. Once changes have been implemented, POMS data can be recollected to assess if the number of inpatients without morbidity has reduced.

7.2.3 Addition of a ‘mobility’ domain to the POMS

The use of the POMS as a discharge tool relies on the fact that it records all morbidity requiring inpatient care. At present the POMS does not assess mobility. For arthroplasty patients, this is an important factor in determining fitness for discharge. The addition of a ‘mobility’ domain may make the POMS more suitable for use as an orthopaedic bed utilisation tool. The addition of a ‘mobility’ domain to the POMS requires further investigation and psychometric evaluation.

7.2.4 Does post-operative morbidity lead to poorer long-term PROMs?

In Chapter 4, univariable analysis revealed an association between morbidity on POD 15 and poorer long-term PROMs. Multivariable analysis showed the association was not statistically significant. The sample sizes in this chapter were small and this may have led to a false negative result. This requires further investigation with a larger sample size. A true association between POMS morbidity on POD 15 and poorer long-term PROMs may be found.

7.2.5 Which definition of surgical site infection should be used?

A systematic review of the commonly used definitions of surgical site infection would be valuable. The modified version of ASEPSIS also needs to be

validated for use in orthopaedic patients. This information could be used to decide which definition of surgical site infection is most appropriate for national wound surveillance and publication on the 'NHS Choices' website.

7.3 Conclusions

Morbidity following lower limb arthroplasty has previously been poorly recorded. The POMS reveals that infection and renal complications are the most common types of morbidity following hip and knee replacement. Pulmonary, pain and gastro-intestinal morbidity are less common. Cardiovascular, wound, neurological and haematological morbidity are the least common. Most types of morbidity decrease with time after surgery.

Many patients remain in hospital with no identifiable morbidity. The POMS can be used prospectively to identify these patients. The use of POMS as a bed utilisation tool is based on the assumption that it identifies all patients with clinically significant morbidity. The most common reason for non-discharge of patients without morbidity in this study was on-going physiotherapy and occupational therapy input. This raises the possibility that patients were not sufficiently mobile for safe discharge. The addition of a 'mobility' domain to the POMS may improve its utility as a bed utilisation tool for orthopaedic patients.

This study found no association between the POMS and PROMs at 18 months post-surgery. There was a tendency towards an association between the presence of morbidity on POD 15 and poorer PROMs at 18-months post-

surgery, but this was not statistically significant on multivariable analysis. Thus, this study does not support the use of the POMS as an early surrogate marker of longer-term PROMS in orthopaedic patients. An association was found to exist between longer inpatient stay and PROMs at 18 months post-surgery. Thus, this study supports the use of length of inpatient stay as an early surrogate marker of longer-term patient-reported outcome.

The wound domain of POMS has high specificity, reasonable sensitivity, high negative predictive value and low positive predictive value compared to the inpatient ASEPSIS score. It may be prudent to consider replacing the wound domain of the POMS with a validated definition of surgical site infection.

On the same series of orthopaedic patients, surgical site infection rates varied widely depending on the definition used. According to the CDC definition the infection rate was 15.45%, according to the NINSS definition it was 11.32% and according to the ASEPSIS definition it was 8.79%. When comparing surgical site infection rates between institutions, the same definition should be used to ensure comparisons are valid.

Publications related to this thesis

1. ‘How should we measure wound infection rates to ensure high quality care for all?

***Ashby E**, Haddad FS, Wilson APR*
The Journal of Bone and Joint Surgery (Br.) 2010 Sept, 92(9): 1294-1299

How will surgical site infection be measured to ensure "high quality care for all"?

E. Ashby,
F. S. Haddad,
E. O'Donnell,
A. P. R. Wilson

From University College London Hospital, London, United Kingdom

As of April 2010 all NHS institutions in the United Kingdom are required to publish data on surgical site infection, but the method for collecting this has not been decided. We examined 7448 trauma and orthopaedic surgical wounds made in patients staying for at least two nights between 2000 and 2008 at our institution and calculated the rate of surgical site infection using three definitions: the US Centers for Disease Control, the United Kingdom Nosocomial Infection National Surveillance Scheme and the ASEPSIS system. On the same series of wounds, the infection rate with outpatient follow-up according to Centre for Disease Control was 15.45%, according to the UK Nosocomial Infection Surveillance was 11.32%, and according to ASEPSIS was 8.79%. These figures highlight the necessity for all institutions to use the same method for diagnosing surgical site infection.

If different methods are used, direct comparisons will be invalid and published rates of infection will be misleading.

E. Ashby, MRCS, Orthopaedic Registrar
Chelsea and Westminster Hospital, 369 Fulham Road, London SW10 9NH, UK.

F. S. Haddad, FRCS(Orth) Consultant Orthopaedic Surgeon

E. O'Donnell, Team Leader Wound Surveillance

A. P. R. Wilson, Consultant Microbiologist
Department of Clinical Microbiology, University College London Hospital, 235 Euston Road, London NW1 2BU, UK.

Correspondence should be sent to Miss E. Ashby; e-mail: elizabethashby@doctors.org.uk

doi:10.1302/0301-620X.92B9.22401 $2.00

J Bone Joint Surg [Br] 2010;92-B:1294-9.
Received 4 February 2010; Accepted after revision 6 April 2010

In June 2008 the Department of Health in the United Kingdom published a report entitled "*High Quality Care For All*".[1] This stated that by April 2010, healthcare providers working for or on behalf of the NHS must publish 'quality accounts'. These will cover three aspects of patient care: safety, patient experience and patient outcome. As part of the safety aspect, rates of surgical site infection (SSI) will be published.

SSI is a major risk in all orthopaedic surgery. It causes pain and can lead to wound dehiscence, deep infection and generalised sepsis. Further surgery and admission to intensive care may be necessary. A patient with an SSI spends twice as long in hospital.[2] It is therefore not only distressing for the patient but also an economic burden.

In the report "*High Quality Care For All*" report, seven steps were described to improve quality. These are to set standards of high quality; measure quality; publish quality performance; recognise and reward quality; raise standards; safeguard quality; and stay ahead by encouraging innovation. The report stated that data on SSI will soon be available on the NHS Choices website (step 3) but did not state what a 'good' rate of SSI is (step 1), nor did it describe how such rates will be measured (step 2). This paper concentrates on step 2 and examines the methods commonly used to measure SSI.

There is a misconception that SSIs are easy to define and diagnose. There are several definitions of SSI and the diagnosis of infection varies between surgeons. They were traditionally diagnosed using the findings on examination of pain, redness, heat, swelling and impairment of function. As surgeons became increasingly accountable for their practice, more reliable and reproducible methods of diagnosing SSI became necessary. There are three definitions in use today (i) the US Centers for Disease Control (CDC), (ii) the English Nosocomial Infection National Surveillance Scheme (NINSS) and (iii) an English system which combines scores for additional treatment, serous discharge, erythema, purulent exudate, separatin of deep tissues, isolation of bacteria and stay as an in-patient (ASEPSIS).

The CDC definition[3] is used throughout the world and divides SSIs into incisional and organ/space infections. Incisional SSIs are further divided into superficial and deep infections (Table I). Several of the CDC criteria are subjective, and on psychometric evaluation have been shown to be unrealiable.[4]

The United Kingdom NINSS definition of SSI is based on the CDC, with two significant modifications. Firstly, pus cells must be present for a wound culture to be classified as positive and secondly, a surgeon's diagnosis of infection is excluded as a sufficient criterion to diagnose SSI. These changes were implemented to

Table I. The Center for Disease Control definition of surgical site infection

	Superficial infection (involving skin and superficial tissues)	Deep infection (involving the fascial and muscle layers)
Time period	Occurs within 30 days of surgery	Occurs within 30 days of surgery or within one year if implant present
Site	Involves only the skin and superficial tissue	Related to the surgical site and involves deep tissues
Further criteria	Must fulfil one of the following:	Must fulfil one of the following:
	Purulent discharge from superficial incision	Purulent discharge from deep incision
	Organisms isolated from incision.	Spontaneous dehiscence or deliberate opening of a deep incision, following fever or pain or tenderness around the wound (unless cultures are negative)
	Pain, tenderness, swelling, redness or heat around the incision AND the incision deliberately opened by a surgeon (unless cultures are negative)	
	Diagnosis by a surgeon or physician	Abscess involving a deep incision
		Diagnosis by a surgeon or physician

Table II. Points scale used to calculate total ASEPSIS score

Criterion	Points
Additional treatment	
Antibiotics	10
Drainage of pus under local anaesthetic	5
Debridement of wound under general anaesthetic	10
Serous discharge	0 to 5
Erythema	0 to 5
Purulent exudate	0 to 10
Separation of deep tissues	0 to 10
Isolation of bacteria	10
Stay in hospital over 14 days	5

Table III. Points scale for ASEPSIS daily wound inspection

	Proportion of wound affected (%)					
	0	>0 & ≤ 9	20 & ≤ 39	40 & ≤ 59	60 & ≤ 79	80 & ≤ 100
Serous exudate	*0*	*1*	*2*	*3*	*4*	*5*
Erythema	*0*	*1*	*2*	*3*	*4*	*5*
Purulent exudates	*0*	*2*	*4*	*6*	*8*	*10*
Separation of deep tissues	*0*	*2*	*4*	*6*	*8*	*10*

Table IV. Breakdown of ASEPSIS scores

Score	Meaning
0 to 10	No infection / Normal healing
• to 20	Disturbance of healing
21 to 30	Minor infection
31 to 40	Moderate infection
³ 41	Severe infection

Table V. Patient demographics / Gender

M:F	44:56
Mean age in years (range)	56.4 (0 to 99)
ASA*[14][11] (%)	
I	2109 (*28.3*)
II	3639 (*48.9*)
III	1595 (*21.4*)
IV	105 (*1.4*)
Type of surgery	
Elective: emergency	86:14

*ASA, american society of anesthesiologists

improves the objectivity of CDC, but the reproducibility of NINSS remains low.[5]

ASEPSIS is a quantitative wound scoring method.[6] The score is calculated using objective criteria based both on visual characteristics of the wound and the consequences of infection (Tables II and III). A score of >10 indicates an increasing probability and severity of infection (Table IV). The original ASEPSIS method has been shown to be both objective and repeatable,[7] but a more recent revised version has not been psychometrically evaluated.

Scoring methods provide more detailed information than the CDC and NINSS but are more complicated to perform, more time-consuming and more costly. It takes an average of 59 minutes to collect the data and calculate an ASEPSIS score.[8]

Table VI. Infection rates according to CDC, NINSS and ASEPSIS

Definition*	Number of uninfected wounds (%)	Number of infected wounds (%)	95% CI† for infection rate
CDC	6297 (*84.55*)	1151 (*15.45*)	*14.63* to *16.27*
NINSS	6605 (*88.68*)	843 (*11.32*)	*10.60* to *12.04*
ASEPSIS	6793 (*91.21*)	655 (*8.79*)	*8.15* to *9.44*

*CDC, center for disease control; NINSS, nosocomial infection national surveillance scheme
†CI, confidence interval

The purpose of this prospective observational study was to assess the rate of SSI in orthopaedics according to these three common definitions.

Patients and Methods

In May 2000 a wound surveillance programme began in the Department of Trauma and Orthopaedics at University College Hospital, London. For the first two years, because of funding restraints, data were collected for six months of each year, from May to October. This represented 35% of total orthopaedic admissions. From 2002 onwards, collection of data became continuous. This study is part of a hospital audit programme and so formal ethical approval was not required.

All trauma and elective orthopaedic patients with a minimum of two-nights stay in hospital and an operation involving the incision of tissue were included in the programme. Traumatic wounds were not included, only incisions made at the time of surgery.

Wounds were assessed by a member of a specialist team, made up of four nurses and a healthcare assistant. The sole role of this team was to collect and record data on wound infection, and all members received specialist training in the different definitions and diagnosis of surgical site infection.

Each patient was reviewed on three separate occasions, once pre-operatively and twice post-operatively. The same standardised data collection sheet was completed for each patient. The information collected pre-operatively included patient demographics and the name of the consultant in charge. Operative information and microbiology results came from a direct interface with other hospital computer databases. Microbiology tests, such as wound swabs or tissue cultures, were performed according to clinical judgement. No specific microbiology tests were requested for study purposes alone.

Surgical wounds were inspected two or three days after operation and again on days four or five if the patient remained in hospital. The proportion of each wound exhibiting erythema, serous discharge, purulent discharge or dehiscence was recorded. Wounds were inspected directly if visible, but if a dressing was present the relevant information was gained by questioning nursing staff. This was done to avoid an unnecessary increase in the risk of infection. The nurses were encouraged to fill in the data collection sheet at the time of dressing change.

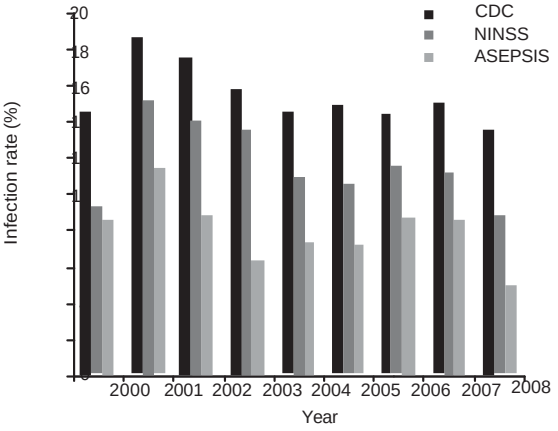


Fig. 1

Graph showing infection rates (CDC, center for disease control; NINSS, nosocomial infection national surveillance scheme).

At each post-operative visit the notes and drug charts of the patient were inspected. The diagnosis of a wound infection by a medical practitioner, the prescription of prophylactic or therapeutic antibiotics and the opening of a wound or drainage of an abscess were recorded.

At the time of discharge patients were given a simple 'yes/no' questionnaire regarding their wound, which they were asked to complete and return in a pre-paid envelope two months later. They were contacted by telephone if no postal questionnaire was returned. The questionnaire was used to ascertain whether a wound infection had been diagnosed since discharge, whether antibiotics had been prescribed for the wound. whether any further surgery had been necessary and whether the hospital stay had been longer than 14 days.

The data were stored on a modified Access 97 database which was only accessible to members of the surveillance team. A single patient episode was defined as an operation with follow-up of either three months or until a further operation was performed, whichever was shorter. At any time point, SSI resulting in readmission was recorded in the database. Sufficient information was gathered to allow each wound to be diagnosed according to the CDC, NINSS and ASEPSIS definitions.

Table VII. 'Superficial' and 'deep' incisional infection rates according to CDC and NINSS

Definition*	Number of wounds with no infection (%)	Number of wounds with superficial incisional infection (%)	Number of wounds with deep incisional infection (%)
CDC	6297 (*84.5*)	689 (*9.3*)	462 (*6.2*
NINSS	6605 (*88.7*)	663 (*8.9*)	180 (*2.4*

*CDC, Center for Disease Control; NINSS, Nosocomial Infection National Surveillance Scheme

Table VIII. Grade of infection according to ASEPSIS

Grade of infection	Number of wounds (%)
No infection	6110 (*82.0*)
Disturbance in healing	683 (*9.2*)
Infection	
Minor	297 (*4.0*)
Moderate	140 (*1.9*)
Severe	218 (*2.9*)

Table IX. Agreement between CDC* and ASEPSIS infection rates

	CDC		
	No infection (%)	Infection (%)	Total (%)
ASEPSIS			
No infection	6133 (*82.3*)	660 (*8.9*)	6793 (*91.2*)
Infection	164 (*2.2*)	491 (*6.6*)	655 (*8.8*)
Total	6297 (*84.5*)	1151 (*15.5*)	7448 (*100.0*)

*CDC, Center for Disease Control

Table X. Agreement between NINSS* and ASEPSIS infection rates

	NINSS		
	No infection (%)	Infection (%)	Total (%)
ASEPSIS			
No infection	6312 (*84.7*)	481 (*6.5*)	6793 (*91.2*)
Infection	293 (*3.9*)	362 (*4.9*)	655 (*8.8*)
Total	6605 (*88.6*)	843 (*11.4*)	7448 (*100.0*)

*NINSS, Nosocomial Infection National Surveillance Scheme

Table XI. Agreement between NINSS* and CDC† infection rates

	NINSS		
	No infection (%)	Infection (%)	Total (%)
CDC			
No infection	6275 (*84.2*)	22 (*0.3*)	6297 (*84.5*)
Infection	330 (*4.5*)	821 (*11.0*)	1151 (*15.5*)
Total	6605 (*88.7*)	843 (*11.3*)	7448 (*100.0*)

* NINSS, Nosocimial Infection National Surveillance Scheme
† CDC, Center for Disease Control

Statistical analysis. Crude infection rates were calculated for each definition of infection and 95% confidence intervals calculated. The agreement between crude infection rates was calculated using the k statistic. Kendall's tau b value was used to assess correlation between the subdivisions of each definition.

Results

Between May 2000 and October 2008, 7448 orthopaedic wounds were assessed in 7299 patients. The patient demographics are shown in Table V. The rate of follow-up of patients two months after surgery was 91%.

The crude infection rates according to CDC, NINSS and ASEPSIS are shown in Table VI and the incidence of infection according to the three definitions over time is shown in Figure 1. CDC and NINSS can be divided into 'no infection', 'superficial incisional infection' and 'deep incisional infection' (Table VII). ASEPSIS can be divided into 'no infection', 'disturbance of healing', 'mild infection', 'moderate infection' and 'severe infection' (Table VIII).

Tables IX, X and XI show the level of agreement of crude infection rates for the three definitions. The agreement between CDC and ASEPSIS is 88.94% (k statistic 0.4861, $p < 0.0001$), that between NINSS and ASEPSIS is 89.61% (k statistic 0.4266, $p < 0.0001$) and that between CDC and NINSS is 95.27% (k statistic 0.7969, $p < 0.0001$).

Tables XII, XIII and XIV show the level of agreement between the subgroups of the three definitions of SSI. Kendall's tau b correlation coefficient between ASEPSIS and CDC is 0.5932, between ASEPSIS and NINSS is 0.4493, and between NINSS and CDC is 0.7870 (all $p < 0.0001$).

Discussion

Wound surveillance in the same series of patients gives varying rates of SSI depending on the definition used. On this series of 7448 patients, the SSI rate according to the CDC was 15.45%, according to the NINSS was 11.32% and according to ASEPSIS was 8.79%. The latter has previously been compared with other definitions of SSI,[9] but never in a series of trauma and orthopaedic patients only.

Further breakdown of ASEPSIS scores reveals that 9.17% of patients had disturbance of healing (score 11 to 20). A score of 11 to 20 is known to include some infected wounds, but in most cases is thought to indicate wound breakdown from another cause.[10] If these patients were regarded as having 'true' infections, the SSI rate according to ASEPSIS would be similar to the rate according to CDC and NINSS.

Table XII. Agreement between CDC* superficial and deep incisional infection rates and ASEPSIS scores

	CDC definition			
	No infection (%)	Superficial infection (%)	Deep infection (%)	Total
ASEPSIS score				
0 to 10	5770 (77.5)	304 (4.1)	36 (0.5)	6110 (82.1)
11 to 20	363 (4.9)	226 (3.0)	94 (1.3)	683 (9.2)
21 to 30	89 (1.2)	98 (1.3)	110 (1.5)	297 (4.0)
31 to 40	18 (0.2)	50 (0.7)	72 (1.0)	140 (1.9)
>40	57 (0.8)	11 (0.1)	150 (2.0)	218 (2.9)
Total	6297 (84.6)	689 (9.2)	462 (6.2)	7448 (100.0)

* CDC, center for disease control

Table XIII. Agreement between NINSS* superficial and deep incisional infection rates and ASEPSIS scores

	NINSS definition			
	No infection (%)	Superficial infection (%)	Deep infection (%)	Total
ASEPSIS score				
0 to 10	5814 (78.1)	281 (3.8)	15 (0.2)	6110 (82.0)
11 to 20	498 (6.7)	166 (2.2)	19 (0.3)	683 (9.2)
21 to 30	172 (2.3)	101 (1.4)	24 (0.3)	297 (4.0)
31 to 40	33 (0.4)	86 (1.1)	21 (0.3)	140 (1.9)
>40	88 (1.2)	29 (0.4)	101 (1.4)	218 (2.9)
Total	6605 (88.7)	663 (8.9)	180 (2.4)	7448 (100.0)

* NINSS, nosocomial infection national surveillance scheme

Table XIV. Agreement between NINSS* and Center for Disease Control (CDC) superficial and deep incisional infection rates

	NINSS definition			
	No infection (%)	Superficial infection (%)	Deep infection (%)	Total
CDC classification				
No infection	6275 (84.3)	22 (0.3)	0 (0.0)	6297 (84.6)
Superficial infection	121 (1.6)	568 (7.6)	0 (0.0)	689 (9.2)
Deep infection	209 (2.8)	73 (1.0)	180 (2.4)	462 (6.2)
Total	6605 (88.7)	663 (8.9)	180 (2.4)	7448 (100.0)

* NINSS, nosocomial infection national surveillance scheme

A positive correlation was found between an increasing ASEPSIS score and deep incisional infection according to the CDC and the NINSS. Therefore, the higher the ASEPSIS score, the greater the likelihood of a deep incisional SSI.

The strengths of this study are that it is prospective, there is a large study population and a dedicated wound surveillance team with specialist training who collected and recorded the data. One weakness is that some of the data was recorded by ward nurses, leading to a possible decrease in inter-observer reliability. In order to minimise this problem, each nurse completed a standardised simple results table with tick-boxes, and all wound observations were verified by the surveillance team asking the ward nurses specific questions.

Wound surveillance in orthopaedic surgery became mandatory in the NHS in England in 2004. Reported rates of SSI depend on the method used for diagnosis, the case mix, the thoroughness of surveillance and documentation, and the duration of patient follow-up. Follow-up after discharge is essential in any wound surveillance programme, as more than half of SSIs present after discharge.[11]

Wound surveillance has several advantages. Firstly, the feedback of rates of infection to surgeons has been shown to reduce them.[10] Secondly, outliers can be identified so that good practice can be recognised and propagated, and poor practice can be highlighted and improved. Thirdly, rates of SSI can be used as a performance indicator to compare surgeons, departments, hospitals and countries. However, a

dedicated programme with full outpatient follow-up can result in apparently high rates of infection being reported. Comparison should only be made within the institution or with hospitals operating to the same standards.

This study illustrates the need for a single accurate and reproducible definition of SSI to allow its use as a reliable performance indicator. The three definitions investigated here are not directly comparable, and at present there is no good evidence as to which is best to use. The CDC is used worldwide, but this popularity may be unwarranted as the definition is subjective and has been shown to be unreliable.[4] The United Kingdom NINSS version of CDC was devised in an attempt to make CDC more objective, but again NINSS is not reproducible.[4] The difference in reported infection rates according to CDC and NINSS in this study shows how a small alteration in the definition of the former has a marked effect on the reported outcome. Methods of wound scoring are claimed to be superior to those of the CDC and NINSS, because they provide more information and are objective.

From 2010, hospitals will no longer be paid a fixed tariff for each procedure: instead, they will receive a variable amount according to 'outcome'. Therefore, hospitals with lower SSI rates will receive higher tariffs. In order for this to be fair, the rates must be reported in a reliable and accurate fashion.

We recommend the use of the ASEPSIS scoring method. Both CDC[4] and NINSS[5] have been shown to be unreliable. The original ASEPSIS method is objective and repeatable,[7] but a more recent revised version has not been psychometrically evaluated and this must be promptly addressed.

Before the publication of SSI rates within the profession and to the general public, the Department of Health must clearly define it. If different diagnostic methods are used by different institutions, any comparison will be meaningless and misleading.

Supplementary material

A further opinion by Professor S. Hughes is available with the online version of this article on our website at www.jbjs.org.uk

APR Wilson was part funded by the UCLH/UCL comprehensive biomedical centre.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References

1. **Darzi A.** High quality care for all: NHS Next Stage Review final report. Department of Health, 2008. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyandGuidance/DH_085825 (date last accessed 7 April 2010).
2. **Coello R, Charlett A, Wilson J, et al.** Adverse impact of surgical site infections in English hospitals. *J Hosp Infect* 2005;60:93-103.
3. **Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG.** CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol* 1992;13:606-8.
4. **Allami MK, Jamil W, Fourie B, Ashton V, Gregg PJ.** Superficial incisional infection in arthroplasty of the lower limb: interobserver reliability of the current diagnostic criteria. *J Bone Joint Surg Br* 2005;87-B:1267-71.
5. **Wilson JA, Ward VP, Coello R, Charlett A, Pearson A.** A user evaluation of the Nosocomial Infection National Surveillance System: surgical site infection module. *J Hosp Infect* 2002;52:114-21.
6. **Wilson AP, Treasure T, Sturridge MF, Grüneberg RN.** A scoring method (ASEPSIS) for post-operative wound infections for use in clinical trials of antibiotic prophylaxis. *Lancet* 1986;1:311-13.
7. **Wilson AP, Webster A, Gruneberg RN, Treasure T, Sturridge MF.** Repeatability of a sepsis wound scoring method. *Lancet* 1986;1:1208-9.
8. **Wilson AP, Helder N, Theminimulle SK, Scott GM.** Comparison of wound scoring methods for use in audit. *J Hosp Infect* 1998;39:119-26.
9. **Wilson AP, Gibbons C, Reeves BC, et al.** Surgical wound infection as a performance indicator: agreement of common definitions of wound infection in 4773 patients. *BMJ* 2004;329:720.
10. **Wilson AP, Hodgson B, Liu M, et al.** Reduction in wound infection rates by wound surveillance with postdischarge follow-up and feedback. *Br J Surg* 2006;93:630-8.
11. **Brown RB, Bradley S, Doptiz E, et al.** Surgical wound infections documented after hospital discharge. *Am J Infect Control* 1987;15:54-8.
12. **No authors listed.** New classification of physical status. *Anesthesiology* 1963;21:111.

2. ‘Outcome measures for orthopaedic interventions on the hip’

***Ashby E**, Grocott MPW, Haddad FS*
J Bone Joint Surg Br. 2008 May; 90(5): 545-9.

3. ‘How efficient is patient discharge following lower limb arthroplasty?’

***Ashby E**, Matejowsky C, Mythen MG, Haddad FS, Grocott MPW*

Submitted to Perioperative Medicine August 2014

How efficient is patient discharge following lower limb arthroplasty?

Authors

E Ashby (Orthopaedic Registrar, Great Ormond Street Hospital, London)

C Matejowsky (Surgical Outcomes Research Centre, University College London Hospitals)

MG Mythen (Smiths Medical Professor of Anaesthesia, UCL Institute of Child Health; Surgical Outcomes Research Centre, Joint UCH/UCL Comprehensive Biomedical Research Centre)

FS Haddad (Consultant Orthopaedic Surgeon, University College London Hospital)

MPW Grocott (Senior Lecturer, Critical Care Medicine, UCL Institute of Child Health; Surgical Outcomes Research Centre, Joint UCH/UCL Comprehensive Biomedical Research Centre)

Correspondence

Miss Elizabeth Ashby

Declarations Nil

Abstract

Appropriately timed patient discharge is essential for optimal patient care and efficient hospital functioning. The Post-Operative Morbidity Survey (POMS) is the only validated prospective method of measuring short-term post-operative morbidity. It has not been used previously as a bed utilisation tool. We collected POMS data from 529 consecutive lower limb arthroplasty patients over a one-year period and recorded the number of patients remaining in hospital without morbidity, together with alternative reasons for remaining in hospital. On post-operative day 5, 53% of all hip arthroplasty patients and 47% of all knee arthroplasty patients remained in hospital with no identifiable morbidity. The most common reason for inappropriate bed occupancy was on-going physiotherapy and occupational therapy. We believe that the Post Operative Morbidity Survey is able to identify patients remaining in hospital with no significant morbidity and has utility as a prospective bed utilisation tool. Addition of a mobility measure to the POMS may improve the utility of this survey in detecting morbidity requiring hospitalisation, particularly following lower limb joint replacement surgery.

Introduction

Appropriately timed discharge of patients following surgery is essential for optimal patient care and efficient hospital functioning. A patient discharged early is at risk of under-diagnosis of medical complications with consequent adverse outcome. A patient whose discharge is delayed is at risk of developing a hospital-associated complication (e.g. hospital-acquired infection) and incurs an unnecessary cost to the health-care

provider. Post-operative patients should be discharged at the earliest safe opportunity to reduce the rate of hospital-associated complications and the cost of each inpatient episode. Appropriate discharge timing should increase the throughput of patients and reduce waiting times.

Historically, hospitals in the UK have been paid according to contracts with no financial incentive to treat increased numbers of patients. This changed in 2000 when the NHS Plan[1] announced that hospital incomes would be directly linked to activity. Payment by Results[2] began in 2003 and now every healthcare provider is paid a sum (tariff) for each procedure undertaken. In the UK, many patients remain in hospital with no medical indication[3]. One study showed that 31% of post-operative patients remained in hospital inappropriately[4]. Payment by Results aims to reduce this figure by rewarding efficiency and encouraging increased activity.

In order to improve efficiency, hospitals must first recognise inappropriate bed occupancy. The Post-Operative Morbidity Survey (POMS)[5] is the only validated prospective method of measuring short-term post-operative morbidity in the literature. The POMS was designed to identify morbidity of a type and severity that could delay discharge from hospital. The survey focuses on indicators of organ systems dysfunction (e.g. inability to tolerate enteral diet) rather than traditional diagnostic categories (e.g. deep vein thrombosis). The POMS assesses 9 domains of morbidity (Table 1). Data is obtained from observation charts, medication charts, patient notes, routine blood test results and direct patient questioning and observation. No additional investigations are required. The data collection process is simple to allow routine screening of large

numbers of patients. The POMS has been shown to be reliable, valid and acceptable to patients[6]. The POMS has been used in outcomes research[7] and in effectiveness research[8].

In the US, over 98% of post-operative inpatients had morbidity defined by the POMS[5]. This implies that patients with a POMS score of zero are fit for discharge. Therefore, as well as providing useful clinical research and audit data, the POMS may have utility for assessing and improving hospital bed utilisation.

The aim of this study is to identify inappropriate bed occupancy following lower limb arthroplasty using the POMS. We report the reasons for delayed discharge and suggest ways to improve bed utilisation.

Methods

Ethics Committee approval was sort prior to commencement of the study. The requirement for consent was waived as collection of the POMS has become a routine part of service evaluation within our institution. All patients aged 18 or over undergoing elective lower limb arthroplasty at University College Hospitals NHS Trust over a 12 month period were eligible for inclusion into this prospective cohort study. There were no exclusion criteria ensuring a consecutive sample was taken. Elective lower limb arthroplasty procedures included unicompartmental knee replacement (UKR), total knee replacement (TKR), revision total knee replacement, hip resurfacing (HR), total hip replacement (THR) and revision total hip replacement.

Data was collected by one of two study nurses. The age, sex, ASA and length of inpatient stay for each patient was recorded. POMS data were collected on post-operative days (POD) 3, 5, 8 and 15 if the patient remained in hospital. Presence of morbidity was defined as occurring in any patient meeting POMS criteria, in one or more domain of the survey, on the day of data collection. For patients remaining in hospital without morbidity on POD 8 and POD 15, the reason was recorded. The use of mobility aids on these days was also noted.

The number and percentage of patients with no identifiable morbidity according to the POMS was calculated for POD 3, 5, 8 and 15. The number of days each patient remained in hospital with no morbidity was calculated by distracting the day on which the patient first had a POMS score of zero from their total length of stay. An overall estimated cost saving was calculated by multiplying this figure by the average cost for one orthopaedic inpatient night.

The number of patients developing post-operative morbidity after a period free of morbidity was recorded. The number of readmissions to the same hospital in the first year following discharge was also recorded.

Results

Data collection was completed on 529 patients. Patient characteristics of the study population are shown in table 2. The mean age of all study patients was 68.9 years,

the median ASA was 2 and 62% of patients were female. The median length of stay was 7 days and the overall inpatient mortality rate was 0.4%.

A) Hip arthroplasty patients

The location of hip arthroplasty patients on POD 3, 5, 8 and 15 is shown in table 3. Many patients undergoing HR remained in hospital with no identified morbidity on POD 3 (75%), 5 (78%) and 8 (16%). All HR patients were discharged by POD 15.

Many THR patients remained in hospital with no identifiable morbidity on POD 3 (46%), 5 (54%), 8(34%) and 15 (7%). Patients undergoing revision THR patients also remained in hospital with no identifiable morbidity on POD 3 (11%), 5 (29%), 8 (31%) and 15 (3%). Discharge status and prevalence of morbidity for all hip arthroplasty patients combined are presented in figure 1.

B) Knee arthroplasty patients

The location of knee arthroplasty patients on POD 3, 5, 8 and 15 is shown in table 4. Many patients undergoing UKR remained in hospital with no identified morbidity on POD 3 (63%), 5 (39%) and 8 (6%). All UKR patients had been discharged by POD 15.

Many TKR patients remained in hospital with no identifiable morbidity on POD 3 (50%), 5 (50%), 8 (19%) and 15 (4%). Revision TKR patients also remained in hospital with no identifiable morbidity on POD 3 (52%), 5 (47%), 8 (17%) and 15 (3%).

Discharge status and prevalence of morbidity for all knee arthroplasty patients combined are presented in figure 2.

C) Overall inappropriate bed occupancy days

Table 5 and figure 3 show the average number of days that post-operative patients remained in hospital with no identifiable morbidity. HR patients stayed an average of 2.36 days, THR patients 4.19 days, and revision THR patients 10.37 days. UKR patients stayed an average of 1.76 days with no identifiable morbidity, TKR patients 2.73 days, and revision TKR patients 14.38 days.

D) Cost of inappropriate bed occupancy days

Overall, 529 patients undergoing lower limb arthroplasty were included in this study. These patients remained in hospital for a total of 1965 days with no morbidity as defined by the POMS. A surgical inpatient bed costs up to £400 per night[9]. If these patients had been discharged when their POMS score was zero, a saving of up to £786,000 could have been made.

E) Reasons for patients with no morbidity remaining in hospital

Of the 529 patients participating in this study, 120 remained in hospital with no morbidity defined by the POMS on POD 8 and 20 patients remained with no identifiable morbidity on POD 15. The reasons for non-discharge are shown in

figure 4. The most common reason is continuing physiotherapy and occupational therapy input. Other reasons include waiting for home equipment, waiting for a rehabilitation bed, waiting for a social services package of care and patients feeling unwell with negative investigations.

Of the patients remaining in hospital with no morbidity identified by the POMS, 24% were mobilising with a zimmer frame, 55% were mobilising with two crutches, 14% with a single crutch and 7% were mobilising unaided. This study did not record how far patients could mobilise or whether they could climb stairs.

F New morbidity and readmission

38 out of 529 patients developed morbidity as an inpatient following a period without morbidity. 5 of these patients underwent a second surgical procedure and developed morbidity in the second post-operative period. 33 patients (6.2%) developed morbidity following a period without morbidity. 25 of these patients developed wound morbidity, 6 developed cardiovascular morbidity, 1 developed neurological morbidity and 1 patient developed neurological morbidity.

24 patients developed wound morbidity by POD 5 having had no morbidity on POD 3. One revision arthroplasty patient developed wound morbidity by POD 8 having had no morbidity on POD 5. Of the 6 patients that developed cardiovascular morbidity, 5 were prescribed new anticoagulation (two for pulmonary embolus and three for deep vein thrombosis) and 1 patient had a myocardial infarction. 1 THR patient developed a

CVA and another THR patient developed infectious morbidity (an infected peripheral cannula site) after a period with no morbidity.

No patient in this study was readmitted to the same hospital in the first year following discharge for any reason relating to their surgery.

Discussion

This study identifies many patients remaining in hospital with no identifiable morbidity following lower limb arthroplasty in a UK teaching hospital. The rate of inappropriate bed occupancy varies with time after surgery and type of arthroplasty.

The most common reason for patients remaining in hospital with no identifiable morbidity was on-going physiotherapy and occupational therapy. This suggests that improving pre- and post-operative planning could reduce inappropriate bed occupancy. Prior to surgery patients could be taught post-operative physiotherapy exercises in group classes. Occupational therapists could assess each patient's home environment and ensure necessary modifications are made. In the post-operative period 'fast-track' pathways could be used to ensure maximum therapy in-put at the earliest possible opportunity. Some physiotherapy and occupational therapy could be provided post-operatively at the patient's home rather than in hospital. This would require safety and cost evaluation prior to implementation.

Three of the top five reasons for patients remaining in hospital with no identifiable morbidity relate to 'social' issues (awaiting home equipment, awaiting a rehabilitation bed, awaiting a package of care from social services). Pre-operative clinics could identify and address these problems prior to admission. Such clinics could also be used to manage patient expectation so they are aware of the difficulties they may encounter in the post-operative period and the expected timing of discharge.

This study has several strengths. A large consecutive dataset was collected prospectively using a validated methodology for measuring post-operative morbidity. This is the first published study to prospectively evaluate the appropriateness of discharge following lower limb joint replacement surgery specifically.

The weaknesses of this study are that it was conducted in a single centre, the POMS is not validated as a bed utilisation tool, there was not daily recording of data so the calculation of excess days are an approximation, and patient mobility was not fully assessed. Data was collected regarding mobility aids, but the distance each patient could mobilise was not recorded.

This is the first time the POMS has been used as a bed utilisation tool. It has not been validated for this purpose but has previously been used to identify patients in hospital without morbidity[5,6]. In the US over 98% of inpatients had morbidity defined by the POMS[5] suggesting that patients with a POMS score of zero were rapidly discharged. In a previous UK study 63% of orthopaedic patients remained in hospital with no

morbidity on POD 3 and 42% on POD 5 suggesting that discharge efficiency was lower in the UK institution.

Use of the POMS as a bed utilisation tool relies on the assumption that it captures all reasons for remaining in hospital. In this study, the main reason for remaining in hospital with no identifiable morbidity was 'on-going physiotherapy and occupational therapy input'. A specific concern in this patient group is that these patients may not have adequate mobility to be discharged safely. Including a specific domain for mobility may improve the sensitivity of the POMS for morbidity requiring hospitalisation following orthopaedic surgery. Criteria for a positive result could include inability to mobilise 10 metres or climb a single flight of stairs. Whilst this domain could be especially relevant for orthopaedic patients, this requires further investigation.

Use of the POMS as a "fitness for discharge" tool rests on the assumption that patients do not develop new morbidity after they have become free from morbidity, either in hospital or following discharge. No patients were readmitted to the study hospital in the first post-operative year for complications linked to surgery. However, 38 patients developed "new" morbidity following a period without morbidity whilst in hospital. This highlights a limitation of prospectively using the POMS as a "fitness for discharge" tool. To overcome this potential problem primary arthroplasty patients should have regular would reviews until POD 5 and revision arthroplasty patients until POD 8. This could be done on an out-patient basis. Patients should be aware of the risk of deep vein thrombosis, pulmonary embolus, myocardial infarction and cerebro-

vascular accident following discharge and receive clear written instructions regarding symptoms and management. As long as these precautionary measures are in place, POMS has potential as a bed utilisation tool.

The most commonly used tool to assess appropriate bed utilisation in the literature is the Appropriateness Evaluation Protocol (AEP)[10]. The AEP is a retrospective tool that relies on data from the inpatient record. It has been shown to be valid and reliable in some studies[10] but not in others[11]. The POMS is a prospective tool that could be used in real time to assist with appropriate patient discharge. The AEP is a retrospective tool that can only be used to evaluate past events. Data for the POMS is collected directly from contemporary data sources whilst the patient is in hospital; the AEP relies solely on past patient records and is therefore dependent on completeness and accuracy of record keeping for reliable functioning.

The AEP has been used in several European countries to examine bed utilization. In Portugal 50% of inpatient days were deemed inappropriate[12], in Italy 37.3%[13], in Germany 28%[14], in Switzerland 8-15%[15] and in France 7%[16]. This study indicates bed utilisation in the UK is comparable to that seen in Portugal and Italy but such a direct comparison may have limited validity since different bed utilisation tools have been used.

The finding that many fewer patients remain in hospital with no morbidity (as defined by the POMS) in the US when compared with the UK suggests that bed utilisation in the US is superior to that seen in the UK. The implementation of ‘payment by

results' in the UK aims to improve appropriate bed occupancy to optimise patient care and improve efficiency. If the patients in this study had been discharged when they first had no morbidity defined by the POMS, a saving of over £750,000 could have been made in one year (based on a cost of £400 per bed-day).

We believe that the POMS is able to identify patients remaining in hospital without clinically significant morbidity and may be used prospectively as a bed utilisation tool. To use the survey for this purpose, it may be useful to add an additional measure to assess mobility.

References

1 **Department of Health**. The NHS Plan. 1 July 2000.

2 **Department of Health.** Response to Reforming NHS Financial Flows: Introducing Payment by Results: Response issued 10 February 2003.

3 **Angelillo IF, Ricciardi G, Nante N et al.** Appropriateness of hospital utilisation in Italy. *Public Health* 2000 January; 114 (1): 9-14.

4 **Alijani A, Hanna GB, Ziyaie D et al.** Instrument for objective assessment of appropriateness of surgical bed occupancy: validation study. *BMJ* 2003 June 7; 326 (7401): 1243-4.

5 **Bennett-Guerrero E, Welsby I, Dunn TJ et al**. The use of a postoperative morbidity survey to evaluate patients with prolonged hospitalization after routine, moderate-risk, elective surgery. *Anesth Analg* 1999 August;89(2):514-9.

6 **MPW Grocott, JP Browne, J Van der Meulen, C Matejowsky, M Mutch, MA Hamilton, DZH Levett, M. Emberton, FS Haddad, MG Mythen.** The Post-Operative Morbidity Survey was validated and used to describe morbidity following major surgery. Journal of Clinical Epidemiology 2007;60:919-928

7 **Bennett-Guerrero E, Feierman DE, Barclay GR et al**. Preoperative and intraoperative predictors of postoperative morbidity, poor graft function, and early rejection in 190 patients undergoing liver transplantation. *Arch Surg* 2001 October;136(10):1177-83.

8 **Wakeling HG, McFall MR, Jenkins CS et al**. Intraoperative oesophageal Doppler guided fluid management shortens postoperative hospital stay after major bowel surgery. *Br J Anaesth* 2005 November;95(5):634-42.

9 **Webb A**. Solutions for reducing length of stay. *Health Service Journal* 14/1/2008

10 **Gertman PM, Restuccia JD.** The appropriateness evaluation protocol: a technique for assessing unnecessary days of hospital care. *Med Care* 1981; 19**:** 855–871.

11 **Kalant N, Berlinquet M, Diodati JG, Dragatakis L, Marcotte F**. How valid are utilization review tools in assessing appropriate use of acute care beds? *CMAJ*. 2000 Jun 27; 162 (13): 1809-13.

12 **Bentes M, Gonsalves ML, Santos M, Pina E.** Design and development of a utilization review program in Portugal. *Int J Qual Health Care* 1995; **7:** 201–212.

13 **Angelillo IF, Ricciardi G, Nante N, Boccia A, Group AC.** Appropriateness of hospital utilization in Italy. *Publ Health* 2000; 114: 9–14.

14 **Sangha O, Schneeweiss S, Wildner M, Cook EF, Brennan TA, Witte J, Liang MH.** Metric properties of the appropriateness evaluation protocol and predictors of inappropriate hospital use in Germany: an approach using longitudinal patient data. *Int J Qual Health Care.* 2002 Dec; 14(6): 483-92.

15 **Santos-Eggimann B, Paccaud F, Blanc T**. Medical appropriateness of hospital utilization: an overview of the Swiss experience. *Int J Qual Health Care* 1995; **7:** 227–232.

16 **d'Alche-Gautier M, Maiza D, Chastang F**. Assessing the appropriateness of hospitalisation days in a French university hospital. *Int J of Health Care Qual Assurance.* 2004; 17 (2): 87-91.

Table 1: Criteria for a positive POMS score

Variable	Criteria for positive result
Pulmonary	Requires supplementary oxygen or ventilatory support
Infection	Currently on antibiotics or temperature >38°C in the last 24 hours
Renal	Oliguria (<500ml/day), elevated creatinine (>30% pre-op level), catheter in-situ (for non-surgical reason)
Gastrointestinal	Unable to tolerate enteral diet for any reason
Cardiovascular	Diagnostic tests or treatment within the last 24 hours for: myocardial infarction, hypotension (requiring pharmacological therapy or fluids >200ml/hour), atrial/ventricular arrhythmia or cardiogenic pulmonary oedema
Central nervous system	Presence of new focal deficit, coma, confusion, delirium
Wound complications	Wound dehiscence requiring surgical exploration or drainage of pus from operative wound with or without isolation of organisms
Haematological	Requirement of blood transfusion, platelets, fresh frozen plasma or cryoprecipitate within the last 24 hours
Pain	Wound pain requiring parenteral opioids or regional anaesthesia

Table 2 Demographics of study population

	Number performed	Mortality rate / %	Age / years		ASA		% Female	Length of stay / days	
			Mean	Range	Median	Range		Median	Range
UKR	66	1	66.1	45-87	2	1-3	45	5	2-52
TKR	226	0	70.3	23-90	2	1-3	36	6	3-37
RTKR	8	0	71.6	46-88	2	1-3	25	13	3-102
BHR	32	0	51.6	22-70	1	1-3	50	6	4-13
THR	162	0	70.7	21-89	2	1-3	36	8	3-51
RTHR	35	3	72.2	26-88	2	1-3	36	14	6-93
TOTAL	529	0.4	68.9	21-90	2	1-3	38	7	2-102

Table 3 – Location of patients following hip arthroplasty procedures

			POD 3	POD 5	POD 8	POD 15
Procedure	**BHR**	Patients discharged	0/32 (0%)	2/32 (6%)	27/32 (84%)	32/32 (100%)
		Inpatients POMS >0	8/32 (25%)	5/32 (16%)	0/32 (0%)	0/32 (0%)
		Inpatients POMS = 0	24/32 (75%)	25/32 (78%)	5/32 (16%)	0/32 (0%)
	THR	Patients discharged	0/162 (0%)	13/162 (8%)	78/162 (48%)	138/162 (85%)
		Inpatients POMS >0	87/162 (54%)	62/162 (38%)	29/162 (18%)	13/162 (8%)
		Inpatients POMS = 0	75/162 (46%)	87/162 (54%)	55/162 (34%)	11/162 (7%)
	Revision THR	Patients discharged	0/35 (0%)	0/35 (0%)	3/35 (9%)	20/35 (57%)
		Inpatients POMS >0	31/35 (89%)	25/35 (71%)	21/35 (60%)	14/35 (40%)
		Inpatients POMS = 0	4/35 (11%)	10/35 (29%)	11/35 (31%)	1/35 (3%)
	TOTAL	Patients discharged	0/230 (0%)	16/230 (7%)	109/230 (47%)	191/230 (83%)
		Inpatients POMS >0	127/230 (55%)	92/230 (40%)	50/230 (22%)	27/230 (12%)
		Inpatients POMS = 0	103/230 (45%)	122/230 (53%)	71/230 (31%)	12/230 (5%)

Table 4 – Location of patients following knee arthroplasty procedures

			POD 3	POD 5	POD 8	POD 15
Procedure	**UKR**	Patients discharged	7/66 (11%)	33/66 (50%)	59/66 (89%)	65/66 (98%)
		Inpatients POMS >0	17/66 (26%)	7/66 (11%)	3/66 (5%)	1/66 (2%)
		Inpatients POMS = 0	42/66 (63%)	26/66 (39%)	4/66 (6%)	0/66 (0%)
	TKR	Patients discharged	0/226 (0%)	22/226 (10%)	145/226 (64%)	211/226 (93%)
		Inpatients POMS >0	114/226 (50%)	90/226 (40%)	38/226 (17%)	7/22 (3%)
		Inpatients POMS = 0	112/226 (50%)	114/226 (50%)	43/226 (19%)	8/226 (4%)
	Revision TKR	Patients discharged	0/8 (0%)	1/8 (13%)	1/8 (13%)	6/8 (75%)
		Inpatients POMS >0	6/8 (75%)	4/8 (50%)	5/8 (62%)	1/8 (12.5%)
		Inpatients POMS = 0	2/8 (25%)	3/8 (37%)	2/8 (25%)	1/8 (12.5%)
	TOTAL	Patients discharged	7/300 (2%)	56/300 (19%)	205/300 (68%)	282/300 (94%)
		Inpatients POMS >0	137/300 (46%)	101/300 (34%)	46/300 (15%)	9/300 (3%)
		Inpatients POMS = 0	156/300 (52%)	143/300 (47%)	49/300 (17%)	9/300 (3%)

Table 5 - Number of inappropriate inpatient days classified by type of arthroplasty

	Total number of patients	Total number of inappropriate inpatient days	Average number of inappropriate inpatient days per patient
BHR	33	78	2.36
THR	162	678	4.19
Revision THR	35	363	10.37
UKR	63	111	1.76
TKR	227	620	2.73
Revision TKR	8	115	14.38
Total	528	1965	3.72

Figure 1 – Discharge status and prevalence of morbidity following all types of hip arthroplasty

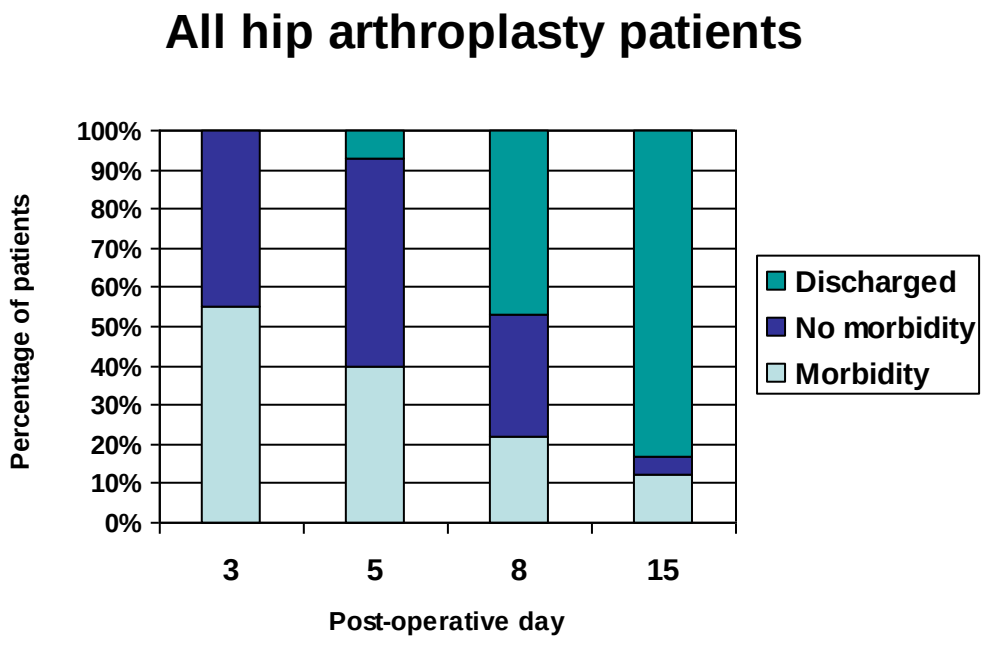


Figure 2 – Discharge status and prevalence of morbidity following all types of knee arthroplasty

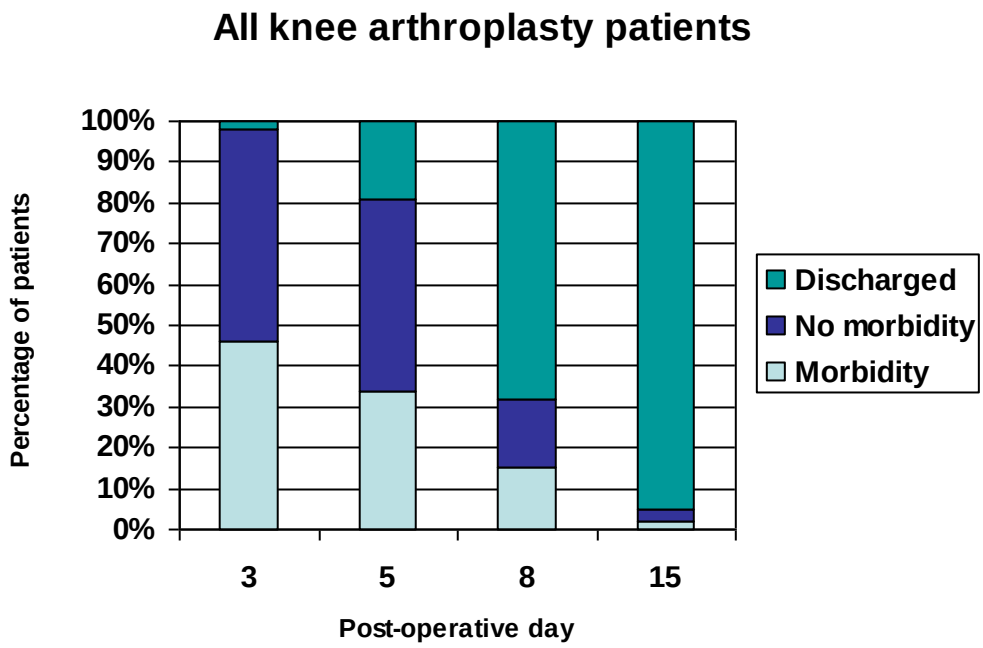


Figure 3 – Average number of inappropriate inpatient days classified by type of arthroplasty

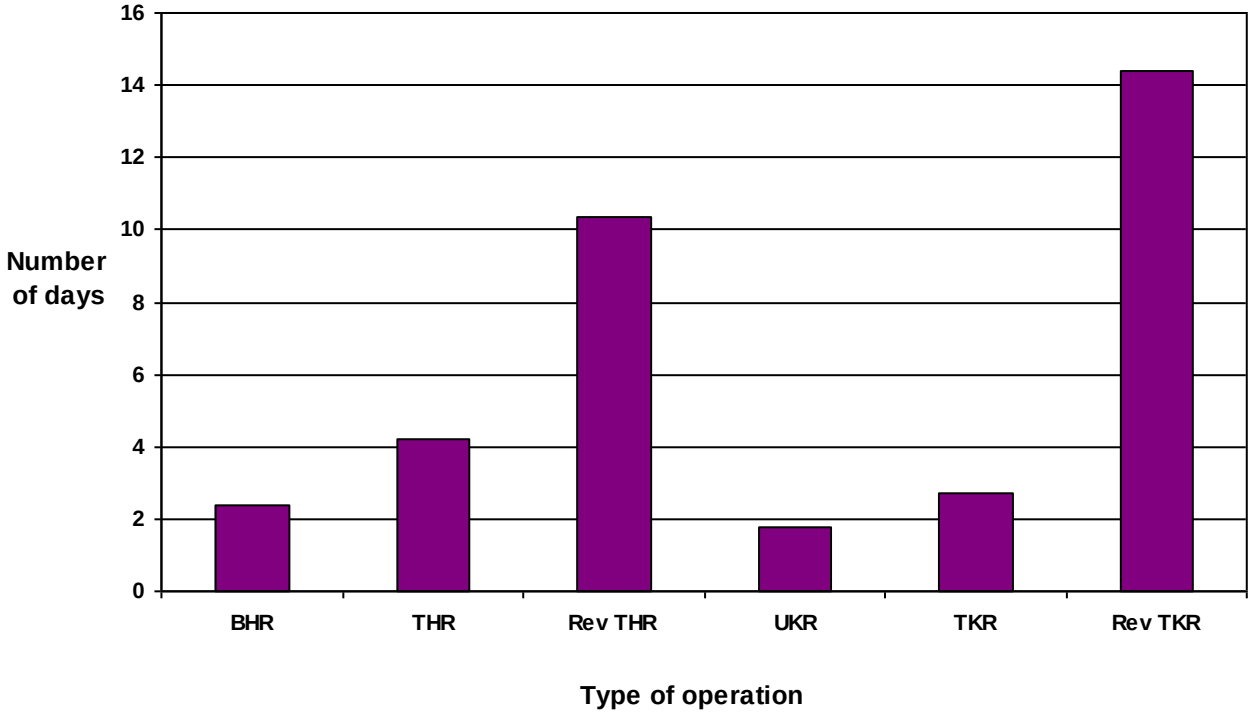
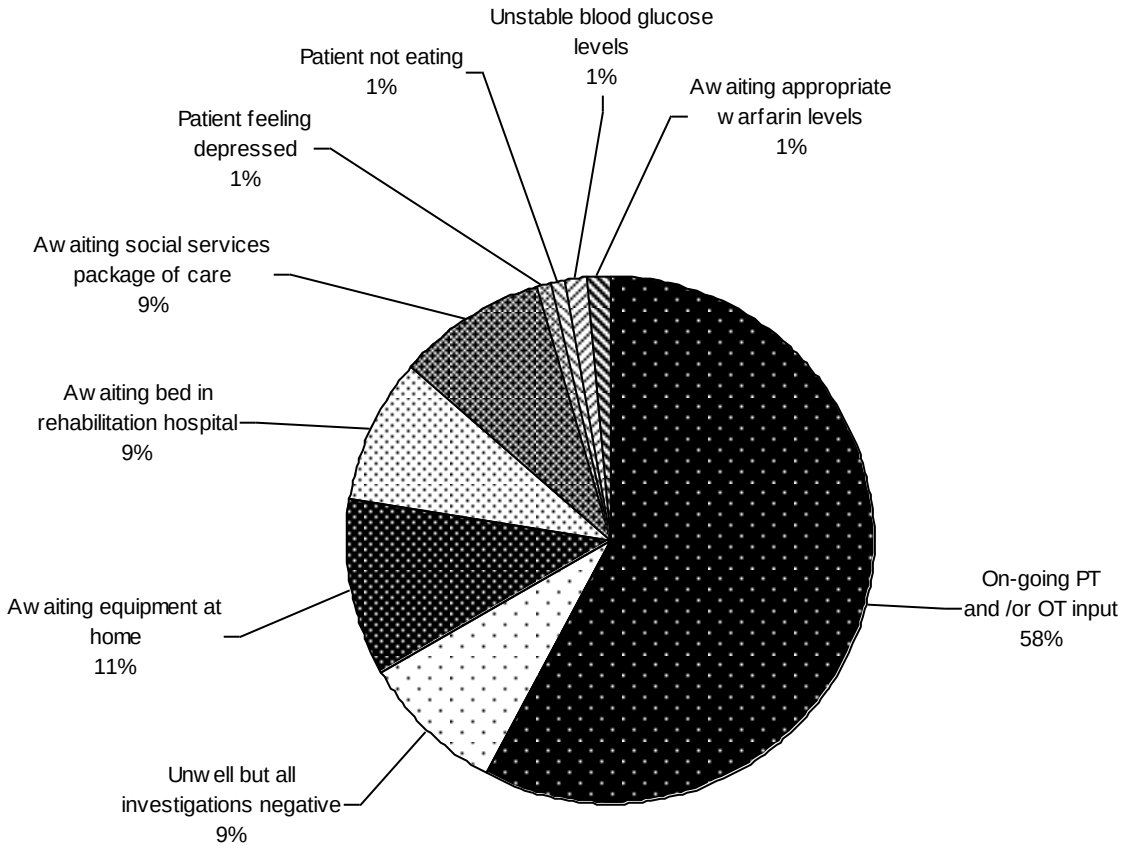


Figure 4 – Reasons lower limb arthroplasty patients with no morbidity remained in hospital on post-operative days 8 and 15



References

1. Weiser TG, Regenbogen SE, Thompson KD, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. *Lancet.* 2008;372(9633):139-144.
2. Wilson AP, Hodgson B, Liu M, et al. Reduction in wound infection rates by wound surveillance with postdischarge follow-up and feedback. *Br J Surg.* 2006;93(5):630-638.
3. Magee H, Davis LJ, Coulter A. Public views on healthcare performance indicators and patient choice. *J R Soc Med.* 2003;96(7):338-342.
4. Skinner TJ, Price BS, Scott DW, Gorry GA. Factors affecting the choice of hospital-based ambulatory care by the urban poor. *Am J Public Health.* 1977;67(5):439-445.
5. Fanjiang G, von Glahn T, Chang H, Rogers WH, Safran DG. Providing patients web-based data to inform physician choice: if you build it, will they come? *J Gen Intern Med.* 2007;22(10):1463-1466.
6. Government U. Response to Reforming NHS Financial Flows: Introducing Payment by Results. In: Health TDo, ed2003.
7. Darzi A. High Quality Care For All. Next Stage NHS Review. Final Report. In: DOH, edJune 30 2008.
8. Burack JH, Impellizzeri P, Homel P, Cunningham JN, Jr. Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. *Ann Thorac Surg.* 1999;68(4):1195-1200; discussion 1201-1192.
9. Iezzoni LI. An introduction to risk adjustment. *Am J Med Qual.* 1996;11(1):S8-11.
10. Saklad M. Grading of patients for surgical procedures. *Anesthesiology.* 1941;2:281-284.
11. Michel JP, Klopfenstein C, Hoffmeyer P, Stern R, Grab B. Hip fracture surgery: is the pre-operative American Society of Anesthesiologists (ASA) score a predictor of functional outcome? *Aging Clin Exp Res.* 2002;14(5):389-394.
12. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg.* 1991;78(3):355-360.
13. Lie SA, Pratt N, Ryan P, et al. Duration of the increase in early postoperative mortality after elective hip and knee replacement. *J Bone Joint Surg Am.* 2010;92(1):58-63.
14. Whiteley MS, Prytherch DR, Higgins B, Weaver PC, Prout WG. An evaluation of the POSSUM surgical scoring system. *Br J Surg.* 1996;83(6):812-815.
15. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. *Br J Surg.* 1998;85(9):1217-1220.
16. Mohamed K, Copeland GP, Boot DA, et al. An assessment of the POSSUM system in orthopaedic surgery. *J Bone Joint Surg Br.* 2002;84(5):735-739.
17. Daley J, Henderson WG, Khuri SF. Risk-adjusted surgical outcomes. *Annu Rev Med.* 2001;52:275-287.
18. Angelillo IF, Ricciardi G, Nante N, et al. Appropriateness of hospital utilisation in Italy. *Public Health.* 2000;114(1):9-14.

19. Bruce J, Russell EM, Mollison J, Krukowski ZH. The measurement and monitoring of surgical adverse events. *Health Technol Assess.* 2001;5(22):1-194.
20. Bennett-Guerrero Eea. The use of a postoperative morbidity survey to evaluate patients with prolonged hospitalization after routine, moderate-risk, elective surgery. *Anesth Analg* 1999;89:514-519
21. Grocott MP, Browne JP, Van der Meulen J, et al. The Postoperative Morbidity Survey was validated and used to describe morbidity after major surgery. *J Clin Epidemiol.* 2007;60(9):919-928.
22. Lebuffe G, Vallet B, Takala J, et al. A european, multicenter, observational study to assess the value of gastric-to-end tidal PCO2 difference in predicting postoperative complications. *Anesth Analg.* 2004;99(1):166-172.
23. Sanders J, Cooper J, Mythen MG, Montgomery HE. Predictors of total morbidity burden on days 3, 5 and 8 after cardiac surgery. *Perioper Med (Lond).* 2017;6:2.
24. West MA, Loughney L, Ambler G, et al. The effect of neoadjuvant chemotherapy and chemoradiotherapy on exercise capacity and outcome following upper gastrointestinal cancer surgery: an observational cohort study. *BMC Cancer.* 2016;16(1):710.
25. Bennett-Guerrero E, Feierman DE, Barclay GR, et al. Preoperative and intraoperative predictors of postoperative morbidity, poor graft function, and early rejection in 190 patients undergoing liver transplantation. *Arch Surg.* 2001;136(10):1177-1183.
26. Wakeling HG, McFall MR, Jenkins CS, et al. Intraoperative oesophageal Doppler guided fluid management shortens postoperative hospital stay after major bowel surgery. *Br J Anaesth.* 2005;95(5):634-642.
27. Ackland GL, Harris S, Ziabari Y, Grocott M, Mythen M, Investigators S. Revised cardiac risk index and postoperative morbidity after elective orthopaedic surgery: a prospective cohort study. *Br J Anaesth.* 2010;105(6):744-752.
28. Ackland GL, Moran N, Cone S, Grocott MP, Mythen MG. Chronic kidney disease and postoperative morbidity after elective orthopedic surgery. *Anesth Analg.* 2011;112(6):1375-1381.
29. Hollowell J, Grocott MP, Hardy R, Haddad FS, Mythen MG, Raine R. Major elective joint replacement surgery: socioeconomic variations in surgical risk, postoperative morbidity and length of stay. *J Eval Clin Pract.* 2010;16(3):529-538.
30. Grocott M. *Measuring Morbidity following Major Surgery*. London, UCL; 2008.
31. Ridgeway S, Wilson J, Charlet A, Kafatos G, Pearson A, Coello R. Infection of the surgical site after arthroplasty of the hip. *J Bone Joint Surg Br.* 2005;87(6):844-850.
32. Bori G, Muñoz-Mahamud E, Cuñé J, Gallart X, Fuster D, Soriano A. One-stage revision arthroplasty using cementless stem for infected hip arthroplasties. *J Arthroplasty.* 2014;29(5):1076-1081.
33. Brown RB, Bradley S, Opitz E, Cipriani D, Pieczarka R, Sands M. Surgical wound infections documented after hospital discharge. *Am J Infect Control.* 1987;15(2):54-58.

34. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol.* 1992;13(10):606-608.
35. Allami MK, Jamil W, Fourie B, Ashton V, Gregg PJ. Superficial incisional infection in arthroplasty of the lower limb. Interobserver reliability of the current diagnostic criteria. *J Bone Joint Surg Br.* 2005;87(9):1267-1271.
36. Wilson AP, Gibbons C, Reeves BC, et al. Surgical wound infection as a performance indicator: agreement of common definitions of wound infection in 4773 patients. *BMJ.* 2004;329(7468):720.
37. Wilson AP, Treasure T, Sturridge MF, Gruneberg RN. A scoring method (ASEPSIS) for postoperative wound infections for use in clinical trials of antibiotic prophylaxis. *Lancet.* 1986;1(8476):311-313.
38. Wilson AP, Webster A, Gruneberg RN, Treasure T, Sturridge MF. Repeatability of asepsis wound scoring method. *Lancet.* 1986;1(8491):1208-1209.
39. Wilson AP, Helder N, Theminimulle SK, Scott GM. Comparison of wound scoring methods for use in audit. *J Hosp Infect.* 1998;39(2):119-126.
40. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30(6):473-483.
41. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ.* 1992;305(6846):160-164.
42. McHorney CA, Ware JE, Jr., Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care.* 1993;31(3):247-263.
43. Feeny D, Wu L, Eng K. Comparing short form 6D, standard gamble, and Health Utilities Index Mark 2 and Mark 3 utility scores: results from total hip arthroplasty patients. *Qual Life Res.* 2004;13(10):1659-1670.
44. Mahomed NN, Arndt DC, McGrory BJ, Harris WH. The Harris hip score: comparison of patient self-report with surgeon assessment. *J Arthroplasty.* 2001;16(5):575-580.
45. Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. *Arch Phys Med Rehabil.* 1999;80(5):572-579.
46. Mangione CM, Goldman L, Orav EJ, et al. Health-related quality of life after elective surgery: measurement of longitudinal changes. *J Gen Intern Med.* 1997;12(11):686-697.
47. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996;34(3):220-233.
48. Jenkinson C, Fitzpatrick R, Argyle M. The Nottingham Health Profile: an analysis of its sensitivity in differentiating illness groups. *Soc Sci Med.* 1988;27(12):1411-1414.
49. Beaton DE, Bombardier C, Hogg-Johnson SA. Measuring health in injured workers: a cross-sectional comparison of five generic health status instruments in workers with musculoskeletal injuries. *Am J Ind Med.* 1996;29(6):618-631.
50. Brooks R. EuroQol: the current state of play. *Health Policy.* 1996;37(1):53-72.

51. Brazier J, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res.* 1993;2(3):169-180.
52. van Agt HM, Essink-Bot ML, Krabbe PF, Bonsel GJ. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. *Soc Sci Med.* 1994;39(11):1537-1544.
53. Luo N, Chew LH, Fong KY, et al. Validity and reliability of the EQ-5D self-report questionnaire in Chinese-speaking patients with rheumatic diseases in Singapore. *Ann Acad Med Singapore.* 2003;32(5):685-690.
54. Tidermark J, Bergstrom G, Svensson O, Tornkvist H, Ponzer S. Responsiveness of the EuroQol (EQ 5-D) and the SF-36 in elderly patients with displaced femoral neck fractures. *Qual Life Res.* 2003;12(8):1069-1079.
55. Bellamy N. *Osteoarthritis - an evaluative index for clinical trials*. Ontario, Canada, MacMaster University; 1982.
56. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol.* 1988;15(12):1833-1840.
57. Bachmeier CJ, March LM, Cross MJ, et al. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. *Osteoarthritis Cartilage.* 2001;9(2):137-146.
58. Stucki G, Sangha O, Stucki S, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. *Osteoarthritis Cartilage.* 1998;6(2):79-86.
59. Ostendorf M, van Stel HF, Buskens E, et al. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. *J Bone Joint Surg Br.* 2004;86(6):801-808.
60. Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. *J Rheumatol.* 1995;22(6):1193-1196.
61. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. *Arthritis Rheum.* 1989;32(1):37-44.
62. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis Rheum.* 1992;35(1):1-10.
63. Selman SW. Impact of total hip replacement on quality of life. *Orthop Nurs.* 1989;8(5):43-49.
64. Harris WH. 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. *J Bone Joint Surg Am.* 1969;51(4):737-755.
65. Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *J Bone Joint Surg Br.* 1972;54(1):61-76.
66. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989(248):13-14.
67. Mackinnon J, Young S, Baily RA. The St Georg sledge for unicompartmental

replacement of the knee. A prospective study of 115 cases. *J Bone Joint Surg Br.* 1988;70(2):217-223.
68. Wright JG, Rudicel S, Feinstein AR. Ask patients what they want. Evaluation of individual complaints before total hip replacement. *J Bone Joint Surg Br.* 1994;76(2):229-234.
69. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br.* 1996;78(2):185-190.
70. Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord.* 2003;4:10.
71. Amstutz HC, Thomas BJ, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am.* 1984;66(2):228-241.
72. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. *J Orthop Sports Phys Ther.* 1998;28(2):88-96.
73. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br.* 1998;80(1):63-69.
74. Crow R, Gage H, Hampson S, et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess.* 2002;6(32):1-244.
75. Soderman P, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? *Clin Orthop Relat Res.* 2001(384):189-197.
76. Dawson J, Fitzpatrick R, Frost S, Gundle R, McLardy-Smith P, Murray D. Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. *J Bone Joint Surg Br.* 2001;83(8):1125-1129.
77. Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement surgery. *Qual Health Care.* 1996;5(2):81-88.
78. Kreibich DN, Vaz M, Bourne RB, et al. What is the best way of assessing outcome after total knee replacement? *Clin Orthop Relat Res.* 1996(331):221-225.
79. Liow RY, Walker K, Wajid MA, Bedi G, Lennox CM. The reliability of the American Knee Society Score. *Acta Orthop Scand.* 2000;71(6):603-608.
80. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes.* 2003;1:17.
81. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS)--validation of a Swedish version. *Scand J Med Sci Sports.* 1998;8(6):439-448.
82. Roos EM, Roos HP, Lohmander LS. WOMAC Osteoarthritis Index--additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and MacMaster Universities. *Osteoarthritis Cartilage.* 1999;7(2):216-221.
83. Weale AE, Lee AS, MacEachern AG. High tibial osteotomy using a dynamic axial external fixator. *Clin Orthop Relat Res.* 2001(382):154-167.
84. Harcourt WG, White SH, Jones P. Specificity of the Oxford knee status

questionnaire. The effect of disease of the hip or lumbar spine on patients' perception of knee disability. *J Bone Joint Surg Br.* 2001;83(3):345-347.
85. Alijani A, Hanna GB, Ziyaie D, et al. Instrument for objective assessment of appropriateness of surgical bed occupancy: validation study. *BMJ.* 2003;326(7401):1243-1244.
86. Gertman PM, Restuccia JD. The appropriateness evaluation protocol: a technique for assessing unnecessary days of hospital care. *Med Care.* 1981;19(8):855-871.
87. Kalant N, Berlinguet M, Diodati JG, Dragatakis L, Marcotte F. How valid are utilization review tools in assessing appropriate use of acute care beds? *CMAJ.* 2000;162(13):1809-1813.
88. Bentes M, Gonsalves ML, Santos M, Pina E. Design and development of a utilization review program in Portugal. *Int J Qual Health Care.* 1995;7(3):201-212.
89. Sangha O, Schneeweiss S, Wildner M, et al. Metric properties of the appropriateness evaluation protocol and predictors of inappropriate hospital use in Germany: an approach using longitudinal patient data. *Int J Qual Health Care.* 2002;14(6):483-492.
90. Santos-Eggimann B, Paccaud F, Blanc T. Medical appropriateness of hospital utilization: an overview of the Swiss experience. *Int J Qual Health Care.* 1995;7(3):227-232.
91. d'Alche-Gautier MJ, Maiza D, Chastang F. Assessing the appropriateness of hospitalisation days in a French university hospital. *Int J Health Care Qual Assur Inc Leadersh Health Serv.* 2004;17(2-3):87-91.
92. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297-234.
93. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-310.
94. Sinha S, Ashby E, Jayaram R, Grocott MP. Quality of reporting in randomized trials published in high-quality surgical journals. *J Am Coll Surg.* 2009;209(5):565-571 e561.
95. Nikolaou VS, Edwards MR, Bogoch E, Schemitsch EH, Waddell JP. A prospective randomised controlled trial comparing three alternative bearing surfaces in primary total hip replacement. *J Bone Joint Surg Br.* 2012;94(4):459-465.
96. Jolles BM, Grzesiak A, Eudier A, et al. A randomised controlled clinical trial and gait analysis of fixed- and mobile-bearing total knee replacements with a five-year follow-up. *J Bone Joint Surg Br.* 2012;94(5):648-655.
97. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA.* 1996;276(8):637-639.
98. Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: can we do better? *Ann Surg.* 2006;244(5):663-667.
99. Ioannidis JP, Evans SJ, Gotzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med.* 2004;141(10):781-788.

100. Shimmin A. The effect of operative volume on the outcome of hip resurfacing. Paper #316. . American Academy of Orthopaedic Surgeons 76th Annual Meeting. Feb. 25-28, 2009. Las Vegas. ; 2009.
101. DH Williams UM, MN Norton. Metal-on-metal hip resurfacing: the effect of component position and size on the range of motion to impingement. *J Bone Joint Surg Br* 2006;88-B(SUPP III):432.
102. Cohen D. How safe are metal-on-metal hip implants? *BMJ.* 2012;344:e1410.
103. Mäkelä KT, Visuri T, Pulkkinen P, et al. Risk of cancer with metal-on-metal hip replacements: population based study. *BMJ.* 2012;345:e4646.
104. van der Weegen W, Brakel K, Horn RJ, et al. Asymptomatic pseudotumours after metal-on-metal hip resurfacing show little change within one year. *Bone Joint J.* 2013;95-B(12):1626-1631.
105. Registry NJ. NJR 8th Annual Report. www.njrcentre.org.uk 2012.
106. Lavigne M, Therrien M, Nantel J, Roy A, Prince F, Vendittoli PA. The John Charnley Award: The functional outcome of hip resurfacing and large-head THA is the same: a randomized, double-blind study. *Clin Orthop Relat Res.* 2010;468(2):326-336.
107. Bolland BJ, Whitehouse SL, Timperley AJ. Indications for early hip revision surgery in the UK--a re-analysis of NJR data. *Hip Int.* 2012;22(2):145-152.
108. Starks I, Gregory J, Phillips S. Revision hip arthroplasty in nonagenarians. *Acta Orthop Belg.* 2010;76(6):766-770.
109. Goodman SB, Hwang K, Imrie S. High complication rate in revision total hip arthroplasty in juvenile idiopathic arthritis. *Clin Orthop Relat Res.* 2014;472(2):637-644.
110. Parvizi J, Pour AE, Keshavarzi NR, D'Apuzzo M, Sharkey PF, Hozack WJ. Revision total hip arthroplasty in octogenarians. A case-control study. *J Bone Joint Surg Am.* 2007;89(12):2612-2618.
111. Reilly KA, Beard DJ, Barker KL, Dodd CA, Price AJ, Murray DW. Efficacy of an accelerated recovery protocol for Oxford unicompartmental knee arthroplasty--a randomised controlled trial. *Knee.* 2005;12(5):351-357.
112. Laurencin CT, Zelicof SB, Scott RD, Ewald FC. Unicompartmental versus total knee arthroplasty in the same patient. A comparative study. *Clin Orthop Relat Res.* 1991(273):151-156.
113. Robertsson O, Borgquist L, Knutson K, Lewold S, Lidgren L. Use of unicompartmental instead of tricompartmental prostheses for unicompartmental arthrosis in the knee is a cost-effective alternative. 15,437 primary tricompartmental prostheses were compared with 10,624 primary medial or lateral unicompartmental prostheses. *Acta Orthop Scand.* 1999;70(2):170-175.
114. Rougraff BT, Heck DA, Gibson AE. A comparison of tricompartmental and unicompartmental arthroplasty for the treatment of gonarthrosis. *Clin Orthop Relat Res.* 1991(273):157-164.
115. Newman JH, Ackroyd CE, Shah NA. Unicompartmental or total knee replacement? Five-year results of a prospective, randomised trial of 102 osteoarthritic knees with unicompartmental arthritis. *J Bone Joint Surg Br.* 1998;80(5):862-865.
116. Seng C, Yeo SJ, Wee JL, Subanesh S, Chong HC, Lo NN. Improved clinical outcomes after high-flexion total knee arthroplasty: a 5-year follow-up study. *J Arthroplasty.* 2011;26(7):1025-1030.

117. Stiehl JB, Komistek RD, Dennis DA, Paxson RD, Hoff WA. Fluoroscopic analysis of kinematics after posterior-cruciate-retaining knee arthroplasty. *J Bone Joint Surg Br.* 1995;77(6):884-889.
118. Banks SA, Markovich GD, Hodge WA. In vivo kinematics of cruciate-retaining and -substituting knee arthroplasties. *J Arthroplasty.* 1997;12(3):297-304.
119. Beight JL, Yao B, Hozack WJ, Hearn SL, Booth RE, Jr. The patellar "clunk" syndrome after posterior stabilized total knee arthroplasty. *Clin Orthop Relat Res.* 1994(299):139-142.
120. Maloney WJ, Schurman DJ. The effects of implant design on range of motion after total knee arthroplasty. Total condylar versus posterior stabilized total condylar designs. *Clin Orthop Relat Res.* 1992(278):147-152.
121. Feinglass J, Koo S, Koh J. Revision total knee arthroplasty complication rates in Northern Illinois. *Clin Orthop Relat Res.* 2004(429):279-285.
122. Kakkos SK, Warwick D, Nicolaides AN, Stansby GP, Tsolakis IA. Combined (mechanical and pharmacological) modalities for the prevention of venous thromboembolism in joint replacement surgery. *J Bone Joint Surg Br.* 2012;94(6):729-734.
123. Pedersen AB, Johnsen SP, Sorensen HT. Increased one-year risk of symptomatic venous thromboembolism following total hip replacement: A nationwide cohort study. *J Bone Joint Surg Br.* 2012;94(12):1598-1603.
124. Pour AE, Keshavarzi NR, Purtill JJ, Sharkey PF, Parvizi J. Is Venous Foot Pump Effective In Prevention of Thromboembolic Disease After Joint Arthroplasty: A Meta-Analysis. *J Arthroplasty.* 2012.
125. Shimoyama Y, Sawai T, Tatsumi S, et al. Perioperative risk factors for deep vein thrombosis after total hip arthroplasty or total knee arthroplasty. *J Clin Anesth.* 2012;24(7):531-536.
126. Singh JA, Jensen MR, Lewallen DG. Patient factors predict periprosthetic fractures after revision total hip arthroplasty. *J Arthroplasty.* 2012;27(8):1507-1512.
127. Della Rocca GJ, Leung KS, Pape HC. Periprosthetic fractures: epidemiology and future projections. *J Orthop Trauma.* 2011;25 Suppl 2:S66-70.
128. Streit MR, Merle C, Clarius M, Aldinger PR. Late peri-prosthetic femoral fracture as a major mode of failure in uncemented primary hip replacement. *J Bone Joint Surg Br.* 2011;93(2):178-183.
129. Poultsides LA, Ma Y, Della Valle AG, Chiu YL, Sculco TP, Memtsoudis SG. In-Hospital Surgical Site Infections after Primary Hip and Knee Arthroplasty - Incidence and Risk Factors. *J Arthroplasty.* 2012.
130. Lopez-Contreras J, Limon E, Matas L, Olona M, Salles M, Pujol M. Epidemiology of surgical site infections after total hip and knee joint replacement during 2007-2009: a report from the VINCat Program. *Enferm Infecc Microbiol Clin.* 2012;30 Suppl 3:26-32.
131. Kosashvili Y, Backstein D, Safir O, Lakstein D, Gross AE. Dislocation and infection after revision total hip arthroplasty: comparison between the first and multiply revised total hip arthroplasty. *J Arthroplasty.* 2011;26(8):1170-1175.
132. Wallace G, Judge A, Alhambra DP, de Vries F, Arden NK, Cooper C. The effect of body mass index on the risk of post-operative complications during the six months following total hip replacement or total knee replacement surgery. *Osteoarthritis Cartilage.* 2014.

133. Mantilla CB, Horlocker TT, Schroeder DR, Berry DJ, Brown DL. Frequency of myocardial infarction, pulmonary embolism, deep venous thrombosis, and death following primary hip or knee arthroplasty. *Anesthesiology.* 2002;96(5):1140-1146.
134. Parvizi J, Mui A, Purtill JJ, Sharkey PF, Hozack WJ, Rothman RH. Total joint arthroplasty: When do fatal or near-fatal complications occur? *J Bone Joint Surg Am.* 2007;89(1):27-32.
135. Pulido L, Ghanem E, Joshi A, Purtill JJ, Parvizi J. Periprosthetic joint infection: the incidence, timing, and predisposing factors. *Clin Orthop Relat Res.* 2008;466(7):1710-1715.
136. Lalmohamed A, Vestergaard P, Klop C, et al. Timing of acute myocardial infarction in patients undergoing total hip or knee replacement: a nationwide cohort study. *Arch Intern Med.* 2012;172(16):1229-1235.
137. Pedersen AB, Mehnert F, Sorensen HT, Emmeluth C, Overgaard S, Johnsen SP. The risk of venous thromboembolism, myocardial infarction, stroke, major bleeding and death in patients undergoing total hip and knee replacement: a 15-year retrospective cohort study of routine clinical practice. *Bone Joint J.* 2014;96-B(4):479-485.
138. Lalmohamed A, Vestergaard P, Cooper C, et al. Timing of stroke in patients undergoing total hip replacement and matched controls: a nationwide cohort study. *Stroke.* 2012;43(12):3225-3229.
139. Mortazavi SM, Kakli H, Bican O, Moussouttas M, Parvizi J, Rothman RH. Perioperative stroke after total joint arthroplasty: prevalence, predictors, and outcome. *J Bone Joint Surg Am.* 2010;92(11):2095-2101.
140. Verlicchi F, Desalvo F, Zanotti G, Morotti L, Tomasini I. Red cell transfusion in orthopaedic surgery: a benchmark study performed combining data from different data sources. *Blood Transfus.* 2011;9(4):383-387.
141. Evans S, O'Loughlin E, Bruce J. Retrospective audit of blood transfusion and comparison with haemoglobin concentration in patients undergoing elective primary and revision lower limb arthroplasty. *Anaesth Intensive Care.* 2011;39(3):480-485.
142. Gombotz H, Rehak PH, Shander A, Hofmann A. The second Austrian benchmark study for blood use in elective surgery: results and practice change. *Transfusion.* 2014.
143. Sivrikoz N, Koltka K, Güresti E, Büget M, Sentürk M, Özyalçın S. Perioperative dexketoprofen or lornoxicam administration for pain management after major orthopedic surgery: a randomized, controlled study. *Agri.* 2014;26(1):23-28.
144. Judd DL, Dennis DA, Thomas AC, Wolfe P, Dayton MR, Stevens-Lapsley JE. Muscle strength and functional recovery during the first year after THA. *Clin Orthop Relat Res.* 2014;472(2):654-664.
145. Horlocker TT. Pain management in total joint arthroplasty: a historical review. *Orthopedics.* 2010;33(9 Suppl):14-19.
146. Petre BM, Roxbury CR, McCallum JR, Defontes KW, Belkoff SM, Mears SC. Pain reporting, opiate dosing, and the adverse effects of opiates after hip or knee replacement in patients 60 years old or older. *Geriatr Orthop Surg Rehabil.* 2012;3(1):3-7.
147. Health Do. The NHS Plan. 2000.
148. webb A. Solutions for reducing length of stay. *Health Service Journal.* 2008.

149. McNicholas MJ, Pengas IP, Assiotis A, Nash W, Hatcher J, Banks J. Total meniscectomy in adolescents: A 40-year follow-up. *J Bone Joint Surg Br.* 2012;94(12):1649-1654.
150. Howells NR, Eldridge JD. Medial patellofemoral ligament reconstruction for patellar instability in patients with hypermobility: A case control study. *J Bone Joint Surg Br.* 2012;94(12):1655-1659.
151. Moonesinghe SR, Harris S, Mythen MG, et al. Survival after postoperative morbidity: a longitudinal observational cohort study†. *Br J Anaesth.* 2014.
152. Khuri SF, Henderson WG, DePalma RG, et al. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Ann Surg.* 2005;242(3):326-341; discussion 341-323.
153. Grocott M. *Measuring Morbidity following Major Surgery*. London: Medicine, UCL; 2008.
154. Hooper GJ, Rothwell AG, Hooper NM, Frampton C. The relationship between the American Society Of Anesthesiologists physical rating and outcome following total hip and knee arthroplasty: an analysis of theNew Zealand Joint Registry. *J Bone Joint Surg Am.* 2012;94(12):1065-1070.
155. Santaguida PL, Hawker GA, Hudak PL, et al. Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review. *Can J Surg.* 2008;51(6):428-436.
156. Singh JA, Lewallen D. Age, gender, obesity, and depression are associated with patient-related pain and function outcome after revision total hip arthroplasty. *Clin Rheumatol.* 2009;28(12):1419-1430.
157. Hamilton DF, Lane JV, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open.* 2013;3(4).
158. Hamilton D, Henderson GR, Gaston P, MacDonald D, Howie C, Simpson AH. Comparative outcomes of total hip and knee arthroplasty: a prospective cohort study. *Postgrad Med J.* 2012;88(1045):627-631.
159. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am.* 2004;86-A(5):963-974.
160. Johnson AJ, Costa CR, Mont MA. Do we need gender-specific total joint arthroplasty? *Clin Orthop Relat Res.* 2011;469(7):1852-1858.
161. O'Connor MI. Implant survival, knee function, and pain relief after TKA: are there differences between men and women? *Clin Orthop Relat Res.* 2011;469(7):1846-1851.
162. Thomsen MG, Husted H, Bencke J, Curtis D, Holm G, Troelsen A. Do we need a gender-specific total knee replacement? A randomised controlled trial comparing a high-flex and a gender-specific posterior design. *J Bone Joint Surg Br.* 2012;94(6):787-792.
163. Khan RJ, Maor D, Hofmann M, Haebich S. A comparison of a less invasive piriformis-sparing approach versus the standard posterior approach to the hip: A randomised controlled trial. *J Bone Joint Surg Br.* 2012;94(1):43-50.
164. Spaans AJ, van den Hout JA, Bolder SB. High complication rate in the early experience of minimally invasive total hip arthroplasty by the direct anterior approach. *Acta Orthop.* 2012;83(4):342-346.

165. Kim YH, Park JW, Kim JS. Computer-navigated versus conventional total knee arthroplasty a prospective randomized trial. *J Bone Joint Surg Am.* 2012;94(22):2017-2024.
166. Burnett RS, Barrack RL. Computer-assisted total knee arthroplasty is currently of no proven clinical benefit: a systematic review. *Clin Orthop Relat Res.* 2013;471(1):264-276.
167. Lau RL, Perruccio AV, Gandhi R, Mahomed NN. The role of surgeon volume on patient outcome in total knee arthroplasty: a systematic review of the literature. *BMC Musculoskelet Disord.* 2012;13:250.
168. Siah RCJ, Childs C. A systematic review of surgical infection scoring systems used in surgical patients. . *JBI Library of Systematic Reviews.* 2011;9(60):2627-2683.