An analysis of choice: a case study on hip prostheses

Charlotte Davies, MSc, BSc

Thesis submitted for the degree of Doctorate of Philosophy

University of East Anglia

School of Medicine, Health Policy and Practice

Date of submission: May 2011

Thesis length: 62,790 words

Charlotte Davies
May 2011
'An analysis of choice: a case study on hip prostheses'

Abstract

Total hip replacement (THR) surgery is a highly successful procedure offering relief of chronic pain and improving physical functioning. Given an ageing population, there is an ever increasing demand for THR, and an increasing need to establish its cost-effectiveness.

This thesis explores two aspects of choice between the alternative prostheses: how choices *should* be made, and what choices are *actually* made.

On the former, a key indicator is the long-term prosthesis survival rate. However, when choosing between prostheses, there is often insufficient evidence on long-term survival. The National Joint Registry (NJR) is an invaluable emerging source of information on this count. Using its Annual Reports, I identify, for example, that the use of cementless prostheses has grown rapidly, despite their performance in terms of **early** revision being inferior to the traditional cemented types.

However, the NJR was only introduced in 2003, and cannot yet provide information on **longer term** prosthesis survival. Previous research has attempted to predict long-term survival by forecasting from short-term data. I assess this approach by revisiting a well-known case-study, examining how well estimated survival curves predict what actually happened. I find that the predictions are very inaccurate, underlining the future value of the NJR as it accumulates more evidence.

On the latter, I employ raw NJR data to examine the actual choices between prostheses made by hospitals. Patients' characteristics explain little variation between hospitals with hospital characteristics appearing more important. I consider how choice might be affected by a highly concentrated oligopolistic manufacturing industry and find evidence of heterogeneous purchasing at the hospital level, consistent with a recent NAO report. I conclude that the NHS is not exploiting its potential buyer power, leaving itself susceptible to manufacturer seller power. I identify evidence potentially consistent with market sharing of regional and product markets by the manufacturers.

List of Contents

List of Tables and Figures

Abbreviations

AIC – Akaike's Information Criterion

BMA – British medical association

BMI – Body mass index

BMJ – British Medical Journal

CASP – Critical appraisal skills programme

CBA – Cost benefit analysis

CEA – Cost effectiveness analysis

CI – Confidence interval

CR2 – Two firm concentration ratio

CR5 – Five firm concentration ratio

CUA – Cost utility analysis

DARE - Database of abstracts of reviews of effects

DGCOMP – Director General for competition

DRG – Diagnostic related group

EC – European commission

EQ-5D – A measure of health outcome

EU – The European Union

EURONHEED - European Network of Health Economics Evaluation Databases

FT – Foundation trust

GDP – Gross domestic product

GP – General Practitioner

HEED – Health economic evaluations database

HES – Hospital Episode Statistics

HHI – Herfindahl Hirschman index

HQIP – Health quality information partnership

HR-QoL – Health related quality of life

IGS – Image guided surgery

IO – Industrial Organisation

ISTC – Independent sector treatment centre

J&J – Johnson and Johnson

JRI – Joint Replacement Instrumentation Ltd

KM curves – Kaplan-Meier curves

MHRA – Medicines and Healthcare products regulatory agency

MIS – Minimally invasive surgery

NHS – National Health Service

NHS EED – NHS economic evaluation database

NHS TC – NHS treatment centre

NJR – The National Joint Registry

NRES – National research ethics service

NSPB – National Specialisation at the brand level

NSPM – National Specialisation at the manufacturer level

OA – Osteoarthritis

ODEP – Orthopaedic data evaluation panel

OFT – Office of fair trading

OPCS-4 - Office of population, censuses and surveys classification of surgical operations and procedures

PbR – Payment by Results

PCT – Primary Care Trust

PPRS – Pharmaceutical price regulation scheme

PROMs – Patient Reported Outcome Measure

QALY – Quality adjusted life years

R&D – Research and development

RCT – Randomised controlled trial

SCAG – Security and confidentiality advisory group

SHA – Strategic Health Authority

SHAR – Swedish hip arthroplasty register

SMEs – Small and medium enterprises

SPB – Specialisation at the brand level

SPM – Specialisation at the manufacturer level

SUR – Seemingly unrelated regression

THR – Total Hip Replacement

UK – United Kingdom

US – United States of America

Acknowledgements

This thesis is dedicated to Ella, Mum & Dad

I would like to thank all of my supervisors:
Professor Alexander MacGregor who has been a huge support for clinical advice, data handling and guidance and encouragement for future career directions
Professor Ruth Hancock who took over at a late stage and has kept me on target.
Professor Miranda Mugford for her guidance and support in the earlier stages of my PhD
and I would especially like to give specific thanks to **Dr Paula Lorgelly**, a superb supervisor, and without her constant support and guidance (no matter where she was based) it simply would not have been possible to complete this thesis. Thanks Paula!

Thanks also to the following for their contributions at various stages in the thesis:
Professor Andrew Briggs
Mr Ian Shemilt
Professor Henrik Malchau
Michael Borroff

I would also like to thank the patients and staff of all the hospitals in England and Wales who have contributed data to the National Joint Registry. I am grateful to the Healthcare Quality Improvement Partnership (HQIP), the NJR Steering Committee and staff at the NJR Centre.
The author has conformed to the NJR's standard protocol for data access. The views 'or its contents', expressed represent those of the authors and do not necessarily reflect those of the National Joint Register Steering Committee or the Health Quality Improvement Partnership (HQIP) who do not vouch for how the information is presented.

I would like to thank the MRC for funding and supporting this PhD studentship.

Finally and very importantly, I would like to especially thank my family:
Thanks to Barney, Mum and Dad for their continuous support and for the help with proof reading. Thanks also to my fantastic daughter Ella for being Ella! Also thanks to Frankie and Simon, George and Amanda and close friends, Jo, Clare and Jane.

Chapter 1, Introduction

1.1 Motivation and Objectives

Musculoskeletal conditions (of which joint disease is one) are the most common cause of severe long-term pain and physical disability[3], and Osteoarthritis of the hip is one of the most common causes of disability in the Western world[4], due to a wearing out of the hip joint. For those with end stage joint disease of the hip, Total Hip Replacement surgery (THR) offers the only effective treatment.

The number of people over the age of 50 years is predicted to double over the next decade, and this will inevitably lead to an increasing incidence of diseases associated with old age, such as joint disease[5]. This ageing population, and increasing presence of disability, places a considerable and rising financial burden on healthcare budgets[6]. Beyond the ageing process, there have been various genetic and constitutional risk factors identified for joint disease in the literature: Obesity is a major risk factor for OA of the knee for middle aged women, alongisde mechanical risk factors such as weightbearing in sport and occupations which involves excessive bending, such as farming. Diet also has been cited as playing a role in the incidence of OA in terms of levels of vitamin D[7]. Moreover, patients with end stage joint disease typically experience chronic pain and loss of physical function which is a drain not only on National Health Service (NHS) resources but also on society at large in terms of lost productivity and an increased burden on domiciliary/informal care.

Necessarily, healthcare providers will be required to make cost-effective resource allocation decisions, and this involves choices. These choices relate not only to the aggregate budget on hip replacement as a whole, as opposed to other areas of healthcare, but also within the hip replacement budget. This thesis explores an important aspect of choice: the decision process regarding which of the many alternative prostheses for THR surgery should be/is chosen for individual patients. This is of relevance not only for the patient him/herself, but also for resource allocation within the overall healthcare budget.

The main of objectives of this thesis are twofold:

a). To examine an important factor in how choices between prostheses *should* be made, bearing in mind that information is scarce on the long-term survival rates of different prostheses.

b). To examine what choices are *actually* being made, and to investigate what this reveals about the buyer-seller relationship between the NHS and the large multinational firms who manufacture the prostheses.

To answer these questions, I use applied econometric methods to analyse data taken from two national hip registries: Sweden, and England and Wales, employing models which have their roots in the academic literatures of economic evaluation, public procurement and Industrial Organisation.

This opening chapter first provides a brief background on THR surgery, hip prostheses and the national joint registries, and then introduces the main academic disciplines underpinning the thesis: economic evaluation; Industrial Organisation and public procurement.

1.2 Total Hip Replacement surgery (THR)

THR surgery was first successfully performed in 1962 in England and is now one of the most frequently performed surgical procedures in the world[8], with over 72,432 operations carried out in England and Wales in 2008/9[9], with the number almost doubling in the last decade.

The care pathway for THR is described by the 'Map of Medicine' for the NHS[10]. It starts in primary care with the GP, who assesses the patient's status. At this point, a decision is made about whether to manage the hip discomfort using excericse, weight loss, pain relief and adaptive aids for self-care, or whether to consider a referral to secondary care (orthopaedic surgeon). A referral to the surgeon should be made particulary if non-surgical treatment has not been beneficial[11] or if the patient has a poor functional status. The surgeon will take a history and examine the patient before recommending surgery and within this, discuss treatment options with the patient including the option of no surgery. Surgery should take place up to 18 weeks from the first appointment. The surgeon will also consider alternative techniques such as: joint preserving techniques as an alternative for younger patients with mild to moderate intra-articular degeneration; arthroscopic surgery; open surgery and pelvic and/or osteotomy. The next stage is consideration of the prosthesis selection, taking into account NICE guidance benchmarks and according to the 'Map of Medicine' "should be goverend by guidance by evidence of its effective performance, and if possible the performance of the operating team using it,

including published evidence from the NJR". The surgeon will also consider the cementing options i.e prosthesis type, the bearing surface and the option of hip resurfacing[12]. In terms of the surgery itself, the surgeon decides on the surgical approach (patient position - posterior or anterolateral) and the surgical techniques (such as whether to use minimially invasive surgery and would closure techniques). Complications arising from surgery include: mortality, infection, dislocation, DVT, pulmonary embolism, dislocation, inadequate fixation and fracture around the implant site. If a complication is suspected, a patient will be radiographically examined to see if further surgery or medical follow-up is required. If no complications arise from surgery then the patient will require early mobilisation, rehabilitation and be discharged with follow up at 6 weeks post-surgery and x-rays at 1 and 5 year time points and subsequently, every 5 years[10]. No clinical pathway is available for revision surgery from the NHS Map of Medicine or in general NHS literature. However, the implication is that prosthesis failure is detected by the surgeon, either through the 5 yearly follow-up X-rays or due to the patient self-reporting indications for prosthesis failure. These indications for prosthesis failure could be identified by the patient in the form of reduced mobility or increased pain which may be a sign of prosthesis loosening or wear and tear of the artificial joint. Other less common indications for prosthesis failure include: infection, dislocation, joint stiffening and blood clots[13]. Once prosthesis failure is detected, the primary treatment option is revision surgery, although alternative options include: resection arthroplasty or fusion/arthrodesis. However, these proceedures are not commonly used as they are reported to require more complex surgical teachniques with often poor clinical outcome[14].

The hip joint itself is made up of a ball and socket, the ball is the top of the thigh bone (the femur) which fits into the socket. THR surgery involves removing parts of the bones which make up the hip joint and replacing them with an artificial joint, which will hereafter be referred to as a ‘prosthesis’ (also referred to as a hip ‘implant’ in the literature). During surgery, the surgeon will saw off the top end of the femur and hammer a metal or ceramic ball on a stem into the femur in place of the removed piece of bone. The hip socket is then drilled into in order to provide a shallow cup for the ball joint to fit into and lined with a material such as polyethylene, metal or ceramic. The joint can be fixed into place with cement to fix the prosthesis to the bone, or it can be secured by other methods of fixation, such as the bone growing into or onto the surface of the component.

THR surgery is widely accepted as a highly successful surgical procedure[4] which has led to a huge growth in the development of the prostheses used for surgery. One of the drivers for this has been to improve late failure rates. The total cost of joint replacement surgery to the NHS in the UK in 2000 was approximately £140 million[15], (£182 million in 2010

prices)[16] with the direct hospital costs of each procedure ranging from £488 to £9,905, and a mean of £4,788[15] (2008 prices). The anticipated cost savings of total joint replacement surgery (relative to no surgery) include the reduced costs of arthritis treatment, medication and community care. The direct and indirect benefits include improved quality of life of the patient (and their families) and increased productivity of the nation's workforce.

A key issue is the durability of the prostheses – how long it will survive before any further revision surgery (where the artificial joint is replaced) is needed. In the long-term, the primary reason for prosthesis failure is loosening of the prosthesis itself. As demand for the procedure has been increasing over time, so too has the associated demand for revision surgery. Revision surgery is technically difficult with reported inferior clinical outcomes[10], and unsurprisingly it is more costly than primary surgery. Prosthesis failure can occur in the immediate postoperative period[17], which is known as 'early failure' and is due to dislocation, primary deep infection and other technical problems[18]. Or it can occur anything up to 20-30 years post-surgery[17], which is known as 'late failure' because of long term wear of the artificial joint "*resulting from the production of prosthetic wear particles characterized by the formation of excessive granulation tissue and osteolysis*" p.941,[19]. Thus, the relative survival rates of different prostheses is of central importance[1] (alongside the patient's age and activity level) when deciding which prosthesis any given patient should receive.

In 1998 there were more than 60 alternative hip prostheses manufactured by 19 companies listed on the market in the UK[20], with total NHS expenditure on hip prostheses of approximately £53 million [15] (£69 million in 2010 prices). Ten years later, by 2008, the National Joint Registry (NJR)[1], listed 124 brands of acetabular cups and 137 brands of femoral stems - a substantial increase in the number of prostheses available from 1998 to 2008. Revision surgery has also increased, with 3,012 revision procedures carried out in 2003/4, rising to 6,581 by 2008/9 [1, 21] and accounting for approximately 9.4% of all THR procedures in England and Wales. Revision surgery is a key element of cost effectiveness, with Briggs et al[22] reporting a mean cost for a standard hip or knee revision procedure in 2000/1 as £5,294 (£6,385; 2008 prices) compared to £3,889 (£4,690; 2008 prices) for a primary procedure.

[1] Alongside the patient's age and activity levels

1.3 The National Joint Registry for England and Wales

There is very little high-quality evidence on the performance of some the alternative prosthesis types, particularly at the disaggregated level of the prosthesis brands[8]. This lack of high quality evidence (highlighted in the National Institute for Health and Clinical Excellence (NICE) guidance of 2000[23]) along with the well publicised problems documented about the early failure of the 3M Capital Hip brand[24] led to the establishment of the NJR for England and Wales in April 2003. It is the largest national joint registry in existence with the primary aim of monitoring the performance of joint prostheses and ensuring patients receive the best clinical care. The NJR records data on hip, knee and most recently, ankle surgery. It is further enhanced by the option to link it to HES (Hospital Episode Statistics) and PROMs (Patient Reported Outcome Measure). HES[25] is a records based system, established in 1989 to collect data on all admissions to NHS hospitals in England including information on: diagnoses and operations; patient characteristics; administrative information, such as waiting time and date of admission and geographical information, such as patient residence and site of treatment. PROMs was introduced into the NHS in 2009 and involves the patient completing a questionnaire both before and after surgery[26]. The aim of PROMs is to use the patient perspective to inform decision-making at all levels of the NHS[26] and is currently in place for four elective procedures: hip and knee surgery, hernia repair and varicose veins. However, neither the NJR nor HES routinely collects data on indirect non-medical costs and resource use (such as patient productivity losses or out-of-pocket expenses)[2].

1.4 Relevant Economic Literatures

Given the broad objectives of this thesis – analysis of choice, both prescriptively and actually - the thesis will draw on a number of relevant academic literatures. This section briefly introduces the three which lie at the heart of the research.

1.4.1 Economic evaluation

In the context of this thesis, economic evaluation is concerned with evaluating the alternative hip prostheses to inform policy decisions regarding which prosthesis should be implanted, given the available resources[27]. Folland et al define 'health economics' as the study of how

[2] The NJR, HES and PROMS databases are managed by Northgate Information Solutions on behalf of the NHS.

resources are allocated to and within the health economy[27], and within this, clinical effectiveness is not sufficient in isolation for this evaluative purpose, costs must also be taken into account[28] and the comparative analysis of costs of alternative treatments of health care is common to all types of economic evaluation[29].

This section provides a brief overview of the literature on economic evaluation by drawing on the seminal text by Drummond et al[29] who define economic evaluation as:

"*the comparative analysis of alternative courses of action in terms of both their costs and consequences*" (p. 8 [29]).

Figure 1 sets out the tasks that characterise economic evaluation according to Drummond et al.

Figure 1, distinguishing characteristics of health care evaluation[29]

Is there comparison of two or more alternatives		Are both costs (inputs) and consequences (outputs) of the alternatives examined?		
		No		**Yes**
		Examines only consequences	*Examines only costs*	
	No	**1A Partial evaluation** Outcome description	**1B** Cost description	**2 Partial evaluation** Cost-outcome description
	Yes	**3A Partial evaluation** Efficacy or effectiveness evaluation	**3B** Cost analysis	**4 Full economic evaluation** Cost-effectiveness analysis (CEA) Cost-utility analysis (CUA) Cost-benefit analysis (CBA)

A study which considers the costs and consequences of an alternative, but not at the same time, is defined as a partial economic evaluation, such as a Cost analysis (figure 1, box 3B), these evaluations do not provide answers to efficiency questions. In comparison, examples of full economic evaluation techniques, in box 4 of figure 1, include:

Cost effectiveness analysis (CEA) where costs are related to one common effect between the alternative programmes, stated either as 'cost per unit of effect' or 'effects per unit of cost' (life years gained per pound spent).

Cost utility analysis (CUA) attaches utility values to the health states produced by the alternative programmes. Utility refers to an individual or society's preferences for a set of health outcomes, allowing for health related quality adjustments to be applied to a given set

of treatment outcomes, alongside providing a generic outcome measure to allow for comparison of the treatment of costs and outcomes of different health programmes.

In the UK, the generic outcome measure which is usually used in CUA is the quality adjusted life year (QALY), which adjusts the length of time affected through the health outcome by the utility value. Utility is measured on a scale from zero, representing death, to 1, being perfect health. Where one intervention generates more QALYS and a lower cost, this is unambiguously preferred. However, where higher gains in QALYs are achieved at a higher cost, a comparison should be made in terms of the cost per QALY ratio 'ICER' (incremental cost effectiveness ratio). In the UK, NICE are reported to use a threshold value of around £20,000 to £30,000 per QALY, therefore where an intervention costs less than £20,000 per QALY it is more likely to be accepted than an intervention costing above £30,000.

Unlike CEA and CUA, *Cost benefit analysis* (CBA) should be carried out where one wants to consider a situation where it may be appropriate to increase the budget. It has its grounding in welfare theory and measures all benefits of interventions in monetary units (typically using willingness to pay, also known as contingent valuation).

As Briggs et al[30] discuss, an increased demand and use of economic evaluation for resource allocation decision making, has led to clear requirements on researchers in terms of the analytic methods they employ to carry out the evaluations. These methods need to incorporate all appropriate evidence into the analysis in order to be able to compare the new intervention with all alternative options and to reflect any uncertainty present. Consequently, economic evaluation has turned to decision analytical modeling as a framework for decision making under situations of uncertainty. Decision analytic modeling involves a set of analytic tools grounded in statistical decision making and closely associated with Bayesian statistics, which has been widely used in business analysis and engineering[30]. As Briggs et al explain "in the context of economic evaluation, a decision analytic model uses mathematical relationships to define a series of possible consequences that would flow from a set of alternative options being evaluated" p.6[30].

There is relatively little economic evidence on the effectiveness of medical devices, including hip prostheses. However, there is clear recognition of the increasing range, innovation and cost associated with medical devices[31]. This has led to a growing debate within the orthopaedic and health economic community on the need for regulation and economic evaluation of medical devices, and how to go about this process - with a focus on whether

these processes should be the same or different to those already in place for pharmaceuticals [32-34]. As Vallejo-Torres et al discuss, conducting health economic evaluations is not a core activity of most of the medical device companies[31] and thus is not integrated into their product development process. In fact, there are a number of reasons why evaluation of medical devices clearly differs to that of pharmaceuticals, including the 'learning curve' effect associated with the use of a device, where the surgeon gains experience at implanting the device over time, and their skill and experience impacts on the success of the intervention[33]. Recently, Vallejo-Torres et al have argued for an iterative economic modeling approach to be used to inform decisions regarding the cost effectiveness of a device in its early stages.

1.4.2 Industrial Organisation (IO)[3]

In assessing whether the patient is receiving the most 'cost-effective' prosthesis on the NHS, it is also necessary to consider what choices the NHS is currently making, and to examine whether it is acting as an efficient purchaser and provider of, in this case, hip replacement surgery. In order to pursue this, a useful starting point is to think in terms of the two sides of the market - of the demand and supply. On the demand side, there is the NHS which is a large organisation with considerable potential buyer power i.e. a single buyer (monopsonist) in the market. On the supply side, there are the manufacturers of hip prostheses who the NHS purchases from. As will be shown later, there is a small number of suppliers, mainly large multinational firms. In this respect, the position is similar to that of the pharmaceuticals markets. Thus the choice decisions in this context are the result of a relationship between mainly a single buyer, with potentially considerable buyer power, and a small number of oligopolist suppliers/sellers, also with potential market seller power. In order to explore the implications of this relationship, the thesis will draw on the theory of Industrial Organisation – the part of micro-economic theory which "studies the operation and performance of imperfectly competitive markets and the behaviour of the firms in these markets". p.7[35].

To help set the scene, it is useful to refer back to the basic micro-economic theory of perfect competition, monopolies and oligopolistic competition, for which I will, in the main, refer to a standard microeconomic text[36][4].

[3] Chapter 6 provides a more in-depth discussion of the theory of IO and competition policy related to this thesis.

Under perfect competition, six main characteristics exist: there are a large number of buyers and sellers; consumers and producers have perfect knowledge; the products sold are identical; firms act independently of each other with the aim of maximising their individual profits; firms are free to enter and exit the market, and finally, firms can sell as much output as they wish at the current market price.

However, in reality, most markets comprise of a only a few firms, who may be very sizeable, in order to realise economics of scale and declining average costs. In the most extreme case, costs may be minimised when there is only one firm, a natural monopoly.

In contrast to a perfectly competitive market, in a monopolistic market the supplier is a price maker (it sets the price at which it sells its output), entry of new firms into the market is blocked, and buyers are price takers. Where a monopolistic market exists there is allocative inefficiency, in that the price exceeds the marginal cost of producing the product, and the monopolist supplies less output than is optimal for society. This is the classic case against monopoly (which forms part of the justification for competition policy). It is traditionally shown by industrial economists using the classic diagram comparing perfect competition and monopoly, which shows the welfare loss from monopoly, figure 2.1, p.41-44[37].

Of course, perfect competition and monopoly are theoretical extremes, and most real world industries entail a small number of suppliers who usually account for a large proportion of the industry, this is known as an oligopoly. Where there are only two suppliers in the market, this is known as a 'duopoly'. In oligopolistic markets, further entry into the market is often difficult, and sometimes completely blocked. An example of an oligopolistic market is the food supermarkets in the UK, where the leading firms, Asda, Tesco, Sainsbury's and Morrisons have a combined market share of 65%[36].

Now, an extra dimension in the analysis concerns the *interdependence* of the firms – when one firm changes its price/output/product range etc. it has an immediate effect on the sales of the others. In this case, a key issue is how the rivals react to this interdependence. One possibility is that they engage in fierce competition, and this may result in a beneficial outcome for consumers which is not very different from that produced under perfect competition. This can happen especially if it is easy for new firms to enter the industry and undercut the otherwise high price of the existing firms, as modelled by Contestable Market

[4] The reader is referred to Morgan et al[30] for a comprehensive description of monopoly, this text is heavily referenced in the remainder of this section.

theory, p.73-75[37]. But other possibilities include much softer competition, under which firms decide to not act aggressively to each other to avoid provoking retaliation. In that case, price may be high and inefficiencies similar to those under monopoly might occur. Traditional oligopoly theory, and its modern equivalent, Game Theory in an IO context, is devoted to such issues.

One example of 'soft competition' under oligopoly is where the group of suppliers effectively join together in order to maximise profits, by forming a 'cartel'. In extreme forms of cartel, the suppliers collectively produce a level of output where the industry marginal revenue equals marginal cost – equivalent to monopoly. In order for a cartel to function successfully, it must be able to prevent its members from cheating by producing too much output and also limit or restrict the new entry of other suppliers. While cartel agreements are illegal in most countries including the UK and the EU, they still occur quite frequently, as can be seen from the cases published by Competition Authorities, and the academic literature analysing the causes and effects of cartels [38, 39]. More generally, other forms of collusion may exist, which are not illegal, because firms do not make formal agreements to collude, but nevertheless amount to an implicit agreement not to act aggressively to each other. These are generally referred to as 'tacit collusion', Motta[37], chapter 4, p.138-141, and are also more likely in markets where there are only a few firms and entry of new firms is difficult. In the context of this thesis, the fact that there is only a small number of suppliers (in fact, as shown later, just two firms have a very large share of the market) raises the *logical possibility* that competition between them may not be fierce. If so, the supply of hip prostheses may be subject to the harmful effects of monopoly like behaviour – high price, market sharing and perhaps slow innovation.

From a public policy perspective, in a situation where a monopoly, duopoly or oligopoly exists, policy should be put in place to constrain suppliers from exploiting their market power. In the UK concerns related to competition and market failure are referred to the Office of Fair Trading and/or the Competition Commission. In Europe, this is the role of DGCOMP within the European Commission.

However, the above discussion has focused only on the supply side of the market, and the implication is that, on the demand side, buyers do not have any power to counter the effects of competition between suppliers. More generally however, Industrial Organisation theory and the Competition Authorities recognise that there will be some markets in which the

buyers too have power. It is argued that such 'buyer power' is more likely to occur in markets where there is only a small number of large buyers who can use their bargaining power to enforce competition between the sellers. In this case, if the buyer is well informed about the prices of the product and about the available alternatives, it may be able to exploit its dominant position in the market to extract low prices by threatening to switch its purchasing from one seller to another – even when faced by a set of powerful suppliers. If so, even a cartel might not be able to exploit the potential for selling power. This possibility is always assessed by the authorities when they conduct investigations of potential competition problems in particular markets, p.121-123 [37].

In the context of this thesis, the NHS is a large organisation with the potential to exploit its buyer power, thus behaving as an effective monopsonist (single buyer). Thus, the question is whether the NHS, which in principle has considerable buying power, when faced with suppliers who also have considerable selling power, is able to achieve an efficient allocation of resources? Below, chapters 6 and 7 explore this question: chapter 6 examines various facets of the buyer and supplying industries, and chapter 7 explores whether the purchasing decisions of different hospitals reveal whether the NHS is a homogenous entity exploiting its buyer power to achieve efficient purchasing of joint prostheses, or whether there are systematic differences between hospitals at the disaggregated local level which might reflect segmentation of the market by the sellers, and perhaps the loss of buyer power by the NHS.

1.4.3 Public Procurement

When considering whether the NHS can/does exploit its buyer power, it is also necessary to have an understanding of public procurement in general, and specifically process of procurement by the NHS in particular. This is also explored in chapter 7, but the following provides a brief introductory overview of some relevant literature.

In the UK, public procurement refers to various areas of government activity - social security, health, education, defence and public order account for almost three quarters of total government public expenditure[6], estimated at over £150 billion per year[40]. Here I focus on some of the issues with specific reference to the NHS.

Public procurement is the process whereby public organisations such as the NHS, purchase goods and services from a third party; and within this, 'Commissioning' refers to the

decision making on which service or product the public sector service requires. Commissioning in the NHS was established as part of the introduction of the ‘internal market’ in 1991, initially with two models of purchasing (i) health authorities - centred on the health needs of the population; and (ii) fundholding - where GPs in individual practices or consortia could purchase elective care for patients[41]. More recently, Primary Care Trusts (PCTs) have been acting as the commissioners of health care for their local population, commissioning services from hospitals (the providers). Each PCT receives a budget based on a complex ‘weighted capitation’ formula designed to link budgets to local needs[41]. However, the recent 'Health and Social Care Bill' white paper of January 2011[42] (under the coalition government formed in 2010), has set out a new restructuring of commissioning, with GPs working in groups of practices called a consortia. Each of these consortia will be responsible for its own commissioning and financial decisions, although these decisions will be overseen by a national NHS commissioning board who will also commission some services directly. Thus, from 2013, GP consortiums will take over from the PCTs currently responsible for commissioning[42].

Earlier, the New Labour NHS reforms (discussed in more detail in chapter 6) also set out policy aims which had a direct impact on procurement. On the supply side of health care, hospital Foundation Trusts were first established in 2004, these are autonomous hospitals operating within the NHS who have greater operational and financial freedoms than NHS Trusts i.e. not performance managed by the Strategic Health Authorities (SHAs). There were also transaction reforms which involved a move from negotiated contracts with hospitals i.e. block contracts and cost and volume contracts, to the system of Payment by Results (PbR), where hospitals are paid on a 'per case basis' and the prices are fixed nationally. Both these major policies will continue to run under the new ‘Health and Social Care Bill of 2011'.

Most recently Sir Phillip Green's report[43] commissioned by the new coalition government, identified large scale inefficiencies at the Central government level, although this report did not specifically apply to the NHS. Even more recently, the National Audit Office (NAO) has published a pivotal report on "The procurement of consumables by NHS Acute and Foundation Trusts"[42] (orthopaedic prostheses are included within 'consumables'). This provides a rare but informative account of procurement practices within the NHS. Amongst other things, it found very limited data to be available on purchasing by individual trusts (p.4) - a finding which has been echoed by my own literature searching for this thesis. This lack of

comparable data means that trusts cannot easily identify how the prices they pay compare with those paid by other individual trusts and therefore whether better 'deals' might be available. In their own data collection, the NAO reports very wide variations in the prices paid for the same item.

Of central relevance to one of the key issues of this thesis - whether the NHS acts as a large homogenous purchaser of hip prostheses - the NAO report finds that most trusts are now outside the DoH's control (due to their Foundation trust status), and thus there is no mechanism to secure any commitment by the separate trusts to purchase in a 'collective' manner. Procurement is the responsibility of individual trusts, which the report suggests means that "significant economies of scale are being lost across the NHS" (p.7). The report also suggests that suppliers to the NHS have no doubt benefited from this lack of price transparency and weak price negotiation which arises from the disaggregated and 'fragmented' purchasing system. I return to this in later chapters, but this evidence does call into question whether the NHS can, in fact, be viewed as a single dominant buyer of hip prostheses. If not, this leaves firmly open the possibility that the main prostheses suppliers may be able to exploit their potential for market power.

1.5 Structure of the thesis

The remainder of the thesis is structured as follows:

Chapter 2 provides a more detailed discussion of THR and more specifically, the prostheses implanted, including a historical context and description of the market for prostheses to date. It also introduces the National Joint Registry for England and Wales in more detail alongside other national joint registries.

Chapters 3, 4 and 5 then address the first objective of the thesis, which examines an important element in how choices between prostheses *should* be made – the survival rates of prostheses. Chapter 3 presents an initial analysis of the data already published from the NJR on early revision (time from when the prosthesis is implanted to when it requires replacing). Chapter 4 provides a literature review of the subject area with specific focus on the economic evaluation literature of the alternative hip prostheses used in THR surgery. In summary, the main finding of chapter 4 is that there is very little empirical evidence on the long term survival of the prostheses. In view of this, Chapter 5 explores the possibility of projecting

survival rates of prostheses into the future a solution to the lack of published long term data. This chapter employs the Swedish hip registry which has been collecting data for considerably longer than any other joint registry. It revisits a previous study by Briggs[18] et al to test the robustness of their extrapolations. The findings of this chapter suggest a lack of robustness of extrapolation methods to date in this context.

This cautionary finding leads to the change in focus at this point in the thesis. Given that chapters 3,4 and 5 find that clear recommendations on prosthesis choice employing long term survival cannot yet be made, the second part of the thesis switches to the ***actual*** decisions which are made by surgeons regarding which prosthesis is implanted (thereby addressing the second objective). Here, the motivation is to explore how the potential buyer power of the NHS, on the one hand, and the potential selling power of the manufacturers, on the other hand, interact in terms of the mix of prostheses used across different hospitals in the NHS.

Chapter 6 first discusses some of the relevant policy and theoretical issues including a discussion of patient choice and principal-agent theory. The chapter then moves on to the issue of public procurement in the NHS and the role of the NHS as a potential monopsonist with significant buyer power, drawing on the theoretical background of Industrial Organisation. The chapter then turns to the supply side industry and its potential seller power, concluding with a brief synopsis of the implications for market power.

Chapter 7 presents the main empirical work in this part of the thesis. Using a set of econometric techniques to test hypotheses, it first establishes whether choice of prostheses is determined mainly by patient characteristics or by the characteristics of different hospitals, and finds very large differences between hospitals. It then investigates whether variables such as hospital size, location and status play important roles, and whether there is any evidence that the manufacturers have segmented the market.

The final chapter brings together the conclusions of all the chapters providing a summary of the main findings, acknowledging the limitations of the thesis and setting out an agenda for future work.

1.6 Summary

The ultimate aim of policy in this area should be to ensure that the patient receives the best quality hip prosthesis available, within the constraints of the NHS budget. This thesis explores two different aspects of choice within the NHS which are relevant to this aim. The first is how to inform decision-making with methods which can identify which prosthesis should be implanted, bearing in mind that little is known about their actual long-term survival rates. The second is to identify the actual choices made by different hospitals in order to attempt to answer whether the NHS is an efficient purchaser of hip prostheses, behaving like a single entity, or whether there are differences between hospitals which may suggest that the dominant prostheses suppliers might have a strong influence on different decisions taken at the local level.

Chapter 2, Contexts and background: Hip prostheses, National Joint Registry and early revision

This chapter provides the background perspectives for the rest of the thesis. It is organized into three parts. Section 2.1 provides a brief history and definition of THR, describes the regulatory setting in the UK, and introduces the key data sources used in this thesis - the National Joint Registry for England and Wales and the Swedish Joint Registry. Section 2.2 draws on the published versions of the NJR to map out a statistical introduction to the current state of THR in England and Wales, in terms of types of prostheses, the patient-mix and hospitals.

2.1 Background

2.1.1 Definition of hip prostheses

Hip prostheses have two components - a stem made of stainless steel or a chrome cobalt molybdenum alloy and a cup made of high density polyethylene[4] . In England and Wales, NICE recognises three broad categories of prosthesis: **cemented**, **cementless** and **hybrid**[20] determined by their method of fixation. Traditionally, components were fixed to the bone with an acrylic cement, known as 'cemented' and first designed by Sir John Charnley in the 1960s. More recently there has been an increase in prostheses which are not fixed with cement, relying on the bone growing into irregularities on the surface of the component[4], these are known as 'cementless'. The third type of prosthesis is a 'hybrid', where one or other of the components is fixed by cemented, and the other is cementless. An alternative to THR surgery is hip resurfacing, this was introduced in the mid-nineties, and involves conserving the femoral bone and only replacing the surface of the joint[8]. The reader is referred to Vale et al[44] for an in-depth review of hip resurfacing.

2.1.2 Brief history of THR

Total hip replacement surgery has revolutionised the treatment options for patients with end stage joint disease of the hip[45] over the past 50 years. Table 2.1 provides an overview of the historical developments in THR, and some of the main manufacturers over the past two centuries (this draws extensively on Anderson et al[46]). It clearly shows the rapid technological change which has taken place, particularly from the 1950's onwards, and that the origins of THR are firmly based in Britain and the United States.

One of the early pioneers of this type of surgery was George Kenneth KcKee who worked in Norwich, England as an orthopaedic surgeon. He was the first surgeon to explore what we now term a 'THR' i.e. using both an artificial ball and socket joint in the 1950s. In the United States, Austin Moore replaced the head of the femur as early as the 1940s. However, it was the development of the low friction arthroplasty produced by Sir John Charnley in Oxford in the 1960s which really led to the modern day prostheses used in THR. Since then, many surgeons and manufacturers of medical devices have developed newer prostheses using different materials and methods of fixation from the original 'Charnley' prosthesis. However, the Charnley and its competitors from the same era, such as the Exeter and the Stanmore, are still widely in use today, albeit in 'updated' forms.

The Table also reveals the evolution of the prosthesis manufacturing industry, in terms of the birth of firms, and consolidations by mergers and acquisitions. Most of the early innovation and development in the hip prosthesis industry started with a single surgeon working with an engineer or small chemist. For example, one of the more famous partnerships was that of Sir John Charnley and Chas F Thackray ltd. Chas F Thackray Ltd was first established in the 1800s as a pharmacy in Leeds, in 1918 they turned their focus to surgical equipment and by the 1940s they had developed a partnership with Charnley to develop hip prostheses. This partnership continued up until the 1990s when Thackray Ltd was acquired by Boehringer Mannehaim, a large multi-national company who had also acquired Depuy in 1974. Smith and Nephew also started as a pharmacy back in 1856, but over time turning their focus to orthopaedic equipment, and today is one of the major multinational suppliers of orthopaedic prostheses.

In the 1970s, a flurry of acquisitions and developments followed, with Bristol Myers Squibb acquiring Zimmer USA and Pfizer acquiring Howmedica (a British company). This period also saw the rise of some of the current major players in the industry, such as Biomet, Joint

Replacement Instrumentation (JRI) and Osteonics (acquired a year later by Stryker). By the 1990s more of the pharmaceutical companies had entered the market.

Table 2.1

Time Period	Action	Category
1800-1900	Chas F Thackray Ltd established, as a pharmacy purchased in Leeds, UK[38]	Est
	Smith & Nephew established as a chemists - Hull, UK (1856), they later enlarge and specialize in elastoplasts and plaster of Paris	Est
	Down Bros (formerly Millikin and Down) established (1881) to supply splints and medical devices to Guy's hospital, London - UK	Est
	Depuy established (1895) - USA	Est
1918	Chas F Thackray focus on surgical equipment[38]	Tech
1920s	First insertion of artificial joint - French and US surgeons	Tech
	Zimmer established (1927) - USA	Est
1930s	Don Richards establishes Richards, Tennessee, USA (1934)	Est
	First metal THR (femoral head and cup replaced) (1938) - London UK	Tech
1940s	*Austin Moore* & *Harold Bohlma*, replacement of cancerous femoral head (1943) - USA	Tech
	Orthopaedic Equipment Company (OEC) established (1943) - USA	Est
	Mushroom shaped acrylic prosthesis to replace head of femur (1946), *Judet* brothers - Paris France	Tech
	Chas F Thackray and *Sir John Charnley* (surgeon) together work on prostheses	Tech
	Introduction of antibiotics and anesthesia	Tech
	Establishment of the NHS in Britain	
1950s	***Wright manufacturing*** established (1950) - Memphis, USA	Est
	Trend emerging in the US for replacement of femoral head	
	Chrome cobalt alloy femoral stem (1950), Frederick *Thompson* - USA	Tech
	George Kenneth McKee (1951) championing THR (artificial ball and socket joint) using metal	Tech
	D Howse established to distribute: Howse-Arden; McKee-Farrar; Monk prostheses	Est
	Introduction of *Stanmore* prosthesis (Royal National Orthopaedic hospital, London)	Tech
	Self-locking femoral replacement with holes to encourage bone growth - (1952) Austin Moore& Frederick Thompson (USA)	Tech
1960s	McKee adds acrylic to the cement for fixation due to loosening of metal components	Tech
	Establishment of Zimmer UK (1964) resulting from a disagreement between Justin.O.Zimmer (Zimmer USA) and one of his salesmen, who went on to Europe and established the Zimmer name in many other countries.	Est
	McKee Farrar adapt the Thompson femoral component (1965), leads to a increase in the success rate	Tech
	Depuy secure rights for the *Muller* hip (1968)	Dev
	Uptake of the *Charnley low friction arthroplasty* (Oxford UK)	Tech
	Introduction of the *Ring* prosthesis (Royal College of Surgeons, UK)	
	Joint Replacement Instrumentation (JRI) established to import *Muller* prosthesis	Est
	Ring change to using polyethylene	Tech
	Research starts on the *Ling-Lee* prosthesis (1967) (later becomes the *Exeter*)	Tech
	Establishment of the Hip society (1968)	
	Howmedica established 1969 (previously Howesound & Howemet)	Est
1970s	Down Bros supply the *Redhill hip* under license to Howmedica	Dev
	General acceptance of the surgical success of THR	
	McKee stops inserting metal on metal hips due to early complications (1972)	Tech
	Howmedica secure rights for the *Harris hip* - 2 component replacement (1972)	Dev

	Bristol Myers Squibb (1972) (pharmaceutical company) acquire Zimmer USA	A
	Pfizer acquire Howmedica (1972)	A
	Deloro Stellite (Swindon, Wilts) establishes ***Deloro Surgical (UK) Ltd*** (1972) manufacturing hip prostheses including *Stanmore*	Est
	Boehringer Manneheim acquire Depuy (1974)	A
	JRI established, Sheffield, UK (1977)	Est
	Biomet established (1978) - Warsaw, Indiana, USA	Est
	Osteonics established (1978) (by engineers from Howmedica)	Est
	Osteonics acquired by ***Stryker*** (1979)	A
	Ring use the polyethylene cup with an offset peg	Tech
1980s	Switch in trend from metal on metal prostheses to metal on plastic	Tech
	Johnson & Johnson open division called Cintor – marketing knee replacements and the Charnley hip in the U.S (1981) J & J acquire D Howse & Co, thus entering the UK market (1982)	Est
	Zimmer UK acquires Deloro Surgical (1980)	A
	Corin (UK) established(1985) supplying the *Freeman Modular hip* and *Cormet* hip	Est
	Orthopaedic Equipment Company (OEC) acquire Zimmer UK	A
	Biomet acquire OEC (1984)	A
	Introduction of the Wrightington THR (Howmedica)	Tech
	C-stem prosthesis (by Wroblewski) (1982)	Tech
	Smith & Nephew acquire ***Richards*** (1986)	A
	Journal of Arthroplasty established (1986) (focus on joint replacement)	
	Aesculap (German) acquire Down Bros (1988)	A
	Exeter modular hip introduced (1988) (copied by others including *C-stem)*	Tech
1990s	***Boehringer Manneheim*** (pharmaceutical company, also own Depuy) acquire Chas F Thackerey	A
	B Braun **Medical ltd** acquire Aesculap	A
	Roche acquire Boehringer Manneheim (Depuy) (1990)	A
	Biomet enter into a manufacturing agreement with ***Merck*** (1997)	M
	Johnson & Johnson acquire Depuy from Roche (1998)	A
	Smith and Nephew acquire ***Midland Medical Technologies*** (1999)	A
	Exeter establishes dominance in the British market(1999)	Tech
2000s	Publication of NICE guidance to Hip prostheses, England and Wales (2000)	
	Zimmer acquire ***Sulzer (known as Centerpulse***) (2003) after bidding war with Smith & Nephew	A
	The National Joint Registry for England and Wales established (2003) - England and Wales	
	Smith & Nephew acquires ***Plus Orthopedics Holding AG*** **("Plus")** (2007)	A
	LVB Acquisition Merge Sub, Inc merges with *Biomet*	A

Est = manufacturer established; Tech = technological development; Dev=manufacturer development; A=acquisition; M=merger

Table 2.1 Timeline of developments in the hip prostheses industry[46]

2.1.3 Regulation

Hip prostheses are categorised as a medical device[47][5] and as such, in the UK, they are monitored by the Medicines and Healthcare Products Regulatory Agency (MHRA), an executive agency for the DoH. This was established in 2003 as a consequence of a merger between the Medicines Control Agency and the Medical Devices Agency, and is responsible for ensuring that medicines and medical devices work and are acceptably safe.

MHRA explain that the main difference between how medicines and devices are regulated relates to how a product gets onto the market; medical devices are approved by the private sector organisations which are called 'notification bodies' and their approval is needed before a CE mark[48][6] is awarded to a device. These 'notified bodies' are private organisations who carry out 'compliance assessments' before certain medical devices can go on the market. They are designated and audited by MHRA[49], a full list of them can be found on the MHRA webpage[50] .

Aside from MHRA, NICE provides guidance to patients and the NHS on best-practice procedures in healthcare. Within this remit, they carry out appraisals of new and existing pharmaceutical and medical technologies, in order to demonstrate the value of the product to the NHS based on proven clinical and cost-effectiveness[51]. In 2000, NICE published guidelines[52] on: *'The selection of prostheses for primary total hip replacement'*, which have remained the primary guidance in the UK on total hip replacement and hip prostheses[7]. The main points in their guidelines are that prostheses should demonstrate a revision rate of 10% or less at 10 years, which should be regarded as the current benchmark. NICE also considers it reasonable to recommend consideration of a prosthesis with a minimum of 3 years revision rate experience if its performance is consistent with the benchmark of a 10% revision rate at 10 years[52].

Following the NICE guidance, the Orthopaedic Data Evaluation Panel (ODEP) was established to 'ensure consistency and enable easy data presentation and comparison'[2] of hip prostheses, as part of the NHS Supply Chain. It provides a rating for prostheses based on

[5] According to the European Union (EU) directive 2007/47/EC. a medical device is defined as: "any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, including the software intended by its manufacturer to be used specifically for diagnostic and/or therapeutic purposes and necessary for its proper application, intended by the manufacturer to be used for human beings for the purpose of diagnosis, prevention, monitoring, treatment or alleviation of disease"[41].

[6] CE mark is a declaration from a manufacturer that a product meets all appropriate provisions of the relevant legalization including those relating to safely and where required has been assessed in accordance with these[42].

[7] NICE states that they will consult the review plans on this guidance in May 2011

data submitted by the manufacturers, Figure 2.1 reports these ODEP ratings. For example, the Charnley cemented cup and stem both have a rating of 10A, indicating strong clinical evidence of prosthesis survival at 10 years.

Manufacturers are required to inform ODEP of all their commerically available prostheses which have been involved in postmarket clinical follow-up studes using seperate proformas for cup and stem. They request data on product details including: prosthesis history; whether it has met the NICE benchmark; information on publications such as conference and peer-reviewed papers; Kaplan-Meier survival curves and revision rates. However, manufactueres are not required to provide information on those products still in development. ODEP provides no explanation about whether or how the manufacturers proforma information is checked or verfied and whether the manufacturer information is publicly available[53]

Figure 2.1, ODEP classification[2]

ODEP Classification:

Pre-entry - Manufacturers are requested to keep ODEP informed of all commercially available prostheses that are involved in post market clinical follow-up studies.

Unclassified

3A - 3 year data, Acceptable evidence. Failure rate of 3% or less.

3B - 3 year data, Weak evidence. Acceptable failure rate.

5A - 5 year data, Acceptable evidence. Failure rate of 5% or less.

5B - 5 year data, Weak evidence. Acceptable failure rate

7A - 7 year data, Acceptable evidence. Failure rate of 7% or less.

7B - 7 year data, Weak evidence. Acceptable failure rate.

10A - 10 year data. Strong evidence. Failure rate of 10% or less.

10B - 10 year data. Reasonable evidence. Failure rate of 10% or less.[2]

10C - 10 year data. Weak evidence. Failure rate of 10% or less. Products given 2 years to improve data or they are deemed unacceptable.

2.1.4 Joint Registries

It is well documented that historically there is very little high quality evidence on the performance of hip prostheses, specifically in terms of their key **survival rates**, i.e. the time until they need to be replaced)[54, 55]. In order to address this gap, many countries have

established their own national joint registries reporting the performance of the prostheses by type and brand name.

The first joint registry was established in Sweden in 1975 (the Swedish Hip Registry) and this is usually viewed as the pioneering joint registry. Its primary aim was to collect data on nationally used prosthesis survival rates. Other subsequent joint registries have since aimed to follow the Swedish Registry's approach, but they have often come under political and practical challenges[56] for example, the German Arthroplasty Register was unsuccessful in sustaining its existence because of "the low rate of participation and the resulting financial problems", p.1567,[56].

Country/Registry name	Year established
Sweden Knee	1975
Sweden Hip	1979
Finland	1980
Norway	1987
Denmark	1995
Germany Knee	1997
Denmark Hip	1997
Australia	1999
New Zealand	1999
Sweden (shoulder & elbow)	1999
Canada	2001
Romania	2001
England and Wales	2003
Slovakia	2003
Switzerland	2004

Source: [56]

Table 2.2 Countries with National joint registries

The NJR and Finnish Registries were set up by government institutions, with the NJR funded by levies placed on the prostheses sold[9]. Most other registries are maintained by national orthopaedic associations. Participation in most registries is voluntary, other than Denmark and Slovakia, where participation is compulsory by government decree[56].

Five registries report that they collect data on clinical scores. The NJR introduced national collection of PROMS in 2009 in the form of a self-reported patient questionnaire, which asks patients about their health from their point of view[26]. The New Zealand registry also collects

data using the Oxford hip, knee and shoulder questionnaire and the Swedish Registry uses self-reported questionnaires in the form of the Western Ontario Osteoarthritis of the shoulder index and the EuroQol (EQ-5D). The Romanian and Swiss registries collect data on radiological findings as part of the post-operative follow-up[56].

The main goal of every register is to measure the outcome of joint replacement. In order to do so, the revision of an implant is set as the endpoint of failure, and the survival rate is calculated according to Kaplan-Meier survival analysis[56]. Several joint registries have demonstrated success in identifying poorly performing prostheses i.e. the Scandinavian registries and the Christiansen Hip [19].

The Swedish Hip Arthroplasty Register (SHAR)[57]

As the Swedish registry is used in chapter 5 it merits a brief separate discussion here. It was established over 30 years ago. Since then it has been routinely collecting data on the alternative types of prostheses implanted and the surgical techniques used, revisions and reoperation rates and demographic data. All 79 hospitals (public and independent sector) are included in the register, mostly reporting their data via a web application.

The register reports prosthesis survival at 3, 5 and 10-year time points, with 10 year survival now reported as being over 95% on average. Recent developments in SHAR include routine collection of patient reported outcome measures (PROMs) since 1992 (routinely collected on the NJR since 2009), and a joint database with Denmark, Norway and Sweden from 1995 onwards. SHAR also reported cost and cost-effectiveness analysis at the hospital level in their 2007 report, however this was not continued in the 2008 annual report due to some apparent difficulties with cost calculations

The National Joint Registry

The NJR for England and Wales is the main data source used in this thesis. It was established in 2003, and collects data on hip, knee and most recently, ankle replacements carried out in the NHS and independent healthcare sectors[58]. It has been managed by the Health Quality Improvement Partnership (HQIP) since 2008 (previously managed by the DoH) and is funded through a levy raised on the sale of hip and knee replacement prostheses. It is the largest international registry recording prosthesis performance and is therefore able to provide data on a scale not previously available. To date, the NJR contains data on up to 7 years survival rates of prostheses. The specific aims of the registry are discussed on the NJR

website[59] and in the 4th annual report[60], however, they can be summarised as the following:

"*to highlight in real time any brand of prosthesis showing high failure rates, and allow prompt removal from the market, if necessary" and to "improve evidence-based purchasing of joint replacement implants for orthopaedic units/hospitals*" p.1[61].

The NJR is linkable to HES and more recently, to the PROMs database, all of which are managed on a day to day level by Northgate Information Solutions. The option to link these three data-sets provides a vast and invaluable source of linkable individual patient level data previously unavailable for England and Wales.

2.2 The current position in England and Wales: prostheses, patients and hospitals

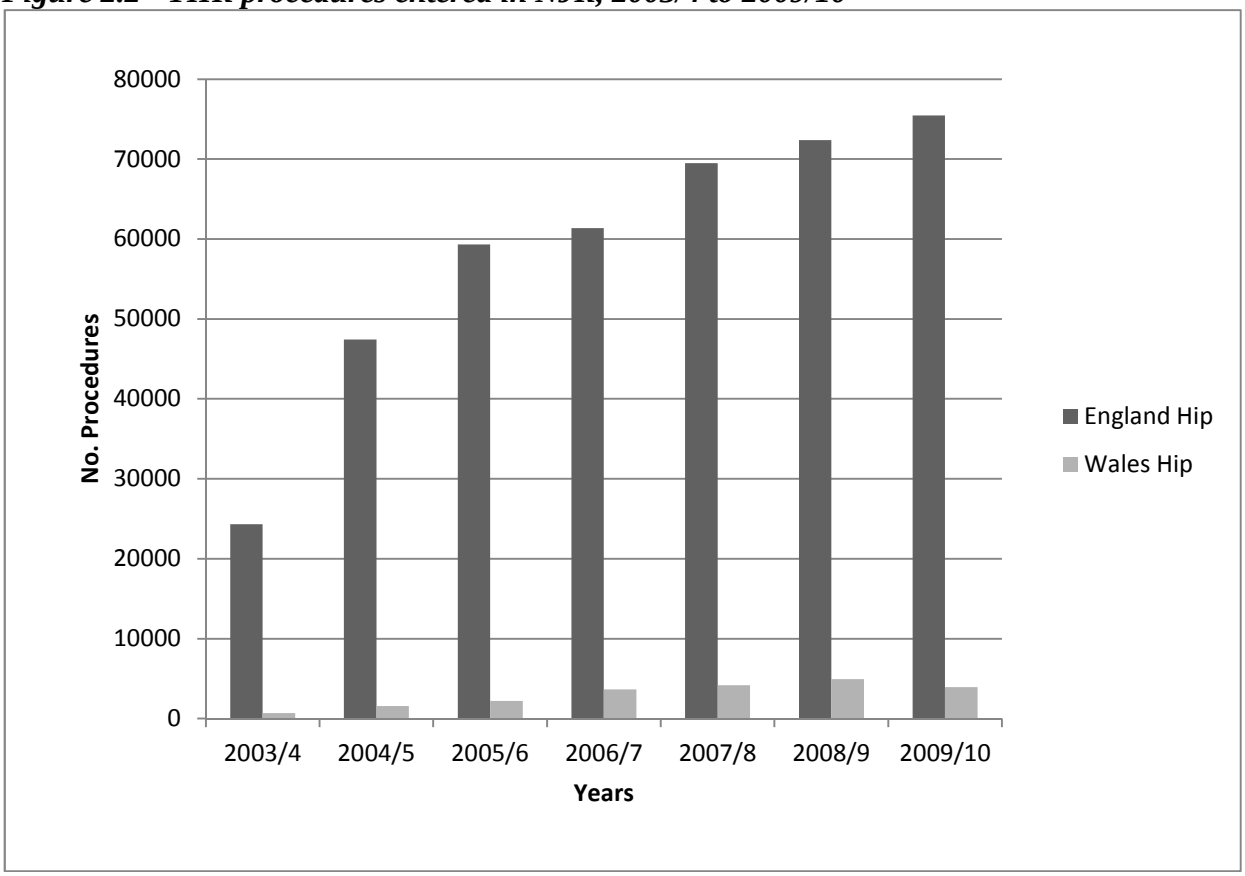
The analysis in the following chapters use, as their source, the raw data on individual patients in the NJR and Swedish joint registries. However, the summary statistics already published in the annual reports of the NJR provide a useful starting point. This section draws selectively from the annual reports to establish an opening picture.

2.2.1 Recent growth in THR

As a consequence of an ageing population[8] and the rising incidence of diseases of old age such as Osteoarthritis, rates of elective total hip replacement have risen year on year for the past two decades[9, 60, 62, 63]. Figure 2.2 shows the growth in THR procedures in the NJR, as reported from 2003/4 to 2009/10. It clearly illustrates the rise in procedures over this 7 year period[9, 60]. It should be noted that some proportion of the increases in reported rates will be due to increases in compliance and consent rates as the NJR became more established - see annual reports for reported compliance and consent rates. In particular, reported compliance was much lower in 2003, 84.5% of NHS Trusts; 91.6% of Independent Sector hospitals and 75% of Treatment Centres, by 2009 this overall figure had risen to 96% participation[9]. As will be seen in later chapters, this sometimes leads to anomalous findings for 2003 relative to the years which follow.

[8] Also as a consequence of surgeons operating on younger patients with new technologies such as hip resurfacing (also included in NJR data)

Figure 2.2 - THR procedures entered in NJR, 2003/4 to 2009/10



Source: NJR annual reports:[9, 58, 60, 64-67]

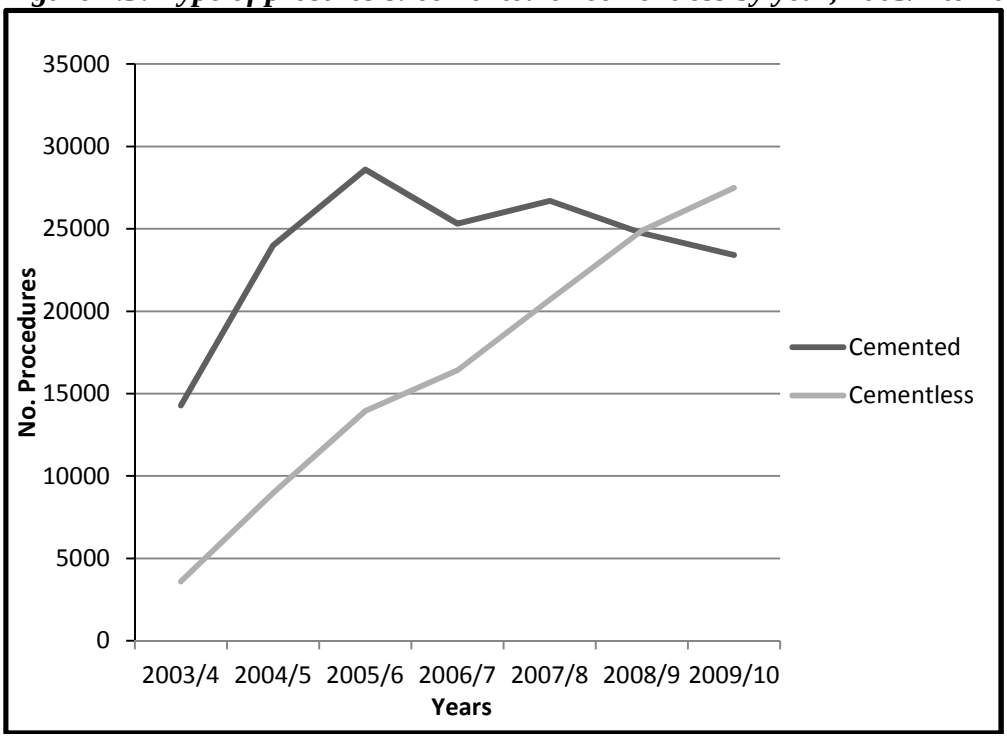
For the years before the introduction of the NJR, data on rates of elective THR were obtainable from HES; Dixon et al's analysis of data from that source show a steadily increasing trend from 1991 to 2000, but with a slight dip in 1996/1997[63].

2.2.2 Broad Prosthesis types

A key feature of competition in many markets is product differentiation: often, this is the main method firms use to compete (e.g. pharmaceuticals, breakfast cereals and cars), this can affect competition both positively and negatively. On the one hand, introducing new innovative brands can be the method by which new firms enter, or existing firms compete with their rivals. Consumers may benefit from more choice and improved product quality. On the other hand, strong differentiation can make the entry of new firms harder because they find it difficult to persuade consumers to switch from their existing brands. It may be the case here, where hospitals and surgeons have strong brand preferences, based largely on previous practice and training.

In this particular market, there is obviously the key distinction between cemented and cementless prostheses. From a competition policy perspective, the European Commission[68, 69] view cemented and cementless prostheses as substitutes and therefore within the same market, but this does not necessarily mean they are very close substitutes: surgeons may have a strong preference for cementless over cemented prostheses for certain types of patients, but the opposite for other patients. Therefore, it is important to distinguish at the outset these two broad types. Figure 2.3 and Table 2.3 show the numbers of cemented and cementless implants recorded on the NJR from 2003/4 to 2008/9. Clearly, the proportion of cementless prostheses increased rapidly over this period: while cemented still accounted for three times as many implants as cementless in 2003, by 2008 cementless have overtaken cemented as the preferred prosthesis type.

Figure 2.3: Type of prosthesis: cemented or cementless by year, 2003/4 to 2009/10



Source: NJR annual reports:[9, 58, 60, 64-67]

Year	Cemented	Cementless	Total
2003	14280	3590	17870
2004	23992	8957	80856
2005	28602	13955	90548
2006	25313	16416	90513
2007	26685	20690	100513
2008	24730	24892	100488
2009	23414	27492	50906
Total	311401	194752	506153

Source: NJR annual reports:[9, 58, 60, 64-67]

Table 2.3: Type of prosthesis by year (cemented and cementless)

2.2.3 Brands

Product differentiation is just not between the two broad categories of cemented and cementless; as can be seen from Table 2.4, in any given year there are more than 200 different brands implanted. While it should be remembered that each patient will receive 2 brands, 1 cup and 1 stem (counted here as different brands), Appendix 1 shows that the most commonly used combinations of cups and stems involve both cup and stem from the same manufacturer. As will become clear in chapter 6, most manufacturers offer a range of brands within each of the four types (cementless/cemented and cup/stem.)

	Number of Brands				
Year	**Cemented cup**	**Cemented stem**	**Cementless cup**	**Cementless stem**	**Total**
2003	48	57	42	41	188
2004	51	71	49	50	221
2005	50	66	51	53	220
2006	48	68	47	49	212
2007	49	66	54	56	225
2008	50	68	54	52	224

Source: Authors' calculations: NJR data-set (see Appendix 2 for a discussion of this data-set)[9]

Table 2.4 Number of prosthesis brands in the market

Of the top 16 most widely used prosthesis brands, 2003-2008, 4 are cemented stems, 4 cementless cups, 6 cemented cups and only 2 cementless stems. The 'best-selling' brand is

[9] Unlike all other tables in this chapter which I derived from the tables published in the NJR annual reports, this table is constructed from the raw data on individual patients provided to me by the NJR. However, this raw data was only available up to the end of 2008. The data-set is described in more detail in Appendix 2

the Exeter V40, which is a cemented stem and accounts for approximately 18% of all prostheses fitted (cemented, cementless, cup and stem).

The main manufacturers of these brands are discussed, along with the brands, in more detail in chapter 6. However, in brief, there are 25 manufacturers recorded on the NJR, but only 5 have a market share of consistently over 5 %. Stryker Howmedica Osteonics and Depuy both account for a third each of the market with the remaining manufacturers: Zimmer, Joint Replacement Instrumentation and Biomet 6% accounting for approximately 7%, 8% and 6% respectively.

2.2.4 Patient Mix

The NJR records data on various characteristics of the patients undergoing THR surgery, these are summarised in table 2.5. This confirms that THR is most frequently performed on older people[10]. Clearly, more females undergo THR surgery than males, around 57% . Only one indication for surgery is included in the table, Osteoarthritis (OA), this is because it is overwhelmingly the primary indication, accounting for approximately 94% of all cases. Other reasons include: avascular necrosis; fractured neck of femur; congenital dislocation and infection[9]. The final patient characteristic included in the table is side of surgery – the hip which is undergoing the replacement surgery: left, right, or bilateral (meaning both sides at the same time). Most patients have a right sided THR (55 % approx), very few patients have both hips replaced at the same time (0-1%). Bilateral surgery is both more lengthy in terms of operation time, and takes longer for the patient to recover due to the increased immobility in the rehabilitation phase post surgery.

[10] Although there has been a small decrease in the average age of patients undergoing THR since 2003/4 most of this change occurred between 2003/4 and 2004/5, and this may simply be a consequence of the reduced coverage of the NJR its first year, as reported above.

	2003/4	2004/5	2005/6	2006/7	2007/8	2008/9	2009/10
Average Age	70	68	68	68	67	67	67
Gender (%)							
Female	56	59	60	60	56	60	56
Male	44	41	40	40	44	40	44
Indication for surgery (%)							
OA	96	94	94	94	93	93	93
Side (%)							
Bilateral	1	0	0	0	<1	-	<1
Left	47	45	45	45	45	-	45
Right	52	55	54	54	55	-	55

Source: NJR annual reports:[9, 58, 60, 64-67], *2008/9 annual report did not report patient mix for side of surgery

Table 2.5 Characteristics of patients undergoing THR surgery, 2003/4 to 20098/10

Further information is provided in table 2.6 on the same patient characteristics, now disaggregated between cemented and cementless. It has already been shown above that increasing numbers of patients are receiving a cementless prosthesis. This table now reveals that it is younger patients who are more likely to receive a cementless prosthesis (approximate average age of 65 years for cementless prosthesis compared to 72 for cemented). This provides part of the explanation for the trend, noted above, for the decline in the overall age of patients undergoing THR surgery. It also shows that male patients are more likely to receive a cementless prosthesis than a cemented, while the reverse is true for females. There is little difference in cemented and cementless for the indication for surgery or the side of surgery.

	2003/4		2004/5		2005/6		2006/7		2007/8		2008/9		2009/10	
	Cem	**C'less**	**Cem**	**C'less**	**Cem**	**C'less**	**Cem**	**C'less**	**Cem**	**C'less**	**Cem**	**C'less**	**Cem**	**C'less**
Age	71	70	72	65	72	65	73	65	73	65	72	66	73	66
Gender (%)														
Female	57	56	65	57	65	57	66	57	66	57	66	57	66	57
Male	43	44	35	43	35	43	34	45	43	34	34	45	34	45
Indication for surgery (%)														
OA	96	96	94	94	95	93	94	93	95	93	93	93	94	93
Side (%)														
Bilateral	1	1	0	0	0	1	0	1	0	1	-	-	<1	1
Left	47	49	44	46	45	45	45	46	44	45	-	-	45	45
Right	53	50	55	54	55	54	55	53	55	54	-	-	55	54

*2008/9 annual report did not report patient mix for side of surgery
Source: NJR annual reports:[9, 58, 60, 64-67]

Table 2.6, Characteristics of patients undergoing THR surgery according to prosthesis type, 2003/4 to 2009/10

2.2.5 Numbers of hospitals[11]

The dominant provider of health care in England and Wales is the NHS, which is free at the point of use and funded by government taxation. However, there is a small private sector (which will be referred to as the independent sector for the remainder of this thesis) provider of health care, which is often funded by employers' medical insurance. There is also an increasing number of procedures which the independent sector is sub-contracted to carry out on behalf of the NHS. More recently, the NHS and independent sector have both established treatment centres (NHS treatment centres (NHS TC) and Independent sector treatment centres (ISTC) respectively), which are small units which carry out day surgery or specific surgical procedures often including hip and knee replacement surgery.

Table 2.7 shows the number of hospitals reporting data for the NJR according to these hospital types. It reveals a steady number of NHS hospitals, at just over 200 hospitals (apart from 2005/6 in which there were slightly more at 220). The number of Independent sector (IS) hospitals reporting their data has declined marginally from 166 in 2003/4 and 2004/5 to 160 in 2009/10. In contrast, there has been an increase in the number of treatment centres reporting data, both IS and NHS, which would be in keeping with their recent and increasing establishment under the Labour government (until 2010).

[11] The term 'Hospital' is used throughout the thesis and refers to the surgical unit of treatment,

	2003/4	2004/5	2005/6	2006/7	2007/8	2008/9	2009/10
NHS hospitals	168	205	220	208	201	201	208
England	156	191	204	191	184	184	191
Wales	12	14	16	17	17	17	17
Independent Sector	166	166	167	165	162	163	160
England	160	161	161	159	157	158	155
Wales	6	5	6	6	5	5	5
ISTC	12*	7	9	10	10	11	13
England	0	7	9	10	10	11	13
Wales	0	0	0	0	0	0	0
NHS TCs	0	3	7	10	10	11	12
England	0	3	7	10	10	11	12
Wales	0	0	0	0	0	0	0

* In 2003/4, No. Of participating ISTC and NHS TCs was given together as a figure for TCs
Sourced from various authors of the following NJR annual reports:[9, 58, 60, 64-67]

Table 2.7, Number of participating hospitals entering data on the NJR, 2003/4 to 2009/10

2.3 Conclusions

This chapter has provided a factual background to the thesis, providing some key empirical findings which can be summarised as follows: Rates of THR have risen year on year in England and Wales for the past two decades. The procedures are carried out within four main providers: NHS hospitals, NHS Treatment Centre's, Independent Sector hospitals and Independent Sector Treatment Centre's, although the NHS remains overwhelmingly, the main provider of THR surgery in England and Wales.

There has also been a rapid increase in the number cementless prostheses implanted (from 2003 to 2008), to the extent that they have now overtaken cemented as the most commonly implanted type. In terms of patient mix, THR is more frequently performed on older female patients with a primary diagnosis of OA. Younger, male patients tend to be more likely to receive a cementless prosthesis.

Beyond broad types of prostheses, product differentiation can be disaggregated at the brand level for cup and stem combinations, where in any given year more than 200 different brands of prostheses are implanted. The 'best-selling' brand of either cup or stem, cemented or cementless, is the Exeter V40 cemented stem, accounting for 18% of all prostheses fitted in England and

Wales (from 2003 to 2008). There are 25 manufacturers of hip prostheses recorded on the NJR, although only 5 of these have a market share of consistently over 5%, with Stryker Howmedica Osteonics and Depuy accounting for a third each of the market.

Chapter 3, Revision in the early years: evidence from the NJR Annual Reports

3.1 Purpose

This is the first of the three chapters concerned with revision rates of different prostheses. The purpose here is to present an opening analysis of the NJR data on revision already available in the public domain, in that it uses only the aggregate summary tabulations available in the published Annual Reviews of the NJR. (In later chapters, I conduct more detailed analysis of the primary unpublished disaggregated data on the individual patients which have been made available to me.) Although the NJR has only been in existence for seven years, and therefore these data can only tell us about revision in the early years after primary surgery, it is already becoming a potentially rich source of information which, to date, does not appear to have been subjected to much detailed academic research. Thus, this chapter helps to fill that gap.

The chapter begins with a brief summary of the only other published (to date) paper[54] using the NJR data. This was based on the first three years data and was largely confined to broad comparisons between cemented and cementless prostheses. My first objective is to update their analysis to 2009 using information now available from the more recent NJR annual reports. The second part of the chapter takes the analysis further by now exploring differences between different brands within each of these broad categories. It establishes a series of stylized facts concerning differences in revision rates between different brands of prostheses, how they change over the early years, and how they relate to reported ODEP ratings.

In terms of the thesis as a whole, the chapter provides a starting point for chapters 4 and 5. In particular, it provides a useful introduction to chapter 5, in which I investigate the reliability of extrapolating long term prosthesis survival rates, when only short-term data are available.

A key, and widely accepted outcome measure, when assessing the success of primary THR, is the time from primary surgery until the prosthesis needs replacement, known as revision surgery. Revision surgery can be necessary sometimes almost immediately post-surgery, but hopefully not until, say, 15 to 20 years, post primary surgery – if at all. The end point for prosthesis failure is usually defined as revision: "*exchange or extraction of at least one part of the prosthesis*"[70] However, it is recognised that this is not the same as the point in time when the prosthesis first

fails. Prosthesis failure is much harder to define for example, whether prosthesis failure should be recorded as the point when the patient first recognsies some discomfort at the hip joint, when they first re-visit their GP or surgeon to report the problem or when they are referred for surgery. The implication of not taking into account the time point when failure first occurs is that a time lag exists from when the prosthesis fails until when the patient undergoes revision surgery. If the lag was the same across all patients then the implication for the analysis on revision surgery would not matter. However, it is unlikely that this is the case, in which case factors such as the (i) propensity of the patient to seek medical advice; (ii) time until GP or surgeon is consulted, (iii) propensity of the GP to refer to the specialist (iv) waiting time until specialist is seen and (v) waiting time for surgery, may vary across regions associated with the alternative prostheses implanted. especially given potential waiting time variation from referral for revision surgery until the date of surgery. Ideally, data would be available on issues by region: number of GPs per head; waiting list until specialist seen and waiting list for revision surgery would be included. However, chapter 3 uses data on revision surgery taken from the NJR annual report which does not report data GPs per head or waiting times. Moreover, the data used in chapter 5 is taken from an extract provided by SHAR which also did not include information on waiting list times or GPs per head, neither is it available in SHAR annual report. Clearly the issue of defining and measure prosthesis failure is an issue for the analysis in this chapter and chapter 5 and warrants further investigation for inclusion in any cost-effectiveness analysis of alternative hip prostheses. As mentioned earlier, NICE guidelines specify that revision should not be necessary in more than 10% of cases, ten years after surgery. When revision is necessary within the first year it is often referred to as an 'early revision or failure' and can be caused by indications such as: infection, dislocation, pain or fracture[9]. Where revision is not required until after the early years, it is usually referred to as a 'late revision or failure', with the primary cause being aseptic loosening principally due to wear of the artificial joint.

Therefore, any meaningful comparison between different prostheses ideally requires long term data on survival. Although long-term registry and peer-reviewed evidence is available for the ‘best-selling; brands such as the Exeter cemented stem and the Corail cementless stem, this is rarely available. This is partly because of the continued emergence of new or updated prostheses for which there is by definition only short term data on survival. But more generally, there is a lack of observational data - studies following a cohort of patients over a 15 to 20 year period are

very few, and not all countries have national joint registries; and even if they do, they have not been collecting data for a 15 or 20 year time period (see Table 2.2, chapter 2). Consequently, when reporting prosthesis survival rates typically the data are right censored as is the case in this chapter. As discussed in the previous section, the Swedish Hip Arthroplasty Register is the only joint registry reporting long term survival data of around 30 years.

With the creation of the NJR, eventually sufficient data will accumulate for England and Wales, and this will permit comparison of actual long-term survival rates. In the meantime, the NJR does now provide up to 7 years post-operation information on some patients, and it is these data which are analysed in this chapter.

3.2 Updating a previous study: Sibanda et al

To date, the only paper (apart from the NJR annual reports themselves) to have analysed the early revision rates reported in the NJR is by Sibanda et al[54]. This reports revision rates after primary hip and knee replacement in England between 2003 and 2006, although here I focus only on their findings on hip replacements.

The authors use data on 170,410 records of primary hip procedures, Their paper was only able to consider revision occurring within the first 3 years of surgery, and only between the four broad prosthesis type: cemented, cementless, hybrid and resurfacing, with revisions identified through the linkage with the HES database due to 'missing patient identifiers in the NJR'. The linkage of HES and NJR data-sets was carried out according to a five hierarchical linkage criteria described in detail in the paper. Revisions were identified using the OPCS-4 codes and revision rates were estimated using the Kaplan-Meier survival analysis method where time of death or September the 30th 2006 was the end of follow-up. Multivariable Cox regression was used to estimate hazard ratios for prosthesis type, age group sex and indication or surgery as risk factors for revision.

They report an overall revision rate for primary THR of 0.7% at 1 year and 1.4% at 3 years. They show that revision rates varied significantly according to prosthesis type ($p<0.0001$), concluding that patients receiving a cemented prosthesis had the lowest 3 year revision rate and

that the highest rate was for resurfacing. They also found that the differences in revision rates by prosthesis type were apparent even at 3 months post-surgery. In terms of patient characteristics, the pattern of revision rates was related to the patients' gender but not to age, and it was almost twice as high for other indications than OA. They also provide a brief comparison of revision rates reported by other national joint registries and conclude that rates observed in Australia and Norway are distinctly higher than those observed in England and New Zealand.

I now update this analysis of Sibanda et al, by using data reported in the 5th, 6th and 7th NJR annual reports. I also now disaggregate by individual brands. First, Tables 3.1 compare the characteristics of patients cited in the Sibanda paper (2003-2006), with the updated data since reported in the NJR annual reports for 2006/7; 2007/8 and 2008/9[9, 66, 67][12]. The two parts of the table confirm the increased popularity of cementless prostheses, at the expense of cemented (cementless rises from 25% to 41%, while cemented falls from 54% to 37%). Both tables show that younger patients are more likely to receive a cementless type and older patients more likely to receive cemented.

[12] It should be noted that there is some overlap with the data - table 2.8 does not report data for a full year, only up to 06/06. Data from the annual reports is only reported for each financial year.

Characteristic		Prosthesis type				
		Cemented	Cementless	Hybrid	Resurfacing	Overall
Age	<55	1634	2937	1313	3098	8982
	55-64	5998	6020	2733	2551	17302
	65-74	16642	6727	3776	516	27661
	75+	16957	3338	2298	37	22630
	Not recorded	1	0	0	0	1
Sex	Female	26512	10840	5969	2359	45680
	Male	14705	8172	4148	3843	30868
	Not recorded	15	10	3	0	28
Indication for surgery	Osteoarthritis	38854	17701	9274	5879	71708
	Other	2378	1321	846	323	4868
Total		41231	19022	10120	6202	76575

Source: Sibanda et al[54]

Table 3.1 (a) Characteristics of patients undergoing primary THR 2003-6 (4/03-9/06)[54]

Characteristic		Prosthesis type				
		Cemented	Cementless	Hybrid	Resurfacing	Overall
Ag e	<55 y	2743	11332	2410	5801	22286
	55-64	9100	22125	6071	4981	42277
	65-74	24429	26144	10576	1187	62336
	75+	29895	14875	9257	125	54152
Sex	Female	44074	42529	17912	3546	108061
	Male	22631	31947	10402	8548	73528
Indication for surgery	Osteoarthritis	66886	74485	27634	12649	181654
	Other	4672	5391	2828	635	13526
Total		66167	74476	28314	12094	181051

Source: NJR annual reports:[66, 67, 71]

Table 3.1(b) Characteristics of patients undergoing primary THR 2006-9 (4/06-3/9)[9]

Table 3.2 turns to revision rates. Part (a) is reproduced from Sibanda et al, reporting 1 year and 3 year revision rates disaggregated by gender, and part (b) updates by now including 2006/7 to 2008/9, and providing estimates of 5 year revision as well. This confirms the Sibanda et al finding that cemented prostheses perform better than cementless, in terms of the 3 year revision

rate. I can now confirm that this is also true after 5 years: 2% for cemented compared to 3.4% for cementless. This is clearly statistically significant: the 95% confidence intervals are non-overlapping (1.8-2.1% for cemented and 3.2-3.7% for cementless.) As can also be seen, hybrid implants (one component cemented and the other cementless) lie somewhere between, and hip resurfacing continues to have the highest revision rates at all time points.

Thus, this updating confirms the somewhat surprising result in Sibanda et al - a continued trend towards implanting more cementless prostheses, despite the fact that they have higher typical revision rates than cemented. It can now also be seen that this applies at 1, 3 and 5 year time points – indeed, the difference between mean revision rates actually increases over time: 0.7 at 1 year, 0.9 at 3 years and 1.4 at 5 years. It is also worth noting that variation in the revision rates by gender exist. In most cases these are not significantly different (the confidence intervals overlap), but the revision rate for cementless prostheses is signficantly different at 1 year for men is 0.4, compared to 1.0 for women and the confidence intervals do not overlap. One possible reason for gender differences in revision rates may be because women are less likely to 'wear' their joints out due to occupation or sport activity.

Category	**Men (C.I)**		**Women (C.I)**	
	1 year revision rate	**3 year revision rate**	**1 year revision rate**	**3 year revision rate**
Cemented	0.4 (0.3-0.4)	1.1 (0.8 to 1.5)	0.3(0.2 to 0.4)	0.8 (0.7 to 1.0)
Cementless	0.4 (0.3 to 0.6)	2.4 (1.9 to 3.1)	1.0 (0.8 to 1.2)	1.6 (1.3 to 2.0)
Hybrid	0.9(0.6 to1.3)	1.9 (1.3 to 2.6)	0.6 (0.4 to 0.8)	1.2 (0.7 to 1.9)
Resurfacing	1.5 (1.1 to 2.5)	1.9 (1.4 to 2.5)	1.9 (1.4 to 2.6)	3.7 (2.8 to 5.0)

Source: NJR annual reports:[66, 67, 71]; 95% confidence intervals in brackets.

Table 3.2(a) Revision rates by prosthesis type 2003-6 Sibanda et al[54]

Category	April 1st 2003 to 31st December 2009 (C.I)		
	1 year revision rate	3 year revision rate	5 year revision rates
Cemented	0.6 (0.6 to 0.7)	1.4 (1.3 to 1.5)	2.0 (1.8 to 2.1)
Cementless	1.3 (1.2 to 1.4)	2.5 (2.4 to 2.7)	3.4 (3.2 to 3.7)
Hybrid	0.9 (0.8 to 1.0)	1.8 (1.6 to 1.9)	2.7 (2.4 to 3.0)
resurfacing	2.1 (1.9 to 2.3)	4.3 (4.0 to 4.8)	6.3 (5.7 to 7.0)

Source: NJR annual reports:[66, 67, 71]

Table 3.2(b) Revision rates by prosthesis type 2003-9: NJR annual report[9]

3.3 A more disaggregated analysis at the individual prosthesis level

One of the limitations of the Sibanda paper, for current purposes, is that it did not compare revision rates at the individual brand level. Since the NJR annual reports do report survival rates for individual prostheses brands, I can now fill in this gap. Tables 3.3(a)-(d) show the revision rates of the most frequently used brands for each of cemented cup and stem and cementless cup and stem at 3 and 5 year time points, ranked in ascending order by 5 year revision rates.

+-

(a) cemented cups

Cemented cup - Brand	Manufacturer	No. patients	3 year Revision rate	5 year Revision rate
Low profile muller	Zimmer	2,316	0.7	1.2
Elite plus cemented cup	Depuy	6,466	0.9	1.3
Stanmore-Arcom	Biomet	1,554	0.9	1.4
Elite plus ogee	Depuy	13,730	1.0	1.3
Opera	Smith & Nephew	4,758	1.1	1.5
Contemporary	Stryker	23,320	1.2	1.9
Charnley cemented cup	Depuy	7,709	1.2	2.1
Charnley ogee	Depuy	7,254	1.5	2.2
Exeter duration	Stryker	7,519	1.6	2.2
ZCA	Zimmer	4,553	1.6	2.4
Cenator cemented cup	Corin	1,896	2.0	2.8
Apollo	Biomet	1,346	2.8	3.7

Source: NJR annual reports:[9, 58, 60, 64-67]

Tables 3.3 Revision rates for individual brands of prostheses (1/4/03 - 31/12/09) [9]

Cemented stem - Brand	Manufacturer	No. patients	3 year Revision rate	5 year Revision rate
MS-30	Zimmer	1,425	0.9	1.3
Stanmore modular	Biomet	2,938	1	1.4
C-stem	Depuy	8,372	1.3	1.6
Elite Plus	Depuy	1,188	1.2	1.8
Exeter V40	Stryker	67,015	1.3	1.9
Charnley	Depuy	13,565	1.3	2.1
CPS-Plus	Smith & Nephew	1,474	1.4	2.3
CPT	Zimmer	10,226	1.8	2.5
Muller STR	Zimmer	1,177	1.4	2.6
SP II	Link Orthopaedics	1,271	2.4	2.9
Muller -Biomet	Biomet	1,469	2.4	3.3
C-stem AMT	Depuy	2,260	1	n/a

Source: NJR annual reports:[9, 58, 60, 64-67]

Table 3.3 (b) cemented stems

Cementless cup - Brand	Manufacturer	No. patients	3 year Revision rate	5 year Revision rate
Trident	Stryker	16,079	1.7	2.4
Reflection cementless	Smith & Nephew	2,730	1.2	2.5
Triology	Zimmer	11,652	2	2.5
Pinnacle	Depuy	24,581	2.2	2.9
Duraloc cementless cup	Depuy	4,911	2.4	3.2
CSF	JRI	10,399	2.6	3.2
Exceed	Biomet	3,396	1.8	3.3
Plasma cementless cup	B Braun/Aesculap	1,296	2.6	3.3
Allofit	Zimmer	1,703	2.3	3.6
EPF-Plus	Smith & Nephew	3,734	3.1	4.7
CSF Plus	JRI	2,957	2.3	n.a

Source: NJR annual reports:[9, 58, 60, 64-67]

Tables 3.3(c) cementless cups

Cementless stem - Brand	Manufacturer	No. patients	3 year Revision rate	5 year Revision rate
Accolade	Stryker	4,184	2.5	2.8
Furlong HAC	JRI	13,977	2.5	3.1
Profemur	Wright Medical UK Ltd	1,004	3.1	3.1
Taperloc	Biomet	3,689	2.3	3.4
Bimetric	Biomet	1,834	2.8	3.4
Synergy	Smith & Nephew	2,156	2.1	3.7
Corail	Depuy	30,093	2.6	3.8
ABG II	Stryker	1,565	2.9	3.8
SL-Plus	Smith & Nephew	4,161	3.3	4.4
Versys	Zimmer	1,064	3.5	4.8
S-ROM	Depuy	1,018	4	5.5
CLS	Zimmer	2,332	3.1	5.9

Source: NJR annual reports:[9, 58, 60, 64-67]

Table 3.3(d) cementless stems

These tables allow me to establish four 'new' facts at the individual prosthesis level.

3.3.1 There are considerable variations in revision rates within each of the four types.

From Sibanda et al, it is only known in broad terms that, cemented prostheses display superior revision performance to cementless prostheses, but these are only averages for each type. What is not known is whether *all* cemented prostheses display superior survival than *all* cementless. In fact, these new tables reveal considerable variability within each type. For cemented cup, 5 year revision ranges from 1.2% for the Low Profile Muller (the 9th most commonly implanted) to 3.7% for the Apollo (the 12th most commonly implanted). For cementless cups, the range is from 2.4% for the Trident (the second most commonly implanted) to 4.7% for the EPF-Plus (the 7th most commonly implanted)[13], For cemented stems, the MS-30 has the lowest 5 year revision rate of 1.3%, while the Muller-Biomet has the highest at 3.3. For cementless stem, the lowest rate is the Accolade, 2.8% (the 3rd most commonly implanted cementless stem), while the highest revision rate is twice that: CLS with a revision rate of 5.9%. In other words, the ratio of lowest-highest 5 year revision rates is in the region of between two or three to one. While these numerical differences in revision rates are fairly small in absolute terms, the importance of the variation is still potentially important: the differences could widen at later time points i.e. at 7

[13] One prosthesis (CSF Plus) does not have a reported 5 year revision rate presumably because 5 year data are not yet available.

and 10 years; moreover, if NICE decided to revise its benchmark downwards, there is some indication that some prostheses may no longer be rated as successful.

This can be summarised concisely with the equations reported in Table 3.4, in which the three and five year revision rates of different brands are regressed simply against dummy variables for each prosthesis type[14]. This confirms that, as already known, both cementless stem and cup have higher revision rates than cemented stem and cup. But equally important, the R squared reveals that there remains an unexplained 50% component of the variance - this represents the magnitude of variations between different prostheses within each type. This suggests that it is misleading to ignore differences between brands within each type and that further analysis at the individual brand level is required[15][16].

	3 year revision rate			**5 year revision rate**		
	Coefficient	Std. Err.	P value	Coefficient	Std. Err.	P value
Cemented cup	-0.075	0.222	0.73	-0.155	0.287	0.59
Cementless cup	0.750	0.212	0.00*	1.005	0.286	0.01*
Cementless stem	1.442	0.213	0.00*	1.820	0.342	0.00*
Constant	1.450	0.145	0.00*	2.155	0.191	0.00*

*significant at the 0.05 level
†R-squared - 3 year revision rate: 0.5964, 5 year revision rate: 0.5488,
Robust standard errors are estimated to control for heteroscedasticity.
Sample size: n=47, covering 345,316 prostheses
Data used in regression analysis is sourced from the following NJR annual reports:[9, 58, 60, 64-67]

Table 3.4, Regressing revision rates against prosthesis type

[14] Cemented stem is the omitted default.
[15] Although it should also be recognised that there may also be some confounding caused by the relative proportions of different bearing types within stem and cup combinations that affect brand comparisons. Alongside this there could also be an element of a performance and intensity bias introduced from high volume centres and surgeons.
[16] In further regression experiments, I regressed the three and five year revision rates against dummy variables for manufacturers, but there were no significant differences between the manufacturers for 3 or 5 years and the R-squared is particularly low (0.2137 and -0.1664 for 3 and 5 year revision rates respectively). Similarly, when manufacturer dummies were added to the prosthesis dummies, this did not add significantly to the overall fit shown in Table 3.4.

3.3.2 The 3 year revision rate is an imprecise predictor of the 5 year rate

Of particular interest, now that there is more data on revision over a slightly longer time period, is whether the three year revision rate is a good predictor of the five year rate. This is particularly relevant to the question is explored in Chapter 5.

Table 3.5 reports the results of regressing five year revision rates against three year revision rates, with the individual brand of prosthesis as the unit of observation. It shows that the three year rate is a positively significantly determinant of the five year rate. This is as expected – a prostheses which requires revision within the first three years is also more likely to within the first 5 years. However, the regression coefficient (1.255) is not only significantly different from zero, it is also strongly significantly greater than unity, and this suggests that the differentials between prostheses tend to widen over time. The R-squared of 0.8492, although a quite good fit, is by no means perfect; in other words, if prosthesis A performs better than B over the first three years after implant, this may not be so after 5 years[17]. Figure 3.1 plots this relationship with the fitted regression line drawn in. As can be seen, there are a number of outliers. For example, there are three prostheses with a 3 year revision rate of 3.1; the fitted regression line predicts that, at that level of 3 year revision, 5 year revision should be 4.2. However, this is a large under-prediction for the CLS (actual 5 year revision of 5.9), a smaller under-prediction for the EPF Plus (actual 4.8), and large over-prediction for the Profemur (actual 3.1). If this magnitude of 'error' can occur with just an additional 2 years of experience, it calls into question whether even a 5 year rate is sufficient to predict what might happen over, say, the next 10 years after implant.

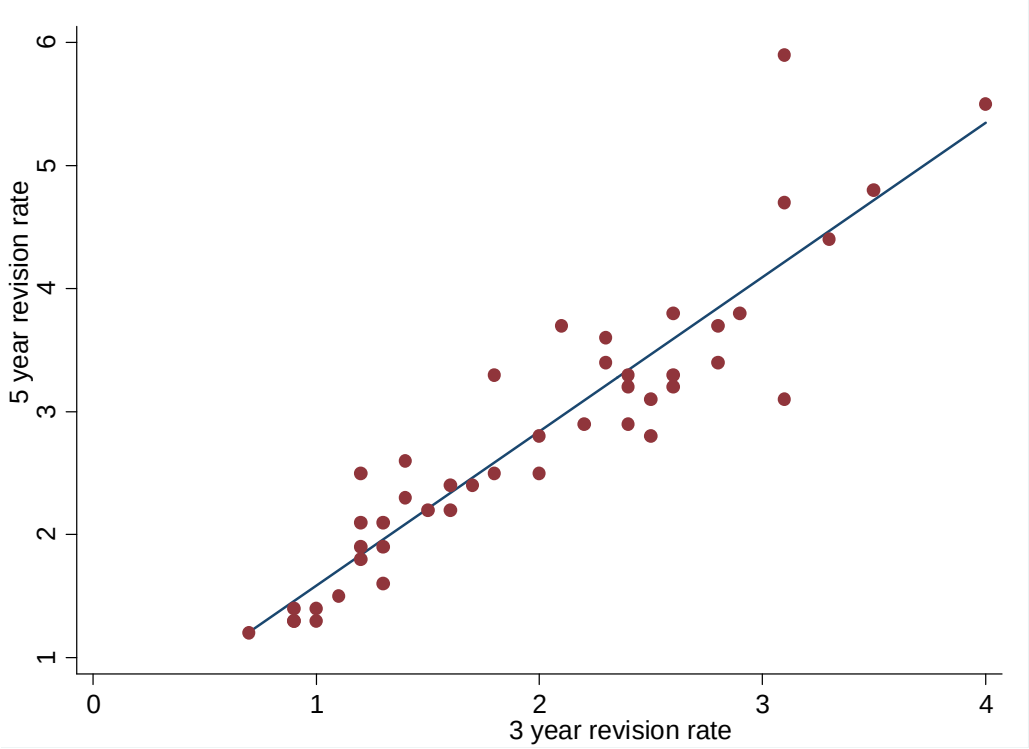
Revision rate 5 year	Coefficient	Std. Err.	P value
Revision rate 3 year	1.255262	0.08927	0.00*
Constant	0.325646	0.14503	0.03

* significant at the 0.05 level
†R-squared: 0.8492
Sample size: n=47, covering 345,316 patientsRobust standard errors are estimated to control for heteroscedasticity.
Data used in regression analysis is sourced from NJR annual reports:[9, 58, 60, 64-67]

Table 3.5 Regressing five year revision rates against three year revision rate (brand level)

[17] The overall rank correlation coefficient is similarly high (ρ = 0.93) and is also high for cemented cup and stem (ρ = 0.95). However, for cementless cup and stem they are noticeably lower (ρ=0.71 and ρ=0.65 respectively).

Figure 3.1 The relationship between three and five year revision rates



Data used in scatter plot is sourced from various authors of the following NJR annual reports:[9, 58, 60, 64-67]

Table 3.6 investigates whether the fit can be improved by adding dummy variables for the prosthesis types and manufacturer, but, as can be seen, none of the prosthesis types is significantly different from the default of cemented stem and only Wright Medical is significant (and negatively) different from the default dummy of Stryker (although Zimmer is significant at the 0.06 level).

	Coefficient	Std. Err.	P value
3 year revision rate	1.232602	0.123587	0
Cemented cup	-0.00468	0.175161	0.98
Cementless cup	0.187707	0.210217	0.38
Cementless stem	0.222212	0.256679	0.39
Braun	-0.18782	0.453351	0.68
Biomet	0.193338	0.223332	0.39
Corin	0.244125	0.450856	0.59
Depuy	0.162823	0.203601	0.43
JRI	-0.29344	0.332187	0.38
Link	-0.15359	0.467972	0.75
Smith & Nephew	0.445836	0.231396	0.06
Wright Medical	-1.03862	0.441671	0.03(†)
Zimmer	0.423043	0.212458	0.06
Constant	0.095347	0.256487	0.71

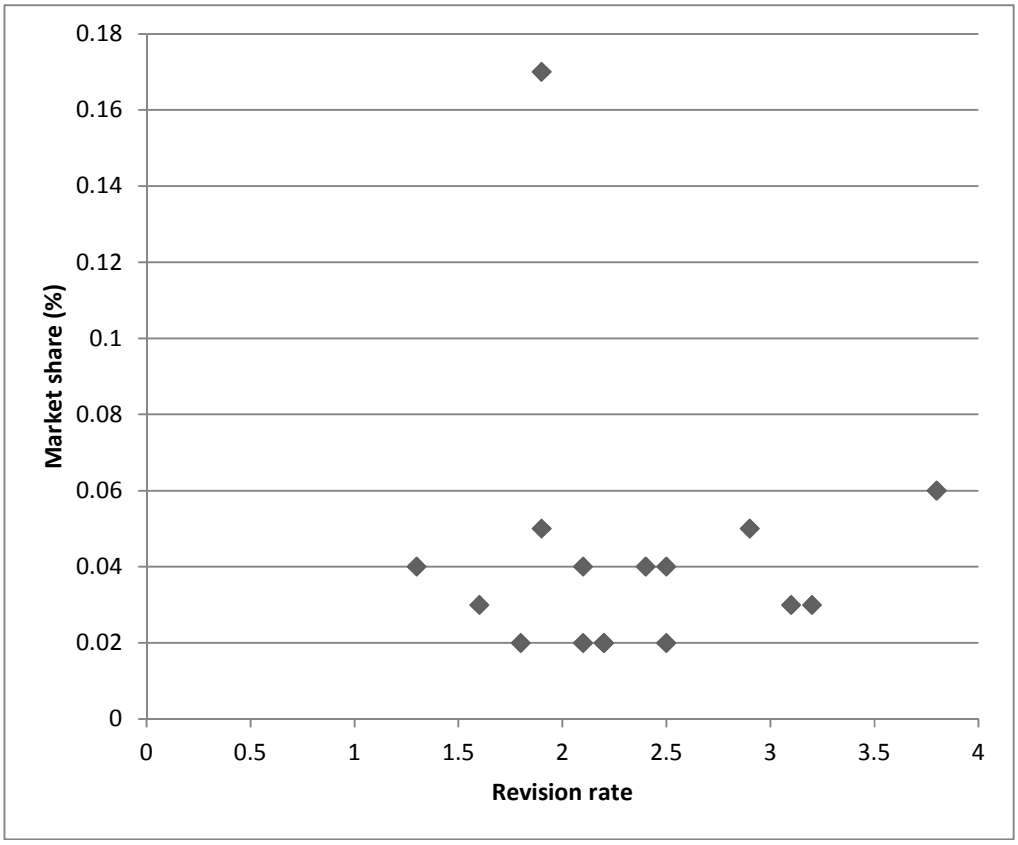
*default manufacturer dummy: Stryker Howmedica Osteonics, default type dummy: cemented stem
†significant at the 5% level
‡R-squared: 0.8735
Sample size: n=47, covering 345,316 prostheses
Data used in regression analysis is sourced from the following NJR annual reports:[9, 58, 60, 64-67]

Table 3.6, Regressing five year rates against three year rates controlling for manufacturers and prosthesis types

3.3.3 There is no evidence that better performing (in terms of lower revision) brands achieve higher market shares (i.e. greater use in hospitals.)

To test for this relationship, Figure 3.2 plots the market share of each prosthesis brand against its 5 year revision rate. The Exeter V40 stands out as the outlier with a 17% market share and a fairly low revision rate of approximately 2%. A simple regression equation reveals that there is no significant relationship between the two (this is true whether or not the outlier is excluded).

Figure 3.2 Relationship between manufacturer market share and revision rate



Data used in scatter plot is sourced from various authors of the following NJR annual reports:[9, 58, 60, 64-67]

3.3.4 There is no clear relationship between ODEP ratings, market shares and NJR revision rates

As explained in Chapter 2, the NJR also reports the ODEP[2] ratings for the most implanted prostheses brands. These new data on revision rates provide an opportunity to assess these ODEP ratings.

First, the ODEP ratings are reproduced here in Tables 3.7.

		ODEP Rating		
Cemented stem - Brand	**Manufacturer**	**2009**	**2008**	**2007**
Contemporary	Stryker	5A	5A	5A
Elite Plus Ogee	Depuy	10A	10A	10A
Elite Plus cemented cup	Depuy	10A	10A	10A
Exeter Duration	Stryker	10A	10A	10A
Opera	Smith & Nephew	10A	5B	5B
ZCA	Zimmer	10A	10A	10A
Marathon cup	Depuy	Pre-entry	Pre-entry	-
Charnley cemented cup	Depuy	10A	10A	10A
Charnley Ogee	Depuy	10A	10A	10A
Low Profile Muller	Zimmer	Unclassified	Unclassified	Unclassified
Stanmore-Arcom	Biomet	10A	10A	10A
Cenator cemented cup	Corin	3A	3A	3A
Apollo	Biomet	Pre-entry	Pre-entry	Pre-entry

(a) cemented cup [2]

		ODEP Rating		
Cemented stem - Brand	**Manufacturer**	**2009**	**2008**	**2007**
Exeter V40	Stryker	10A	10A	10A
CPT	Zimmer	10A	10A	7A
Charnley cemented stem	Depuy	10A	7A	10A
C-stem cemented stem	Depuy	10B	10B	10B
C-stem AMT cemented stem	Depuy	3A	3A	-
Stanmore modular	Biomet	10A	10A	10A
MS-30	Zimmer	10B	10B	10B
CPCS	Smith & Nephew	3A	Pre-entry	Pre-entry
CPS-Plus	Smith & Nephew	7A	7A	7A
Muller-Biomet	Biomet	5B	5B	5B
Muller STR stem	Zimmer	10A	10A	10A
CCA cemented stem	Mathys Orthopaedics Ltd	10A	10A	10A
Furlong cemented stem	JRI	10C	10C	10C

(b) cemented stem[2]

		ODEP rating		
Cementless cup - Brand	**Manufacturer**	**2009**	**2008**	**2007**
Pinnacle	Depuy	7A	5A	3A
Trident	Stryker	5A	5A	3A
Triology	Zimmer	7A	7A	7A
CSF Plus	JRI	Unclassified	Unclassified	Pre-entry
Exceed	Biomet	5A	5A	5A
CSF	JRI	10A	10A	10A
EPF-Plus	Smith & Nephew	3A	3A	3A
Reflection cementless	Smith & Nephew	7A	7B	7B
Duraloc cementless cup	Depuy	10A	10A	10A
Allofit	Zimmer	5A	5A	5A
Procotyl	Wright Medical UK Ltd	Pre-entry	Pre-entry	-
Trabecular metal cementless	Zimmer	5A	Pre-entry	Pre-entry
Plasma cementless cup	B Braun/Aesculap	5A	5A	5A

(c) cementless cup [2]

		ODEP rating		
Cementless stem - Brand	**Manufacturer**	**2009**	**2008**	**2007**
Corail	Depuy	10A	10A	10A
Furlong HAC	JRI	10A	10A	10A
Accolade	Stryker	5A	5B	3B
Taperloc cementless stem	Biomet	10A	10A	10B
SL-Plus cementless stem	Smith & Nephew	10A	10A	10A
CLS cementless stem	Zimmer	10A	10A	10A
Synergy cementless stem	Smith & Nephew	5A	5A	5A
Profemur cementless stem	Wright medical UK Ltd	Pre-entry	Pre-entry	Pre-entry
Bimetric cementless stem	Biomet	10A	10A	10A
Versys cementless stem	Zimmer	Various	Various	3A
S-Rom	Depuy	10A	10A	7B
ABG II cementless stem	Stryker	5B	5B	5B

Sourced from NJR annual reports:[9, 58, 60, 64-67]

(d) cementless stem[2]

Table 3.7- ODEP ratings for the most popular prosthesis brands

According to the NHS Supply Chain[2], ODEP ratings are based on information from the NJR, peer-reviewed literature; and from the manufacturers themselves. Manufacturers are requested to keep ODEP informed of all commercially available prostheses which are involved in post market

clinical follow-up studies. The two dimensions to the ODEP classification are explained in Figure 2.1 in chapter 2; the numerical rating indicates the number of years evidence on which the rating is based, and the alphabetic rating indicates the strength of the evidence (10A being strongest.)

Inspection of Table 3.7 reveals that over half (26/51) of the listed prostheses are awarded a 10 rating, and 23 of these are rated as A. More interestingly, 3 prostheses are given only a B or C rating after 10 years, and one of these (Furlong cemented stem with 10C) now only has 2 years to improve its data before it is deemed unacceptable. Second, some of these (most popular) prostheses are being implanted despite there being no outcome data available to ODEP (CPCS cemented stem; profemur cementless stem; CSF, procotyl and Trabecular metal for cementless cups). Some of the other most implanted prostheses are based on only weak evidence, for example, at the 5 year level i.e. the Accolade and ABG II cementless stems and the Muller Biomet cemented stem.

Table 3.8 reinforces these findings by presenting the ODEP ratings alongside the NJR 5 year revision rates and market shares for the 16 most used prostheses. Although ten of the 16 are rated 10A, as noted earlier, the Furlong only receives a 10C rating from ODEP (and S-Stem is only rated 10B); two others (Trident and Contemporary) only have a 5 rating – in other words, their popularity is in spite of a limited evidence base. Interestingly, some of these most popular brands are recording relatively high revision rates, notably the Corail, which has secured a large market share and a 10A rating, in spite of a relatively high revision rate of 3.8%[18].

[18] It should be noted that some brands of prostheses are implanted with a relative mix of bearings, for example: the Corail is often implanted with hard-on-hard bearings which are reported to have a slightly higher revision rate than metal on polyethylene bearings.

Brand	ODEP rating	Market share	5 year revision rate
Exeter V40	10A	0.17	1.9
Corail	10A	0.06	3.8
Contemporary	5A	0.05	1.9
Pinnacle	7A	0.05	2.9
Charnley stem	10A	0.04	2.1
Elite Plus Ogee	10A	0.04	1.3
Trident	5A	0.04	2.4
Trilogy	7A	0.04	2.5
CSF	10A	0.03	3.2
C-Stem	10B	0.03	1.6
Furlong	10C	0.03	3.1
Charnley cup	10A	0.02	2.1
Charnley Ogee	10A	0.02	2.2
CPT	10A	0.02	2.5
Elite Plus	10A	0.02	1.8
Exeter Duration	10A	0.02	2.2

Source: NJR reports: [58, 60, 64, 65, 71-73]

Table 3.8, ODEP rating, market share and revision rate by top 16 prosthesis brands

3.4 Conclusions

Ideally, if alternative prostheses are to be evaluated comparatively, we need reliable evidence on their long run survival/revision performance. However, long run survival data are still unavailable on a widespread scale, and this is why the extrapolation techniques analysed in chapter 5 are so potentially important. As a preliminary to that analysis, this chapter assesses what we can already learn from the early results on shorter-term survival, as published in the first seven years of the NJR.

The only published paper to date on survival rates using NJR data is by Sibanda et al. Their main finding was that cemented prostheses perform better, in terms of their 3 year survival rates, than do the increasingly popular cementless. This chapter now updates their analysis to include 5 year revision data, and finds that the Sibanda result still holds. Beyond this, it extends the depth of the analysis by examining revision rates for individual brands of prosthesis. It has four main results: (i) there are quite large variations in revision rates between different prostheses within each of these broad types; (ii) the three year revision rate is an imperfect predictor of the 5 year

rate, in other words, data on early revision may give a misleading impression of revision in the longer run, (iii) there is no apparent tendency for the prostheses with the lowest revision rates to be the most commonly implanted in the NHS; and (iv) the ODEP classifications do not appear to be closely related to the emerging evidence on revision rates from the NJR[19].

[19] Consideration of the articulation type is beyond the scope of this thesis.

Chapter 4, A review of the economic evaluation literature comparing the alternative hip prostheses used in THR surgery[20]

4.1 Background

The objective of this chapter is to critically appraise and summarise current published evidence on the costs and cost-effectiveness of using alternative prostheses in THR surgery[74, 75].

Economic evaluation (see section 1.4.1, for an introduction to economic evaluation) is widely used to inform policy decisions regarding which new healthcare technologies should be adopted given the available resources[76]. NICE provides guidance to the NHS in England and Wales on the clinical and cost-effectiveness of new and already developed technologies and within this, provides recommendations on the principles and methods of health technology appraisal[77].

From an economic perspective, some or all of the direct medical costs of implanting a new or alternative hip prosthesis may be offset by reductions in the subsequent direct medical costs associated with complications and/or secondary intervention and also by an earlier return to productive activity.

Health care purchasers (in the NHS, surgeons and clinical or finance managers) are motivated by a desire to buy the most effective prostheses for patients but are also constrained by health budgets, meaning they increasingly demand greater 'value for money' from the prostheses. Potential important differences in non-medical resource use and costs may also result from the use of different prostheses. These include productivity losses (absence from paid/unpaid work) associated with differing lengths of rehabilitation/functional status; other patient out-of-pocket expenses (e.g. travel costs); impact on social care services (both publicly and privately funded; community and domiciliary care).

In the UK, the ODEP[53] (section 2.1.3) provides a rating for prostheses based on data submitted by the manufacturers. For example, the Charnley cemented cup and stem both have a rating of 10A, designating strong clinical evidence of prosthesis survival at 10 years (NICE

[20] A version of this chapter has been published in a peer reviewed journal[66], a parallel piece of work on knee prostheses was published as an editorial in 2009[67].

benchmark)[20]. However, to date, no studies have systematically summarised current economic evidence to compare the impact of different types of prostheses on costs and cost-effectiveness.

This chapter specifically aims to:

1. Assess the completeness of the evidence base for resource use, costs and cost-effectiveness;
2. Assess the applicability of the available evidence to inform resource allocation decisions in the UK NHS.

Section 4.2 describes the methods used; section 4.3 describes the identified studies and their results; section 4.4 discusses and section 4.5 concludes. There are also four appendices (Appendices 3-6).

4.2 Methods

The search strategy criteria to identify relevant papers and approach to data extraction is described below:

4.2.1 Criteria for considering studies for this review

Types of studies: Full economic evaluation studies (cost-effectiveness analyses, cost-utility analyses or cost-benefit analysis), defined as the comparative analysis of alternative courses of action (e.g. healthcare treatments) in terms of both their costs and their consequences (e.g. clinical effects)[78]. Partial economic evaluation studies which compare alternatives in terms of their costs only (i.e. cost analyses)[78]. (See figure 1. Chapter 1.)

Types of participants: Adults 18 years or over.

Types of Interventions: Any THR surgery using any type of hip prosthesis (using any surgical technique) compared to THR surgery using any other type of prosthesis (any surgical technique)[21]. [11]

Types of outcome measures:

[21] Cost effectiveness studies comparing THR surgery with 'no intervention' were not included in the review. This review is concerned with the alternative prostheses used in THR surgery and their comparative success. Arguably, in fact surgery is the only real treatment option for patients with end stage joint degeneration of the hip [11].

1. Direct medical resource use: prosthesis, operative time, post-operative care, length of post-operative hospital stay (los), management of surgical/ implant/post-operative complications, medication, use of therapy services, use of adult social care services, revision surgery within follow-up period, long-term revision surgery (prosthesis failure)

2. Non-medical resource use: productivity losses (sick days, lost wages) - patient: productivity losses (sick days, lost wages) - informal carer(s): other patient/family out-of-pocket expenses (travel to hospital visit)

3. Health effects; Post-operative pain, surgical/implant/post-op complications, physical functioning, health related quality of life (HR-QoL), mortality/survival, quality adjusted life years (QALYs),

Note that direct assessments of revision and bilateral surgery are excluded as they are not within the scope of the thesis.

4.2.2.Search methods for identification of studies

Electronic searches

I searched MEDLINE (1950 to May 2010); EMBASE (1980 to 2010 week 20) Cinahl (1971 to May 2010); The Cochrane Library (Issue 5, 2010): The Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment (HTA) database; Health Economic Evaluations Database (HEED) (1992 to 6 June 2010); the NHS Economic Evaluation Database (NHS EED) (1992 to 6 June 2010) and the European Network of Health Economic Evaluation Databases (EURONHEED) (2000 to 6 June 2010).

A search strategy was developed and adapted for use in each electronic database. An example of the search strategy used in OVID Medline is provided in Appendix 3.

Searching other resources

Grey literature searching was outside the scope of this review. However, I reviewed bibliographies of the included economic evaluations to identify additional eligible economic evaluations.

4.2.3. Data collection and analysis

Selection of studies

The titles and abstracts of the literature search results were screened for eligible economic evaluations. Full text reports of all eligible studies were sought. Excluded studies were listed with the reasons for their exclusion. Articles published in languages other than English were excluded since translation was outside the scope of the current review.

Data extraction and management

Data extraction was carried out using a two-stage process[79]. First, risk-of-bias in generating clinical effect estimates utilised in each economic evaluation (if applicable) was assessed using a tool endorsed by the Cochrane Bone, Muscle and Joint Trauma Group[49]. Study quality was assessed using a more general tool, the Critical Appraisal Skills Programme (CASP) checklist for: (i) cohort studies [80] and (ii) randomised controlled trials[81]. Next, an overall assessment of the methodological quality of each economic evaluation was made, informed by applying the guidelines for authors and peer reviewers of economic submissions to the British Medical Journal (BMJ) and, in the case of model-based full economic evaluations, a checklist for best practice guidelines in decision-analytic modeling[82]. An example of a completed data extraction form is presented in Appendix 4.

Data Synthesis

The extracted data were synthesised by summarising the methodological quality of each study in tables, these tables were then supplemented with a narrative summary. All estimates of costs reported in the literature were converted to British currency values (GBP) using exchange rates based on Purchasing Power Parities and inflated to 2008 prices using a web-based conversion tool[83]. Results are reported according to: study type, perspective, comparator, study design, time horizon, data sources, health benefit measures, discount rate, uncertainty and sponsorship.

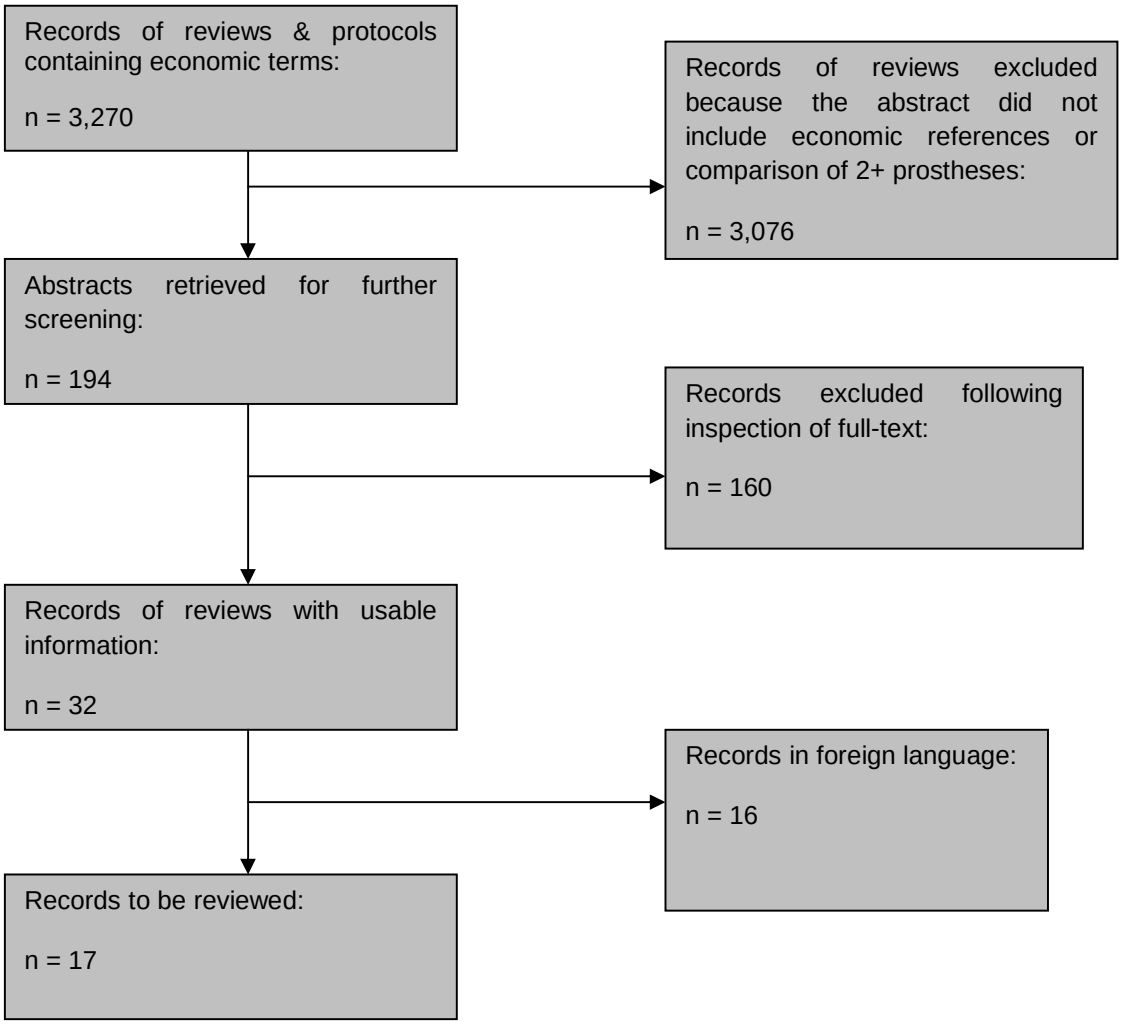
4.3 Results

4.3.1 Description of studies

Results of the search

3,270 papers were retrieved by electronic searches (Figure 4.1). Of these, 194 potentially eligible abstracts were retrieved for further screening. Papers were excluded if they did not compare two or more prostheses, or were not a full or partial economic evaluation. 16 studies identified for possible inclusion are not reported in English and in some cases did not include an English language abstract, these studies are not included in this review. A total of 17 potentially eligible studies were identified amongst the 194 abstracts and are therefore included in this review.

Figure 4.1 - Quorum statement flow diagram [79]



Included studies

Appendix 5 provides a summary of the 17 included studies based on the Drummond et al checklist for economic evaluation studies [84]. A narrative summary of the characteristics and methods of included studies is presented below.

Study Design: Ten studies are classified as full economic evaluations (cost-effectiveness analyses [85-89] and cost-utility analyses [22, 90-93]; no eligible cost-benefit analyses were identified. These studies either employ the survival rate of the prosthesis as the measure of health benefit [85-89], or combine survival and HR-QoL measures to calculate QALYs [22, 90-93]. Nine studies are model-based evaluations and these can be further classified into two sub-groups: (i) deterministic models (e.g. Daellenbach et al [87]) and (ii) probabilistic Markov model (e.g. Briggs et al [22]). The stated purpose of some of these studies is largely methodological [22, 86, 87, 92]; they aim to develop a methodology which can also be applied to other healthcare interventions, using THR and the specific prostheses as an illustrative example to demonstrate a more widely generalisable modelling approach. However, this fact does not limit the reliability of the findings of these studies. Indeed, results from Briggs et al [22] have been used to inform NICE guidelines on hip prostheses [20]. One CUA is a retrospective cohort study conducted using additional questionnaire data [93].

Seven studies [94-100] are classified as cost analyses. Average total costs per patient by treatment group (surgery or prosthesis type) are the main outcome measures reported in these seven studies.

Country: Seven studies were based primarily on UK data, with the others based primarily on data from Australia, USA, Sweden, New Zealand, Germany, Italy Israel and Belgium. Full economic evaluations using revision rates for prostheses derived from populations outside of the UK [22, 86, 87, 93] would need to be further examined for differences in patient characteristics and surgical implantation techniques before results could be applied to the UK setting. Cost analysis studies [94, 96, 98, 100] using data from outside of the UK are based on different health care systems with differing study populations, thus generalisability of these results to the UK setting are of limited use other than to explore cost variation of prostheses as a component of THR

surgery. Furthermore, some of the older studies using UK data are of limited use in terms of the relevance to current NHS practice[101].

Interventions: Only one full economic evaluation conducts a head-to-head comparison between two specific brands of hip prostheses[22]. Four studies compare the Charnley prosthesis with an unspecified alternative (Appendix 5) and ten studies report the comparison as either 'cemented vs cementless' or 'cemented/or hybrid' (see Appendix 5), with no brand information. Scheerlink et al[96] make cost comparisons across three different brands of prostheses and an unnamed 'other'.

Time horizon: NICE[77] recommends using a time horizon sufficiently long to reflect all important differences in costs and outcomes between the alternatives under evaluation. In this case, hip prostheses can last for up to approximately 20 years following implantation[53]. As Appendix 5 reports, a variety of time horizons are used for model-based economic evaluations included in this review, ranging from five years[91] to 60 years[22, 90, 92].

Analytic perspective: General guidance on conducting an economic evaluation recommends adopting a broad societal analytic perspective as the gold standard, but it is widely recognised that a narrower analytic perspective (e.g. health care system) may be sufficient if the purpose of the evaluation is to inform decisions that will be made within a narrower constituency (e.g. health care system)[84]. All studies identified in this review consider only those costs (resource use) relevant from the perspective of the health care system. One study[87] mentions the wider perspectives of society and the patient but resource use and costs that would be relevant from these perspectives are not included in the analysis.

Outcome measures of health gain: Five of the full economic evaluation studies report survival rate of the prosthesis as the primary measure of health benefit; either as an observed rate (see Appendix 5), or a rate statistically extrapolated over a longer time horizon. Three studies[88-90] report survival rates for prosthesis types, varying the length of years through sensitivity analysis of the extrapolated survival rates at which survival was recorded. In general, there is a lack of long-term prosthesis survival data. In order to overcome this difficulty, studies employ statistical extrapolation of prosthesis survival data over a longer time horizon. Briggs et al[22] examine a

range of parametric survival models and conclude that the Weibull distribution fits best to the data; the data are then extrapolated over 60 years.

While survival is a useful measure of health gain, QALYs have the advantage that they combine length of survival with quality of life. Thus they enable comparisons between different health-care interventions in terms of a single measure of relative efficiency (i.e. cost per QALY), informing resource allocation decisions based on considerations of allocative efficiency across interventions[102]. Five economic evaluation studies used QALYs as their composite measure of health benefit[22, 90-93]. However, only Briggs et al[22] and Givon et al[93] conducted primary research on HR-QoL in a THR patient population to inform QALY estimates. Briggs et al used the EQ-5D questionnaire and Givon et al used the Rosser index to inform QALY estimates.

Direct medical resource use, unit costs and costs

Table 4.1 records the unit costs of the prostheses reported in each study: it shows the range between the cheapest and most expensive for the two broad types of prosthesis, and then for specific named prostheses within each type. In general, cemented prostheses were cheaper than cementless, ranging (in this literature) from £691 (Multicentre)[99] to £2,845 (Beuchel Pappes)[99] for cementless, ranging from £455 (Stanmore)[99] to £1,693 (Titan)[99] for cemented.

	Min cost (literature)	Max cost (literature)
CEMENTED (Mean)	£515 [21]	£1,084 [30]
Charnley	£395 [8]	£943 [29]
Stanmore	£455 [33]	£990 [29]
Titan	£1,693 [33]	£1,693 [33]
CEMENTLESS (mean)	£1,819 [31]	£5,785 [34]
Multicentre	£691 [33]	£960 [33]
Spectron	£903 [8]	£1,134 [22]
Buechel Pappes	£2,845 [33]	£2,845 [33]
HYBRID (mean)	£1,886 [32]	£4,452 [34]

Table 4.1 - Prosthesis costs (rebased to 2008 prices, in GBP) [83]

The average total cost of the THR procedure per patient reported in the studies ranges from £4,599[89] to £8,078[96]. Most studies reporting resource use and costs alongside the cost of the prosthesis assume these to be equal for each prosthesis type[99].

According to Scheerlink et al[96] implantation of the prosthesis (including the prosthesis itself), accounts for the second largest component of the total cost of THR surgery (21.3%), with hospital length of stay (LOS) being the largest component. The reported range of mean LOS in days is from 7.3[99] to 23[97] with mean costs varying from £2,101[89] to approximately £7,081[88] (obtained through sensitivity analysis).

The range for duration of surgery (theatre time) is 60 to 246 minutes[96]. Unnanuntana[94] is the only study to report duration of surgery separately for cemented, cementless and hybrid (femoral stem), finding that operative time for a cementless stem is approximately 20 minutes less than for both hybrid and cemented stems. Reported costs for duration of surgery show wide variation from £1,128[90] to £6,176 (obtained through sensitivity analysis) [88]. Scheerlink et al[96] reports medication costs as approximately 9% of the total cost of the procedure, breaking them down according to prosthesis brand, but reporting no apparent differences.

Non-medical resource use: No studies report non-medical resource use. However, it is anticipated that if they were included in an economic evaluation, the overall impact would be to increase the cost-effectiveness of THR surgery as an intervention. However, it is not anticipated

that this would lead to large additional variations between the alternative prostheses because the majority of patients receiving surgery are of retirement age and their main costs incurred are 'out of pocket' expenses such as cost of travel and alternative medicines. These costs are not anticipated to differ dramatically between alternative types of prostheses.Ideally it would be good to confirm this nonetheless.

Data sources used to populate the model: Nine studies used primary research to inform their analysis (for example, as discussed above, Briggs et al elicited HR-QoL data from THR patients). The remaining eight all used purely secondary data sources.

Sensitivity analysis: Only one of the full economic evaluation studies [93] does not conduct sensitivity analysis to address uncertainty. In their 2009 guidance, NICE describe three types of potential selection bias or uncertainty to consider: Structural uncertainty (categorisation of different states of health and the representation of different pathways of care); source of values to inform parameters and parameter precision (uncertainty around the mean health, and cost inputs in the model).

Daellenbach et al[87] perform sensitivity analysis on the 'less-reliable' input data defined as: the intangible costs of re-operation surgery (implicitly including those of the patient) and the expected failure rate of the prosthesis. Baxter and Bevan[88] perform sensitivity analysis on many of the parameters of their model, identifying the main cost drivers (hospital costs, prosthesis price and revision rates). Gillespie et al[86] conduct sensitivity analysis on the 'break-even price ratios' for hypothetical prostheses at various years using four hypothetical rates of prosthetic failure. Briggs et al[22] and Spiegelhalter and Best[92] use probabilistic sensitivity analysis (PSA) applied to parameter uncertainty in the model, conducting sub-group analysis by age and gender. Marinelli et al[91] also perform sensitivity analysis on revision rates, prosthesis costs, preoperative mortality, infection rates and utility values, however the details of the approach employed are not fully reported.

Risk of bias

The reliability of any full economic evaluation depends in part on its use of reliable clinical data, including data on beneficial and adverse effects, complications and secondary interventions[79]. Most of the included studies use observational data, such as from joint registries, to inform their

analysis. Although Randomised Control Trials (RCTs) are often thought of as the gold standard to inform economic evaluation studies[103], evaluation of THR is a context where the use of RCTs is of limited use in terms of the nature of the procedure, due to the long-term follow-up to observe time until revision surgery. Appendix 6 tabulates the assessment of risk of bias. No studies report blinding or randomization of participants. Appendix 6 shows that of the seventeen studies, inclusion and exclusion criteria are stated in five studies, and the intervention and outcome measures are defined in thirteen and fourteen respectively.

Discount Rate: All but one[93] of the full economic evaluation studies use a discount rate to account for time preference of costs and benefits which accrue in the future, varying from 5 to 6% for costs and 1.5 to 6% for benefits.

4.3.2 Summary of main findings

Table 4.2 reports the Incremental cost effectiveness ratios (ICERs)[22] for economic evaluations studies which report them[22, 91, 92] (the extra cost per unit of outcome obtained, in comparing one treatment with another)[104].[93] It is important to note here that the limited reporting of the methods for Marinelli et al[91] makes the strength of their findings difficult to assess. Although Speigelhalter and Best[92] calculate they state that their results should "not be taken as contributing in any way to guidance as to an appropriate prosthesis" (pg 3692) as they use . The remaining 13 studies do not report ICERs as they do not include a HR-QoL outcome in their study. The results in table 4.2 reveal that it is difficult to draw any conclusions from these disparate results.

[22] All ICERs reported in this table are incremental costs per QALY gained.

Study	Age of patient (yrs)	ICER	Age of patient	ICER
	Males		**Females**	
Briggs (2004): Charnley vs Spectron	80 years	3,768	70 years	673
	90 years	11,697	80 years	7,000
			90 years	18,839
Spiegelhalter (2003):** Charnley vs hypothetical alternative	55-64 years	581	55-64 years	537
	65-74 years	5,190	65-74 years	4,710
	75-84 years	13,220	75-84 years	12,030
	> 84 years	21,830	> 84 years	18,790
	Pre-op QALY score at baseline	**Cemented**	**Cementless**	**Hybrid**
Givon (1998)*:** cemented vs hybrid vs cementless with and without hydroxyapatite coating	0.50	7,749	10,241	10,352
	0.60	10,329	13,108	13,290
	0.70	15,484	18,203	18556
	0.80	30,732	29775	30,732
Marinelli (2008): cemented vs cementless broad types	Cementless prosthesis £48			

*Costs rebased to 2008 prices, in GBP [83]
** illustrative only, authors state results should "not be taken as contributing in any way to guidance as to an appropriate prosthesis" [92]
*** reported at baseline assessment.

Table 4.2, Incremental cost effectiveness ratios from reported studies

Other Results

Daellenbach at el[87] conclude that the higher cost cementless prostheses must last 6 to 9 extra years before revision surgery in order to yield the same expected present value as a cemented prosthesis. Fitzpatrick et al[90] report that of the cemented prostheses, the Charnley, Stanmore and Exeter perform relatively well in terms of time until prosthesis failure. Based on their model, they report that a cementless prosthesis costing approximately 300% more than the Charnley or other established prostheses was unlikely to reduce the revision risk sufficiently to produce any cost savings. Two studies[88, 89] report results for the Stanmore and Charnley by calculating the total expected cost of the prostheses over 20 years, reporting that the Stanmore is slightly more cost-effective than the Charnley.

4.4 Discussion

This review has systematically searched for, assessed and summarised literature on the costs and cost-effectiveness of using alternative prostheses in THR surgery. It has identified several methodological problems in the literature including a lack of observed long term prosthesis survival data, limited up-to-date UK based evidence and exclusion of patient and societal perspectives.

Several limitations of this systematic review should be highlighted when interpreting these principal findings. Foreign language studies were considered outside the scope of this review, thus sixteen studies were excluded. For all foreign language studies, English language abstracts were sought to further determine whether the study met the inclusion criteria, in some cases no abstract at all or no English language abstract was available. In the remaining cases it was not clear from the abstract whether or not the study would meet the inclusion criteria. From screening titles, all foreign language studies appear to be partial economic evaluations and thus the generalisability of the study to the UK context (for the purpose of this review) is anticipated to be limited due to international differences in health care settings. Hand searches and grey literature searches were not undertaken.

Only seven studies were based primarily on UK data, with some of the older studies being of limited use in terms of the relevance to current NHS practice. Where studies were non-UK based, revision rates for prostheses derived from populations outside of the UK require further detail of patient characteristics and surgical implantation techniques before results can be applied to the UK setting. Cost analysis studies have generally been based on different health care systems with differing study populations, thus limiting the applicability of these results to the UK, NHS context.

One of the methodological limitations of the studies identified in this review is the different types of economic models used, making comparability across studies difficult: none of the studies compared alternative models to answer the same question. The main difference between the types of model identified in this review is in the description of disease progression. Markov modelling[22, 90-92] involves dividing a patient's possible prognoses into a series of health states. The probabilities defining the transitions between each of these states are specified over a

single cycle of the model[90]. The model is then run over a number of cycles to view how a typical patient would move between states over a specified time period, consisting of several cycles. The transition probabilities reported in the Markov models in this review are calculated based on data obtained from a range of different sources, including life tables, clinical trials and other published sources. Crucially, because the empirical studies typically observe data used to generate transition probabilities over a limited follow-up period, the authors also employ statistical methods to extrapolate beyond the time horizon of observed data, for example the risk of revision. The Markov models identified in this review, are also fully probabilistic in their approach to managing uncertainty in the model parameters, NICE now requires the use of PSA for all cost effectiveness submissions[77].

On the other hand, the deterministic cost-effectiveness models (Daellenbach et al)[87] use more simplified assumptions. A key difference relates to the treatment of prosthesis survival rates. While studies using a Markov approach allow for the possibility that a prosthesis may fail at any point in time (according to a probability distribution), deterministic models assume a range of values for the expected life of a cemented prosthesis, and then determine, for each of these values, the increase in the expected life of a cementless prosthesis required in order for the two to have the same net present value cost (for various age groups). This assumes that a prosthesis will fail at a specific point in time. Other studies[85, 86, 88, 89] use a similar approach. Faulkner et al[89] estimate expected costs over twenty years using data from other studies and using statistical extrapolation to predict future revision rates.

A significant knowledge gap and challenge to research in this area relates to observed survival rates. NICE currently define their benchmark for revision rate as being 10% at 10 years[20]. Some studies in this review have employed methods of extrapolation of the data in order to estimate survival rates into the future. However, these are based on very short time periods of observed data. This highlights the need for more trials comparing different prostheses with long-term follow up. Only one full economic evaluation carried out a head-to-head comparison between two different manufacturer named prostheses[22]. Further economic evaluations of the prostheses according to their manufacturer rather than type (cemented/cementless) are needed, given the large number of prostheses, the likely variability within specific types of prostheses and the technological changes that have occurred over time. It is recommended that clinical trials

should include an economic evaluation during pre-trial modeling (employing a Bayesian iterative approach), which would inform the trial design and subsequent extrapolation of trial data[105].

In order to comprehensively assess whether an intervention provides value-for-money, information on non-medical resource use and productivity losses should also be sought and taken into account, even though not required in assessment guidelines for some agencies (e.g. NICE). Failure to take into account these costs and benefits may hide the fact that they are being merely shifted onto another sector[106]. We have identified very limited consideration of the patients' and society's costs and resource use in the literature. Baxter and Bevan[88] recommend further research combining prosthesis survival and HR-QoL.

This review also highlights the lack of up-to-date published studies using UK data, fourteen out of the seventeen studies included in this review were conducted over five years ago. The recent development of the NJR will provide an opportunity to produce more up-to-date analysis using data from England and Wales.

Finally, the range of costs of prostheses from table 4.1 provides an interesting perspective regarding the NHS national tariff for primary THR (an individual tariff is derived for each hospital patient episode, represented by the average cost of providing a particular procedure)[107]. This tariff specifies how much hospitals are reimbursed for treatments, in 2008/9 this was £5,220 for cemented and £5,587 for cementless prostheses (2008/9)[108]. The tariffs include a component for length of stay (currently £4,262 and £4,193 respectively)[108], implying very low tariffs for the surgical procedure itself (about £1,000 and £1,400 respectively). This is deserving of further research, to understand the potential tradeoffs that could occur across the range of prostheses in terms of 'profit' versus effectiveness.

4.5 Conclusions

This review highlights the need for more clinical trials including economic evaluations [109] and comparing different prostheses with long-term follow up. The establishment of the NJR for England and Wales provides a unique opportunity to address this gap as the registry collects

longer term data into the future. It also enables international comparisons of those countries with existing joint registries. Moreover, the recent introduction of PROMs (Patient Reported Outcome Measures) nationally for hip and knee replacement surgery will help to address the observed gap in the literature on the perspective of the patients undergoing surgery.

The next chapter explores methods for extrapolating current survival rates into the future in order to address the problem of lack of long term survival data on hip prostheses.

Chapter 5, Extrapolating survival curves to predict future prosthesis failure[23]

5.1 Introduction

Chapters 2 and 3 explained that the NJR was established to provide data on prosthesis survival, but that, to date, the registry only has published survival rates of up to 5 years post surgery which is insufficient to inform us about the long term survival of a prosthesis. As a consequence, chapter 4 provided a review of the published economic evaluation literature on hip prostheses, concluding that it is hampered by the lack of long term survival data on brands of hip prostheses. Although ODEP provides summarised evidence on the survival of prostheses used in England and Wales, this evidence is dependent on the manufacturers submitting data and also on the published literature. Thus there are some prostheses which are widely implanted but with no or little evidence of their success in terms of survival rate. The purpose of this chapter is to explore whether the techniques of survival analysis can be employed to forecast long-term survival of prostheses on the basis of only relatively short-run data.

Survival analysis, which is concerned with the time until the occurrence of an event, is widely used in medical research,[110] and in other academic disciplines for analogous purposes. Typically, at any point in time, many individuals in the dataset have not had the event of interest: this is known as right censoring of the data.[111] Standard non-parametric survival analysis methods are used to handle this, such as the Kaplan-Meier survivor function. Semi-parametric methods, such as the Cox proportional hazards model can be employed to allow survival to depend on patient characteristics[111], and for descriptive purposes these are entirely satisfactory. However, they are of limited use for extrapolation and predictive purposes because they leave the baseline hazard function unspecified.[110] Instead, parametric models which assume a baseline hazard can be fitted to the data and then used to extrapolate into the future. There are a range of alternative parametric distributions which vary in how precisely the baseline hazard is modelled.

[23] A version of this chapter has been submitted and reviewed by Medical Decision Making, it is currently under a 'review and re-submit' status.

5.2 Aims

The aim of this chapter is to assess the accuracy of survival analysis in projecting future revision rates beyond the sample estimation period. A previous well-known study[112, 113] is used, which is the only paper to have extrapolated prosthesis survival rates comparatively using registry data, theirpaper contributed to the NICE guidance on hip prostheses in 2000[114]. The original dataset used in that study has been extended to include eight years more data and thus enables assessment of the accuracy of predictions which would have been made in 2000, in the light of a longer time series, up to 2007. More specifically, extrapolations of the original estimated parametric curves are used to predict the outcomes over the subsequent years, and then compared with the actual outcomes. Since the predictions prove to be disappointing, I investigate three possible explanations: (i) the proportionality assumption is inappropriate for modeling differences between prostheses, (ii) the survival curves do not follow the Weibull distribution, (iii) the Weibull may be appropriate, but its parameters cannot be estimated accurately with only little early data.

5.3 Methods

The purpose of Briggs et al.[112, 113] was to construct a probabilistic Markov cost-effectiveness model for primary total hip replacement. An important requirement for this model is information on the transition probabilities between states of the model. Briggs et al. derive these by estimating survival functions for two illustrative prostheses, the Charnley (cemented) and Spectron (cementless). It is these estimated survival functions which are the focus of this chapter – examining how well they predict the subsequent survival of the prostheses for the period 2000-2007.

Briggs et al. model the risk of revision using the Weibull survival function, and in order to compare two alternative prostheses, they make a proportionality assumption with respect to the effect of prosthesis type. That is, they assume that the survival of each prosthesis could be described by a Weibull distribution with the same Gamma parameter (γ), but allowing the Lambda (λ) parameter to differ between the prostheses. To aid interpretation, recall that the

Weibull distribution is characterised by two parameters: λ (scale) and γ (shape), with a hazard function: $h(t) = \gamma\lambda t^{\gamma-1}$, which reduces to the exponential distribution $h(t) = \lambda$, if $\gamma = 1$.

Briggs et al. make a distinction between 'early' and 'late' failures: early failures are identified with reasons such as primary deep infection and dislocation, while late failures are due to failure of the prosthesis itself (general 'wear and tear'). They employ data on late failures from the SHAR to estimate survival: and use them alongside health outcomes QALYs and cost data to assess the relative cost-effectiveness of the two prostheses.

Recall from chapter 2, that SHAR contains information on prosthesis type, surgical procedure, patient characteristics, health outcomes (such as the EQ-5D), time to event data (prosthesis failure and revision surgery), and reasons for revision. Briggs et al. employ data from SHAR for the period 1992-1999, for all patients receiving either a Charnley or Spectron prosthesis. Both the Charnley and the Spectron prostheses are implanted in England and Wales, although the Spectron cementless cup has more recently been updated and renamed 'Reflection'. The NJR only reports ratings for the Spectron cemented stem and not the cementless stem. Survival rates from SHAR and NJR annual reports are not easily comparable as SHAR reports survival rates at 5 and 10 years (96.4% at 5 years and 92.5% at 10 years and 97.6% at 5 years and 93.3% at 10 years for the Charnley and the Spectron respectively), and the NJR reports 3 and 5 year survival rates for the Charnley stem only (98.7% at 3 years and 97.9% at 5 years). The NJR reports ODEP ratings for the Charnley cup and stem as 10A (10 year data, failure of 10% or less) and 7A (7 year data, failure rate of 7% or less) for the Reflection cementless cup and 10A for the Spectron cemented stem. Thus exact comparison of the rating/rates of the two prosthesis types for the two countries is not easily possible, although it does indicate rough similarity and so confirms the appropriateness of applying Swedish data to patients in England.

For the purpose of this thesis, I have been provided with - access to an additional eight years subsequent data was possible, extending the period to 2007. Appendix 7 provides a detailed description of the process of obtaining these updated data.

As a first step, I recovered those patients who would have appeared in the original Briggs et al. dataset, that is having had their primary surgery before 2000. I then constructed two forms of the dataset for this sample:

SHORT which shows their outcome as it would have appeared at the end of 1999

LONG which shoes their outcome as it appears at the end of 2007.

Thus, both forms refer to the same set of patients, but LONG differs from SHORT in that those prostheses which are shown as surviving at the end of 1999, are now revised – recording a date of revision if this subsequently occurred between 2000 and 2007, or date of patient death, or continued survival if the patient had still not required a revision by the end of 2007. This allows me to undertake non-parametric, and parametric analysis over both an 8 year period (for SHORT) and a 16 year period (for LONG).

For clarity below, I will also refer to:

ORIGINAL which is Briggs et al's original sample of patients. If I was able to perfectly match the data sent to me with the sample originally used by Briggs et al, SHORT would be identical to ORIGINAL.

My analysis is undertaken in five stages. **First**, I assess whether the matched 8 year dataset for 1992-1999 (SHORT) appropriately replicates the original Briggs et al. sample (ORIGINAL). I do this by comparing patient characteristics, revision rates, the non-parametric Kaplan-Meier and results of the Cox proportional hazards model.

Secondly, the same non-parametric methods are applied to the extended LONG dataset for the same patients for 1992-2007 (LONG). This provides a comparison between K-M and the Cox model computed on 8 and 16 years of observations, and allows a test of Briggs et al.'s original assumption of proportionality with respect to prosthesis type.

The **third** stage involves re-estimating Briggs et al's parametric models. Focusing on just the late failures Briggs et al. use parametric approaches to quantify the baseline risk and extrapolate beyond the observed eight years to a lifetime time period. They fit six alternative distributions (Weibull, Exponential, Gompertz, Log-logistic, Log-normal and Gamma) to the data, and compare the 'goodness of fit' using the AIC (Akaike's Information Criterion) and Cox-Snell residual plots. Their chosen alternative is the Weibull distribution. Therefore in this stage I fit the Weibull distribution to the extended LONG data and compare parameters (I also test the fit of the other distributions).

Since a key finding in the third stage is that the assumption of proportional hazards for prosthesis type is rejected, this is dropped in the **fourth** stage, and the Weibull model is re-estimated for each prosthesis separately for LONG and SHORT.

Finally, the **fifth** stage assesses the predictions of the original Briggs et al model, by comparing the estimated Weibull equations for the SHORT and LONG data. Ideally of course, the extrapolations (predictions) from Briggs et al.'s model should be assessed against the actual survival rates for 2000-2007, but in practice, this is infeasible given that predictions depend on the characteristics of individual patients, and will therefore differ across all patients in the sample[24]. Therefore I employ a more practicable alternative approach by taking an illustrative patient (a female aged 60 years old without a fracture) and compare her predicted survival curve, projected from SHORT with the within-estimation period predictions for the years 9-16 taken from LONG. This is conducted in two alternative ways. First proportionality is assumed, as do Briggs et al, then the proportionality assumption is relaxed, that is the original model specification is used, but estimated separately for the Charnley and Spectron. The first approach assesses the predictive performance of Briggs et al.'s original estimated equations, while the second assesses what their predictions would have been, had they not assumed proportionality.

5.4 Results

5.4.1 Matching data and replicating Briggs et al.'s results

98.3% of all the patients in ORIGINAL were successfully identified. The remaining 1.7% (n=350) is due to occasional minor coding discrepancies between the form of the data originally made available to Briggs et al. and the form in which I received it. In all such cases, caution is exercised, by omitting these patients (i.e SHORT has 350 fewer patients than ORIGINAL). Table 4.1 reports the sample size, patient characteristics and number of revisions in the ORIGINAL and matched (SHORT) datasets. The descriptive statistics show that the two samples are virtually identical in terms of patient characteristics, age and gender.

[24] Even in quite large samples, the number of patients displaying a specific set of characteristics will often be too small to compute meaningful survival curves.

	ORIGINAL*		SHORT	
	Charnley	Spectron	Charnley	Spectron
Patients	18,505	1,990	18,178	1,967
Mean age (sd)	72 (9.2)	74 (8.1)	71 (9.2)	74 (8.1)
Age distribution (%)				
<40 years	70 (0.4)	5 (0.3)	66 (0.4)	5 (0.3)
40-50 years	264 (1.4)	16 (0.8)	251 (1.4)	15 (0.8)
50-60 years	1,418 (7.7)	60 (3.0)	1,389 (7.7)	60 (3.0)
60-70 years	4,836 (26.1)	391 (19.7)	4,753 (26.1)	385 (19.6)
70-80 years	8,090 (43.7)	1,014 (51.0)	7,945 (43.7)	1,000 (51.0)
80-90 years	3,630 (19.6)	481 (24.2)	3,581 (19.6)	479 (24.3)
>90 years	197 (1.1)	23 (1.2)	193 (1.1)	23 (1.2)
Gender (%)				
Female	12,337 (66.7)	1,472 (74.0)	12,108 (66.7)	1,453 (73.9)
Male	6,168 (33.3)	518 (26.0)	6,070 (33.3)	514 (26.1)
Initial diagnosis (%)				
Osteoarthritis	12,970 (70.1)	1,348 (67.7)	12,826 (79.5)	1,329 (69.8)
Fracture	1,692 (9.1)	319 (16.0)	1,662 (10.3)	317 (16.6)
Other	3,843 (20.8)	323 (16.2)	1,628 (10.1)	258 (13.5)
Revisions (%)				
1992-1999	552 (2.98)	22 (1.10)	528 (2.90)	21 (1.07)
1992-2000	-	-	1,255 (6.90)	98 (4.98)

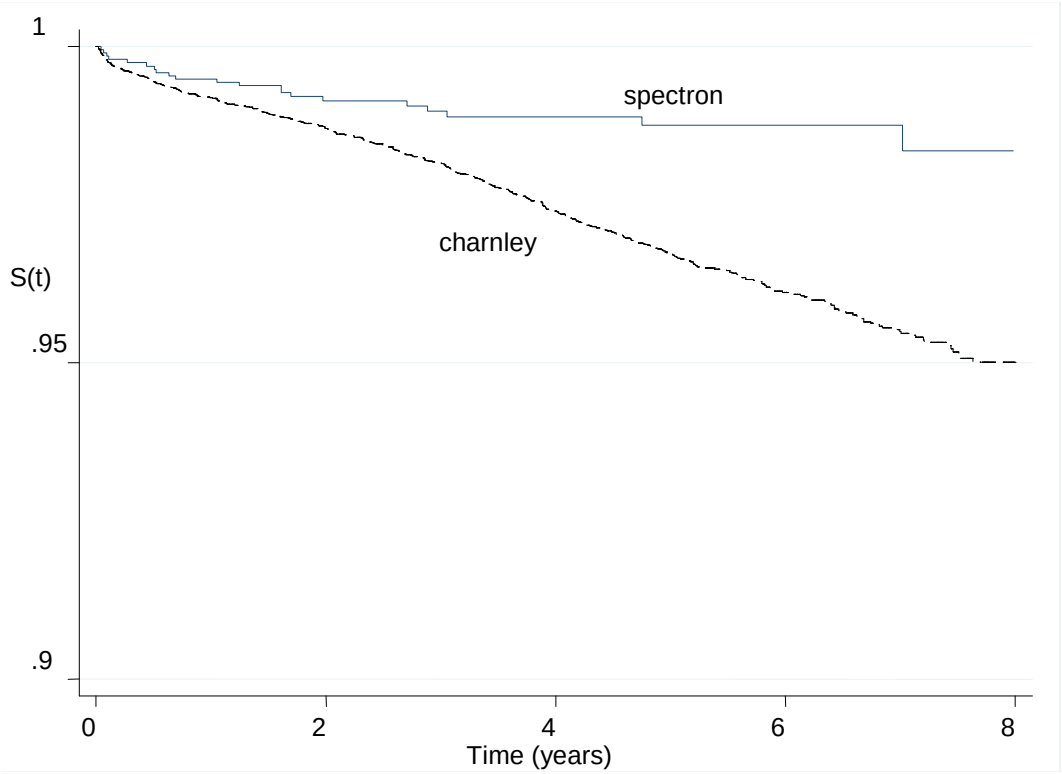
*Source: Briggs et al.[113] Table 2, p. 42

Table 5.1: Summary statistics 1992-9, Briggs et al.* compared to matched data

With comparability of the two datasets confirmed, the next stage is to attempt to replicate Briggs et al.'s non-parametric survival analysis. Figure 5.1 shows the Kaplan-Meier (KM) survival curves by prosthesis. These are identical to those reported in Briggs et al.[113] for the matched data. Crucially, as found by Briggs et al., there is a clear separation of the two survival curves. The significance of this separation is confirmed using a log-rank test for equality of survivor functions, which shows a highly significant difference ($p<0.001$), as is also found in Briggs et al.[113] (cf. Table 1, pg. 42). The results of the reduced form proportional hazards model (including age, gender, and fracture as covariates in addition to prosthesis type) for SHORT are also very similar, to those for ORIGINAL (the first two columns of Table 5.2). Note that the different hazard ratios for fracture merely reflect a changed coding used in the new database: it

appears that there has been a change in how fracture is classified, this results in fracture in SHORT having a smaller hazard ratio than in ORIGINAL)

Figure 5.1: Kaplan-Meier survival curves for matched sample, 1992-1999 (SHORT)



	ORIGINAL*		SHORT		LONG	
	Hazard Ratio	SE	Hazard Ratio	SE	Hazard Ratio	SE
Spectron	0.435†	0.095	0.437†	0.097	0.876	0.092
Age	0.974†	0.004	0.977†	0.004	0.968†	0.002
Male	1.785†	0.15	1.715†	0.147	1.683†	0.092
Fracture	1.718†	0.221	1.217	0.167	1.009	0.009

* Source: Briggs et al.[113] Table 3, p.43
† significant at the 5% level

Table 5.2: Cox proportional hazards model

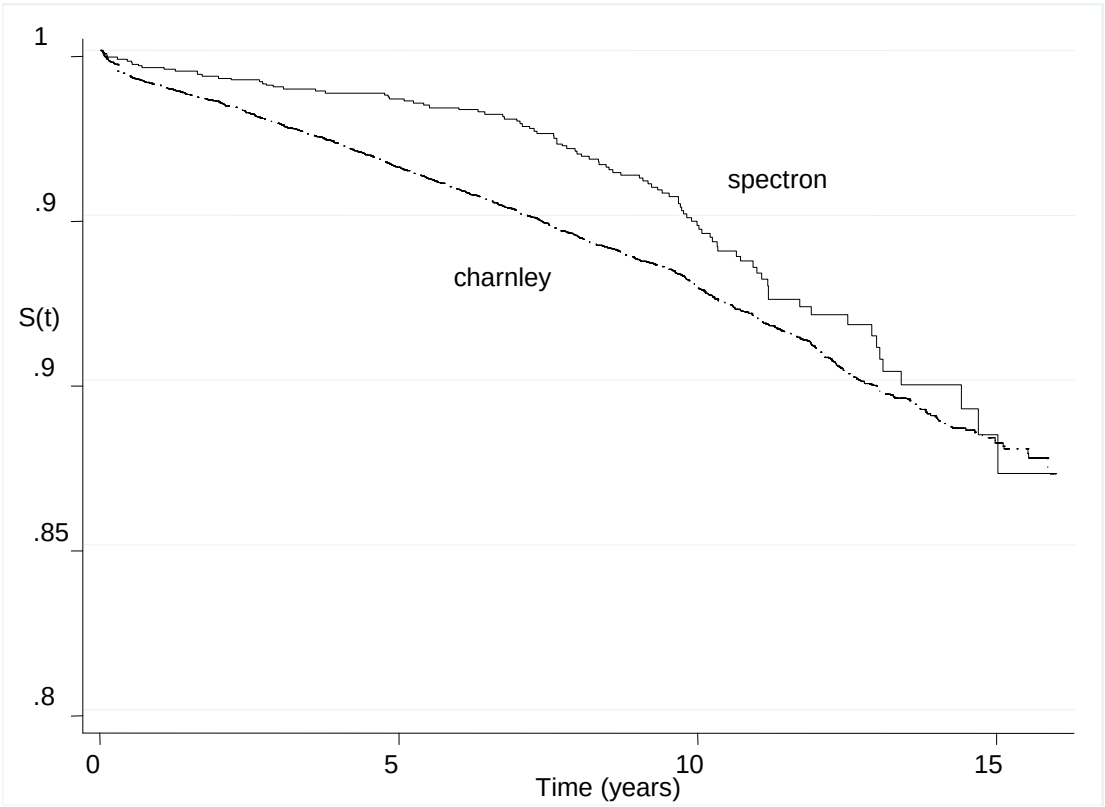
Thus, this matched sample (SHORT) generates substantively identical results to Briggs et al. in terms of the KM survivor function and Cox regression results and nothing appears to have been lost by excluding the 350 unmatched patients. Their headline conclusion is confirmed: the Spectron is unambiguously and significantly[25] superior to the Charnley. This justifies proceeding with the subsequent analysis, in which the success of extrapolations from SHORT are examined over the years, 2000-7.

5.4.2 Re-estimation using the extended dataset (LONG)

Figure 5.2 shows the KM curves taking account of the extra 8 years data for the years 2000-7 (thus the total time period is 1992-2007). The striking contrast with the KM curves in Figure 5.1 for the initial 8 year period is that the curves for the two prostheses now cross after approximately 15 years. In other words, contrary to expectations from Briggs et al., the Spectron survives less well than the Charnley after 15 years. Closer examination of Figure 5.2 reveals that the difference between prostheses widens over time up to eight years (as in Briggs et al.), but thereafter, the gap narrows fairly rapidly, until the crossing at 15 years. The log-rank test for equality of the two survivor functions still shows a significant difference between the two curves at the 2% level; this reflects the fact that the Spectron curve lies above that for the Charnley for most of the 16 years. However, this test no longer has any real meaning, given the crossing of the curves.

[25] In the Cox proportional hazards model all statements on statistical significance relate to the null hypothesis that the coefficient in question has a true value of 1. For example, in this case, the coefficient on Spectron is significantly different from 1, indicating its superior survival (relative to the Charnley).

Figure 5.2: Kaplan-Meier survival curves for matched sample, 1992-2007 (LONG)



The impact on the proportional hazards model of extending the dataset on the sample covariates is presented in the last column of Table 5.2. While the hazard ratios (for ORIGINAL and LONG) are almost identical for age and gender (differing for fracture due to the different coding system used in the new database), the hazard ratio for the Spectron is twice as large, but no longer significant, implying no significant difference between the two prostheses. Moreover the test of Proportional Hazards is now rejected at the 1% level for the Spectron.

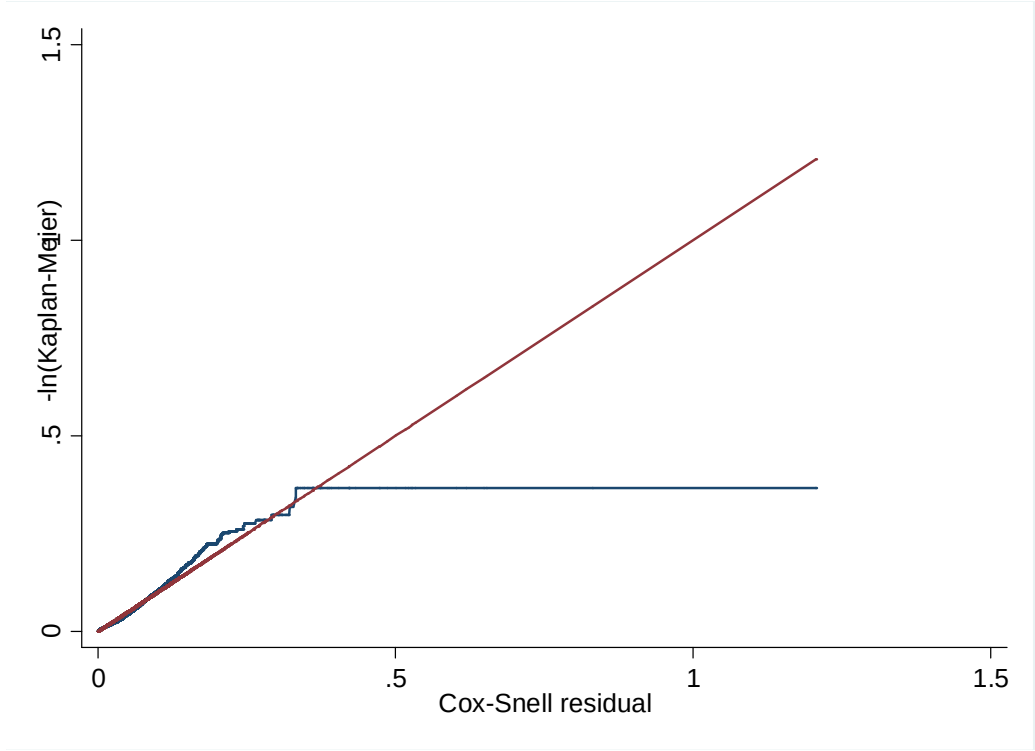
5.4.3 Parametric analysis of LONG

In the original paper by Briggs et al, six alternative parametric distributions were fitted and Cox-Snell residuals were calculated and plotted to consider the fit of the distributions to the data. They found that the fit of most of the models looked very similar, with the lognormal assumption for survival times providing a particularly poor fit to the data. They were also unable to fit the generalised Gamma to the data due to a lack of convergence, resulting in no standard errors for

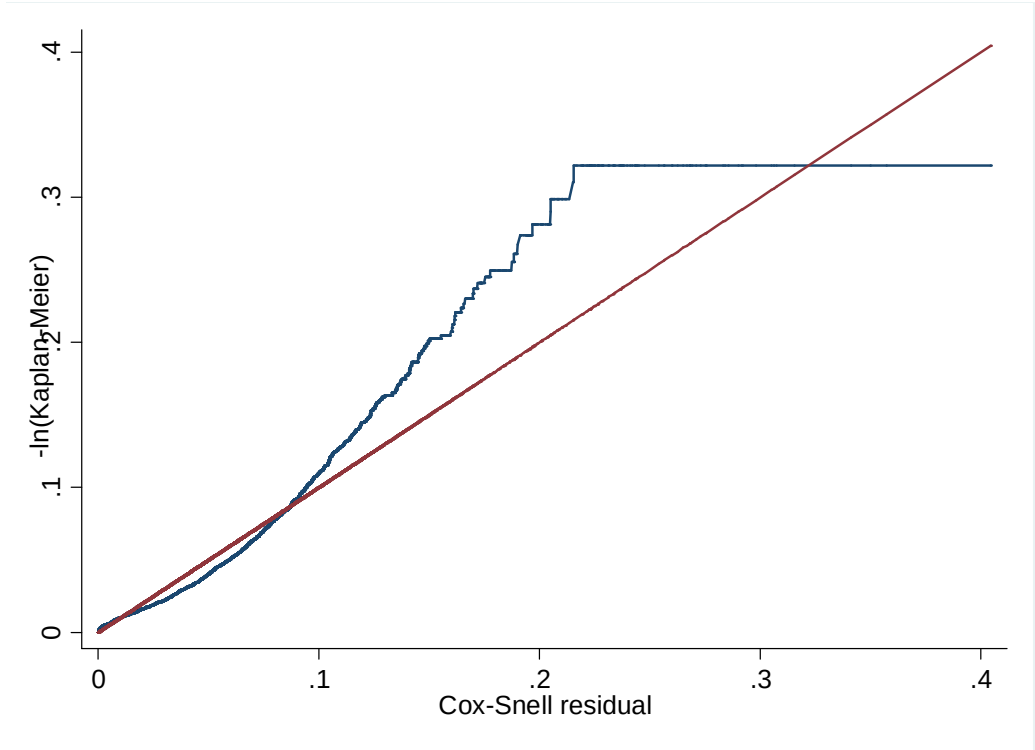
the estimated coefficients. The Additional (Aikake's) Information Criterion (AIC) was also calculated to test the fit of the distributions. Their Table 5, p.44[18], reports that, after the Gamma (which did not converge), the Weibull had the lowest AIC, although they point out that there is little to choose from between the models.

For the purpose of the current analysis, the same six parametric distributions are tested on the LONG data. Five are shown in figures 5.3a-e (the Gamma is not reported as it did not converge and thus could not be used). Figures 5.3 explore the relative fits of the five remaining alternatives by examining Cox-Snell residuals. Here, closer fits are indicated the nearer is the curve to the 45 degree line [111]. As can be seen, the best fits are provided by the Gompertz and Weibull - in that order, but with little difference between the two. Therefore, for comparability with Briggs et al. I proceed with the Weibull distribution. Table 5.3 reports the results of fitting a Weibull model to the extended dataset and compares with the results reported in Briggs et al. As can be seen, the estimates of γ are very similar: 1.402 (SE 0.021) for LONG and 1.454 (SE 0.069) in ORIGINAL. However, the coefficient of Spectron clearly differs between the two: 0.957 (and not significant) for LONG as opposed to 0.258 ($p<0.001$) in ORIGINAL. This is not unexpected given the non-parametric results from the KM curves above and Cox Proportional Hazards model. It is clearly inappropriate to make the proportional hazards assumption regarding prosthesis type as did Briggs et al.; in other words, it cannot be assumed that the Spectron survival function is simply an upward shift of the Charnley function.

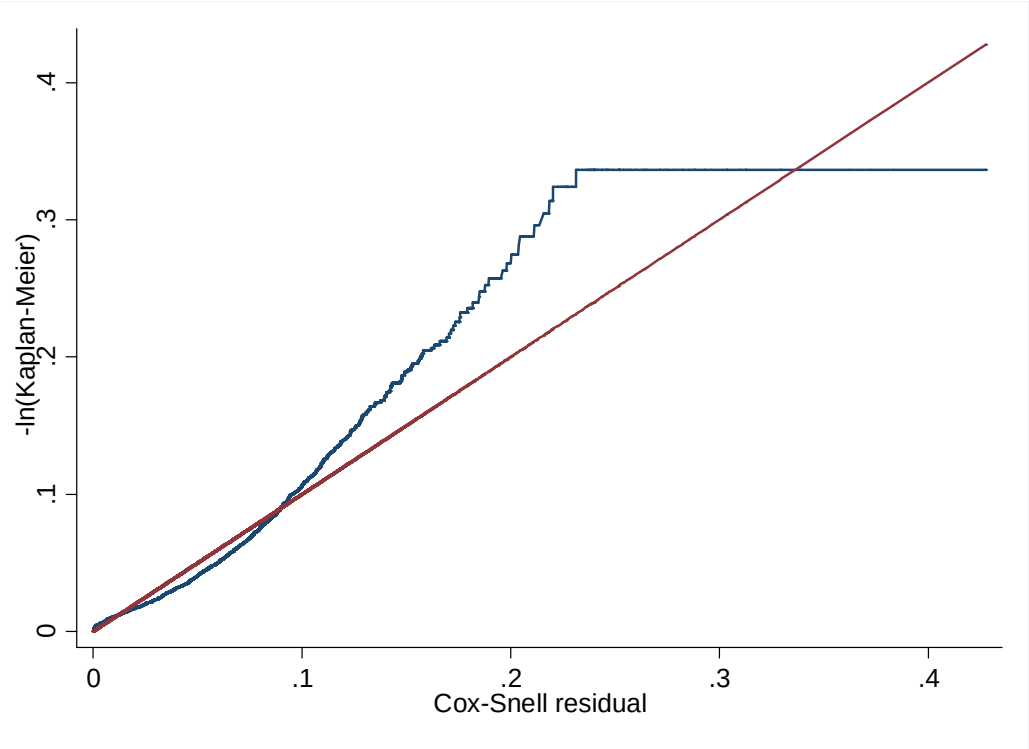
Figures 5.3: (a) Weibull



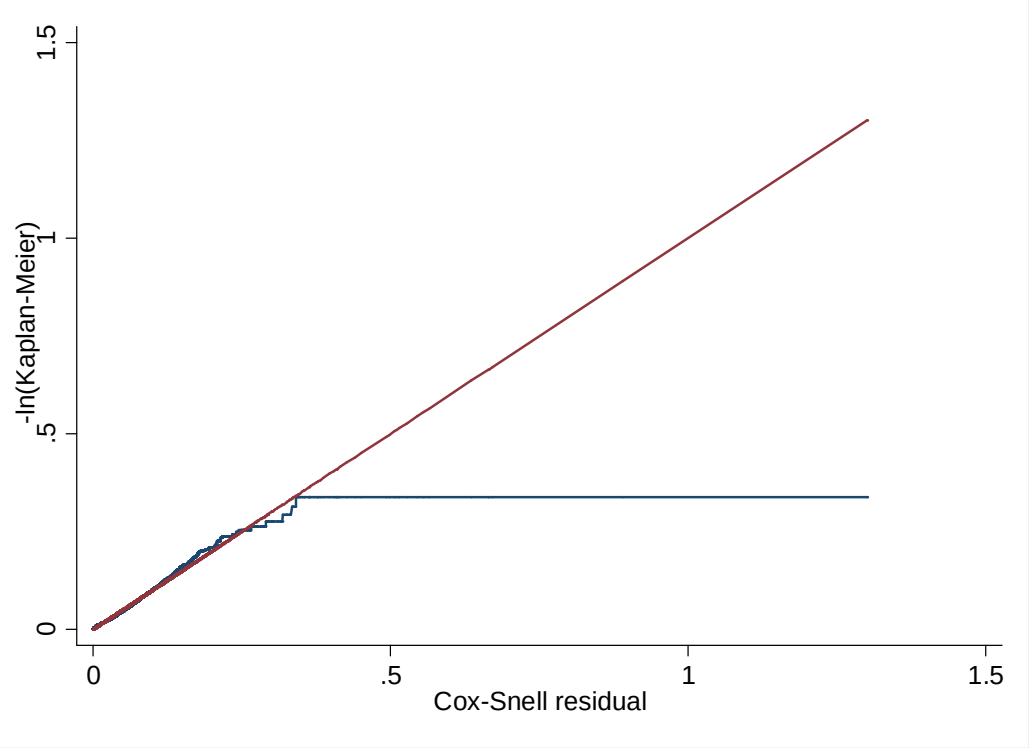
(b) Log-Normal



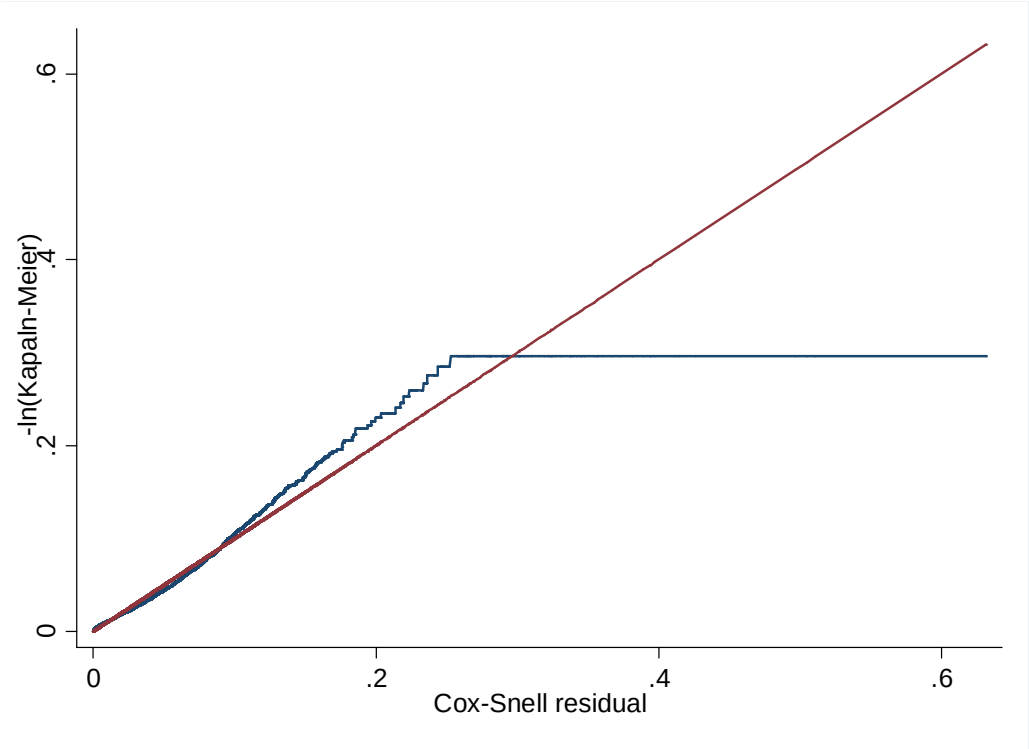
(c) Log-Logistic



(d) Gompertz



(e) Exponential



	ORIGINAL*		LONG	
	hazard ratio (SE)	p-value	hazard ratio (SE)	p-value
Spectron	0.258 (0.099)	<0.001	0.957 (0.119)	0.73
Male	2.177 (0.238)	<0.001	1.806 (0.119)	<0.001
Age	0.963 (0.005)	<0.001	0.959 (0.003)	<0.001
Fracture	1.303 (0.251)	0.17	0.840 (0.099)	0.14
Gamma	1.454 (0.069)		1.402 (0.021)	

* Source: Briggs et al.[113] Table 7, p. 45

Table 5.3: Weibull survival functions, proportionality assumed

5.4.4 Dropping the proportionality assumption

Table 5.4 reports the results of now re-estimating the Weibull model separately for both the Charnley and Spectron using the LONG dataset. The results confirm that there is a noticeable difference in the γ parameter estimates: 1.356 for the Charnley and 2.521 for the Spectron. In

other words, while the γ for the Charnley is little changed from the γ estimate in Table 5.3, it is noticeably increased for the Spectron.

	Charnley		Spectron	
	hazard ratio (SE)	p-value	hazard ratio (SE)	p-value
Age	0.962 (0.00)	<0.001	0.911 (0.01)	<0.001
Male	1.882 (0.12)	<0.001	0.981 (0.25)	0.94
Fracture	0.879 (0.10)	0.29	0.388 (0.17)	0.03
Gamma	1.356 (0.03)		2.521 (0.25)	

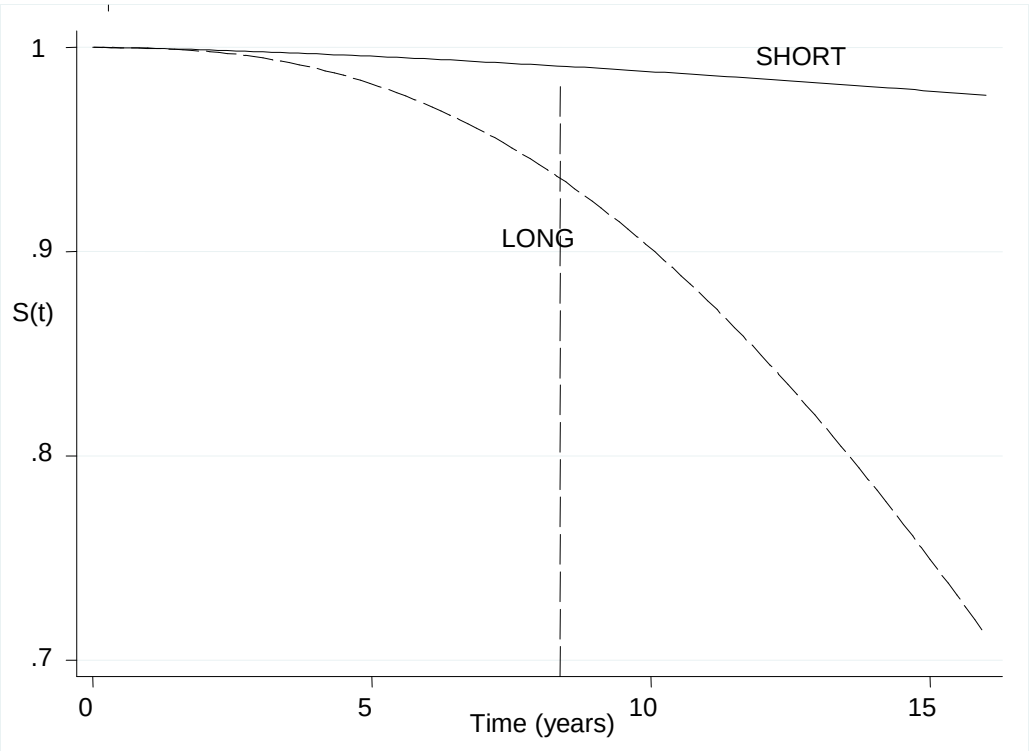
Table 5.4: Weibull survival function: LONG, without assuming proportionality

5.4.5 Assessing predictions

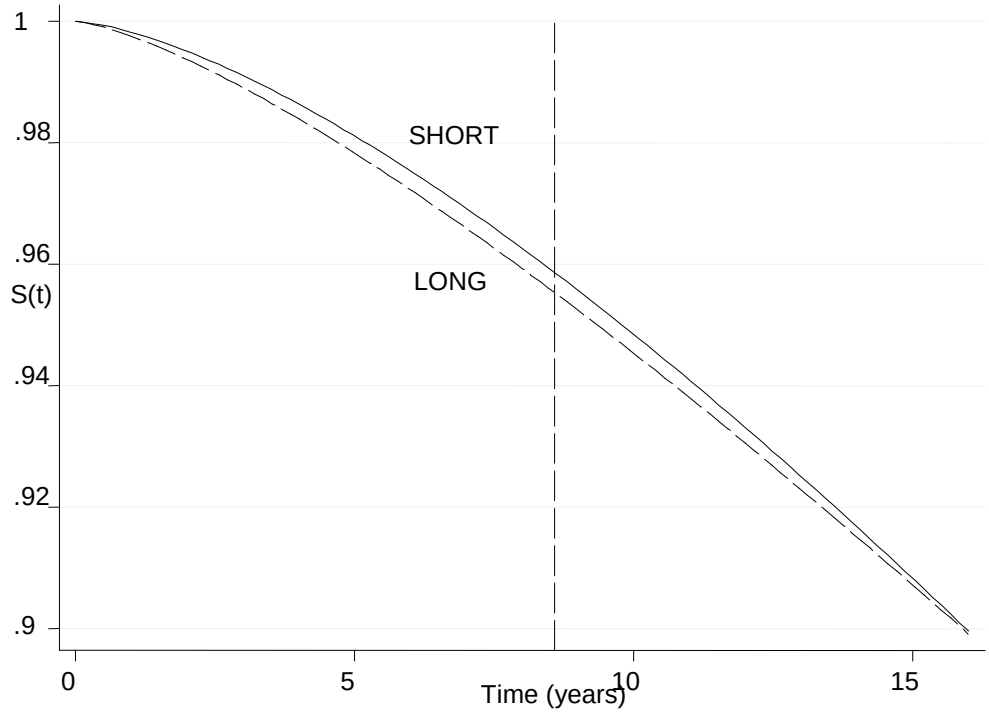
The results call into question the robustness of Briggs et al.'s model for predicting the future survival of the two prostheses over the subsequent years, 2000-2007. In order to test this, prostheses specific survival rates are estimated for an illustrative patient (female, age 60, non-fracture) using the most appropriate specification, that is with the extended dataset (LONG) and not assuming proportionality (as presented in Table 5.4). These within-estimation period predictions can be interpreted as a proxy for the actual values or, less speculatively, simply as 'better' predictions because they are based on 16 years' data and without the inappropriate proportionality assumption. These predictions are then compared to the predictions using the initial dataset as available to Briggs et al., first assuming (as they did) proportionality and then relaxing this assumption.

Figure 5.4 shows the results of the first attempt (assuming proportionality) at assessing the predictive ability of Briggs et al. For LONG, these are all within-period predictions estimated from the equations presented in Table 5.4. For SHORT, years 1-8 are within-period (Table 5.3, first column), while 9-16 are beyond period, from the equations in Table 5.3 (final column). As can be seen, the Weibull for the Charnley predicts future survival very accurately, the difference is never more than 0.33%. For the Spectron however, a difference of nearly 5% has already emerged by year 8, and this then increases dramatically as we move into the extrapolation period. By year 16, the original Briggs et al. model predicts survival at nearly 98%, compared to only 71% from the revised LONG Weibull without proportionality. An error of more than 26% percentage points could lead to considerable errors in any cost-effectiveness evaluation.

Figure 5.4: Estimated Weibull survival curves: SHORT assuming proportionality (as in Briggs et al.) compared to LONG without proportionality*
(a) Spectron



(b) Charnley



*For late failure, female, aged 60, non-fracture.
Fitted values are within-period estimates for LONG, within-period for years 1-8 but beyond-period extrapolations for SHORT for years 9-16

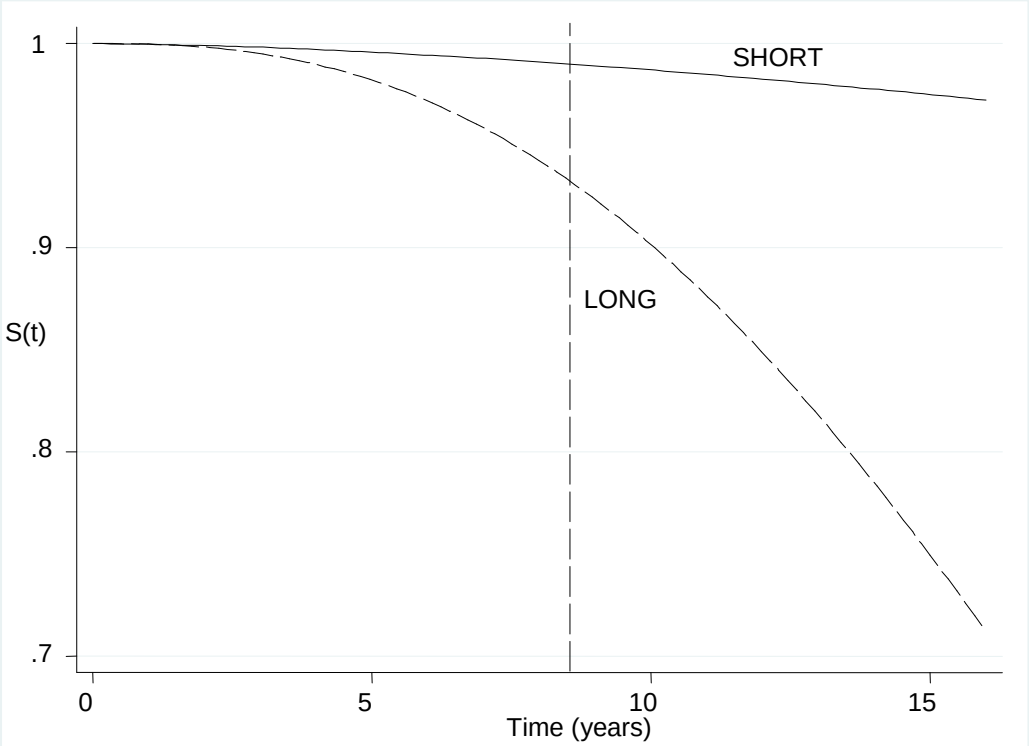
	Charnley		**Spectron**	
	hazard ratio (SE)	p-value	hazard ratio (SE)	p-value
Age	0.970 (0.00)	<0.001	0.966 (0.05)	0.51
Male	2.101 (0.23)	<0.001	2.192 (1.81)	0.34
Fracture	1.072 (0.20)	0.71	1.339 (1.50)	0.79
Gamma	1.471 (0.07)	-	1.629 (0.55)	-

Table 5.5: Weibull survival functions: SHORT re-estimated without proportionality

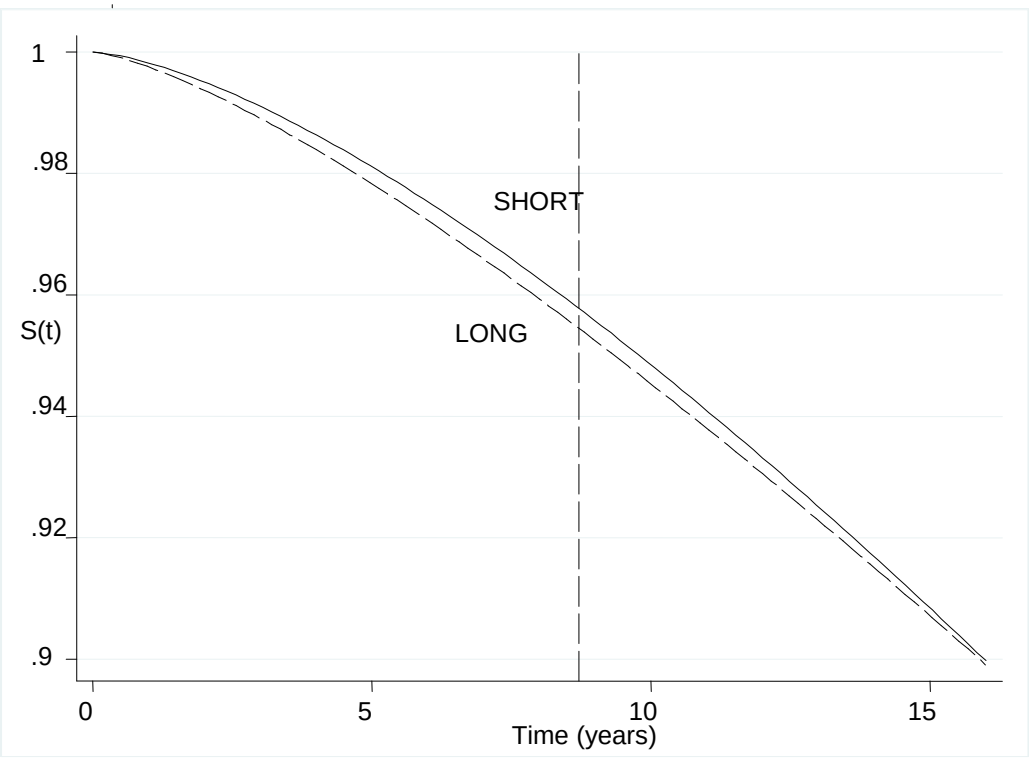
To apply the second approach, I first re-estimate the Weibull using the SHORT data but now without making the proportionality assumption. These results are shown in Table 5.5 and Figure 5.5 shows the predictions. These should be interpreted as the predictions Briggs et al. would have derived had they not assumed proportionality. In fact, the predictions are almost unchanged from Figure 5.4; the Weibull continues to predict without noticeable errors for the Charnley, but over-predicts substantially for the Spectron. For instance the 'error' by year 16 is 25%, only slightly lower than that presented in Figure 5.4.

Figure 5.5: Estimated Weibull survival curves: SHORT without proportionality compared to LONG without proportionality*

(a) Spectron



(b) Charnley



*For late failure female, aged 60, non-fracture.
Fitted values are within-period estimates for LONG, within-period for years 1-8 but beyond-period extrapolations for SHORT for years 9-16

5.5 Discussion

The analysis identifies that, with the addition of eight years more data, it is no longer the case that the survival of the Spectron prosthesis is superior to that of the Charnley prosthesis. After about 15 years, the survival rate of the Charnley is found to be similar, if not, better than the Spectron. This finding is supported by a recently published SR annual report,[115] which reports 10 year survival rates of 92.7% for the Charnley and 92.0% for the Spectron, notably the confidence interval around this survival rate is larger for the Spectron (1.5%) than the Charnley (0.4%).

Of course, whether these results are generalisable to other prostheses still remains to be seen and should be the subject of further work. However, the predictions for the Charnley prostheses do appear to have been robust in comparison to those for the Spectron. This suggests the cautious

conclusion that newer prostheses, where only a few have been implanted, or where they involve using a new technology, are those which maybe less robust to this methodology.

In the original work by Briggs et al, a full cost-effectiveness analysis was carried out comparing the Charnley and Spectron. The results were based on mean costs and QALYs and indicated that the Spectron is cost effective in younger patients, with the probability of the Spectron being more cost-effective than the Charnley ranging from 70% to 100%. It is likely that using the new estimates (LONG), the cost effectiveness of the Charnley will be enhanced. The new results suggest that the Charnley is more likely to be generally preferred. In particular, the deterioration of the Spectron from 15 years on will be most relevant for patients with longer own life expectancy (typically the young). For older patients this may be less relevant since they are less likely to outlive their prostheses, but even the original Briggs results already show the Charnley to be more cost-effective than the Spectron for older patients. In other words these new results mainly reinforce the superiority of the Charnley for the old, but call into question the superiority of the Spectron for younger patients.

Relating this to the methods used by Briggs et al, the assumption of proportionality by prosthesis for the Charnley and Spectron is violated, and their original estimate of the key parameter, γ, is substantially revised upwards for the Spectron. Moreover, extrapolations of the estimated Weibull curve (as given in Briggs et al.) yield very poor predictions for the survival of the Spectron in the 8 years after the original estimation period; the magnitudes of the 'error' is about 25% by year 16.

In section 5.2, I raised three possible sources of error and I now re-visit them. First, as just mentioned, it is now clear that Briggs et al.'s assumption of a proportional impact of the prosthesis type is inappropriate – this is revealed very clearly by the crossing of the KM curves in Figure 5.2. This finding implies that such a specific assumption should probably not be employed in future comparative studies of this type, although in itself, it does not invalidate the use of survival analysis.

Secondly, turning to the performance of the Weibull distribution, there is no evidence that any of the alternative distributions suggested by Briggs et al. would have performed any better. Indeed,

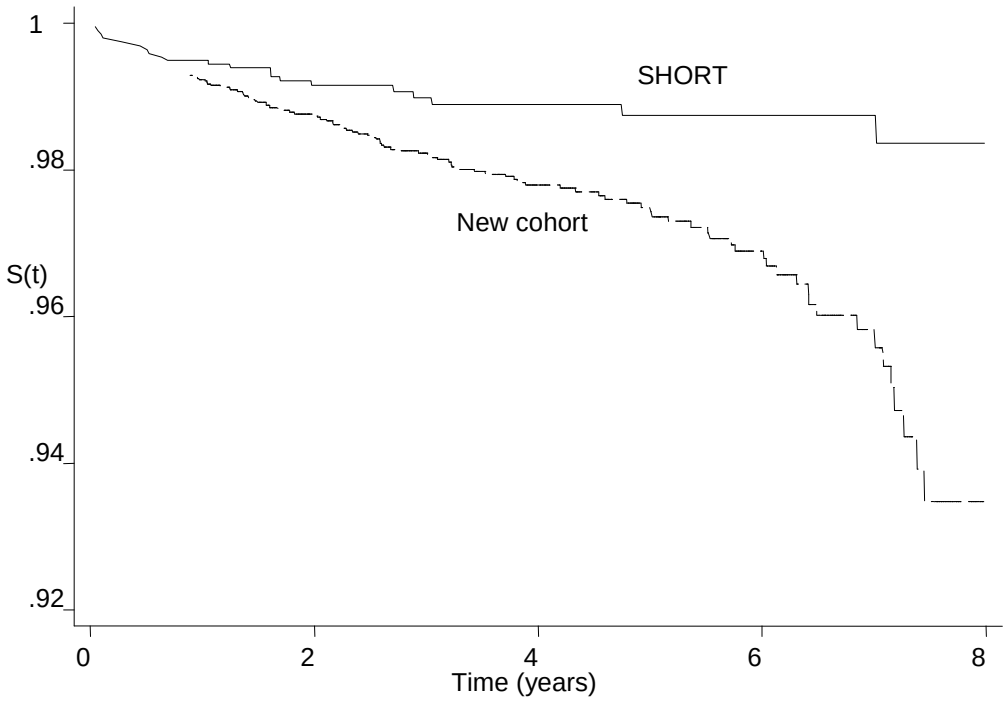
for the Charnley prosthesis, the originally estimated Weibull 'predicts' remarkably accurately (Figures 5.4 and 5.5); the problem is confined to the Spectron.

Third, I raised the possibility that, while the Weibull distribution may be appropriate, its estimated parameters are unstable – in the sense that their magnitudes are sensitive to the addition of more data. For the Charnley, this is not the case: without assuming proportionality, the estimated γ and other coefficients appear to be fairly robust between the eight and sixteen year data: $\gamma = 1.471$ (SHORT) and $\gamma = 1.356$ (LONG) (Tables 5.5 and 5.4 respectively). However, for the Spectron, the γ estimate rises substantially from 1.629 for SHORT to 2.521 for LONG. This appears to be the crux of the matter – such a substantial shift in the γ parameter would inevitably lead to large revisions in predicted values. This also explains why the assumption of proportionality might be acceptable with only 8 years data, as in Briggs et al., but then rejected in the light of 16 years of data (the γ for Charnley and Spectron are similar using the SHORT dataset, but very different when using LONG.)

Finally, in this chapter I have used this specific case comparison, between the Charnley and Spectron prostheses, to illustrate a general point: however, if the interest were also in this specific case *per se*, then the results do raise some doubts about the general trend to implant cementless prostheses (shown earlier in chapter 3) [116, 117]. It should be remembered that the Spectron was a relatively new prosthesis at the time of the Briggs et al. study, so it is possible that the prosthesis itself was experiencing some unexpected 'teething troubles' during this period, and that high failures rates in post-1999 reflect some unforeseen deterioration in the prosthesis itself. In this respect, some very indirect evidence is provided by returning to the Swedish registry but now examining the survival of Spectron implants made between 2000 and 2007, i.e. *after* the Briggs et al. estimation period and therefore not considered in the above analysis. Figure 5.6 compares the Kaplan-Meier curves for the Spectron for this set of patients compared to the Briggs et al. cohort. Quite clearly, survival in this later cohort is much poorer. The causes can only be speculated, however, these results are highly suggestive of confounding by indication – as the use of Spectron has increased, so the outcomes appear to have deteriorated. There may also be some confounding by intensity i.e. a wider range of less specialized, lower volume providers such as community hospitals) This would suggest that relatively new prostheses (and more generally any new intervention where the event of interest occurs at a point

far into the future) should be handled more cautiously than more established prostheses (and interventions) for which both clinical learning and effectiveness may be more established.

Figure 5.6: Kaplan-Meier survival curves for the Spectron: Briggs et al.'s cohort (SHORT, 1992-1999), compared to a new cohort (2000-2007)



5.6 Conclusions

In summary, two main conclusions can be drawn from the analysis. First, when estimated using a large sample of patients (Charnley, with 18,000+ observations), the Weibull survival curve appears to provide accurate predictions of future survival rates when using the first 8 years after implant. However, when estimating survival using a smaller number of patients (about 2,000 Spectron prostheses were implanted) predictions are unstable and we argue should not be extrapolated. Secondly, when comparing alternative prostheses, the assumption that prosthesis type has a simple proportional impact on survival is extremely speculative and potentially dangerous. Thus it can be concluded that it is more robust to estimate separate survival curves for each prosthesis.

Generalisation of these results should be qualified by emphasizing that this analysis focuses only on data from one country, Sweden, and only on two brands of prostheses, with a relatively small dataset for one of them. Further research is called for. Unfortunately, there is very little other evidence in the existing literature on extrapolation of prosthesis survival rates[74, 75]. Fitzpatrick et al. use 14 year observed survival data on two prostheses to extrapolate to a 60 year time horizon, but provide no tests of the extrapolations[118]. An alternative approach is to assume a range of alternative values for the expected life of a prosthesis, estimating for each value the increase in the expected life of a cementless prosthesis required for the two types of prostheses to achieve the same expected net present value cost.[119] Most studies, however, merely use the observed or existing survival data from published sources, and thus do not extrapolate to a lifetime horizon;[120, 121] but this approach can only be used when long term follow-up data already exists.

It is obvious that future research will require more long-term follow-up data, and the introduction of the NJR in 2003/4 will provide an invaluable source for data to conduct future economic evaluations of alternative prostheses. However, as I have shown, predicting prosthesis survival based on data from a registry of five/six years is unlikely to be sufficient to provide robust extrapolations. This was also the implication of the comparison of 3 year and 5 year revision rates in chapter 3 (section 3.3.1). Further research should identify and consider more reliable ways to extrapolate revision rates into the future in the absence of longer follow-up data. The adoption of a Bayesian evidence synthesis approach which employs registry data on a range of prostheses from different countries and combines this with other routinely collected data is one option. This should more fully exploit the data by considering linkage of registry data-sets together, such as the NJR, SHAR and perhaps the New Zealand Registry. This would help to overcome the problem that because there are so many alternative prosthesis options, even in an extensive registry such as the NJR, the actual numbers using a particular prosthesis may still be quite small.

In terms of what this implies for how decisions between alternative prostheses should be made in the mean-time i.e. until more and longer data become available, these results do not offer any direct answers. Here the purpose has been confined to assessing whether the methodology used by Briggs et al is a viable option for dealing with the short-term data available from the NJR to predict future revision rates for economic evaluations of alternative prostheses. The answer to

this question is that it does not appear to be a robust method for newer prostheses where there is short-follow up with a small number of patients who had it implanted. Of course, for decision makers or NICE guidance this finding is unhelpful. The implication is that current decisions will still have to be beased on whatever existing evidence is available from joint registries and the literature, based on the NICE benchmark of 10% at 10 years where possible, or 3 years revision rate experience if their performance is consistent with the benchmark of a 10% revision rate at 10 years. Future work should also further determine the value of this methodology to other prostheses and settings and establish models for incorporating multiple international evidence sources.The results of this and the preceding chapters suggest that there are considerable doubts as to whether current methods can satisfactorily be used to answer 'what prosthesis should be implanted'. This continues to be unresolved, given the problems of limited long term survival data and unreliable methods of extrapolating. As a consequence, the remainder of the thesis turns the focus of the thesis from what '*should be'* to '*what is'* implanted using the NJR data as a primary data source.

Chapter 6, Actual choices in the NHS: the potential for buyer and seller power?

6.1 Introduction

The previous two chapters focused on choice *prescriptively*, in the sense of exploring analytical techniques designed to establish which prosthesis should be implanted. In this and the next chapter, the emphasis is changed to examining the choices which are *actually* currently being made within the NHS in England and Wales. Unfortunately, I am unable to test the effects on price as there is no published or national data on costs of prostheses (other than the NHS reference costs, which are at an aggregate level for THR - see chapter 3). However, I can investigate the pattern of purchases, in particular, exploring a previously overlooked issue, which is how the structure of the main buyer (the NHS) interacts with the structure and nature of competition within the supplying industry (the manufacturers of THR) to influence choice at the hospital level. In order to do this, I will draw on the analysis and methodologies of Industrial Organisation, which is the branch of Economics which underpins most of Competition Economics and Policy. The purpose of this chapter is to introduce, and in some instances measure the theoretical concepts used in Industrial Organisation, and this provides the background for the econometric estimation of models designed to test a number of hypotheses in the next chapter.

This chapter is organized into six sections. Section 6.2 discusses what is meant by patient choice, and some of the relevant recent policy literature; it concludes that choice in this context is usually not by the individual patient, but by their 'agent' (the surgeon, hospital or NHS). This raises the possibility that patients may benefit from the considerable buyer power which the NHS has in the market (as in other health care markets.) This is true, at least in principle, because the NHS is overwhelmingly the main buyer. Section 6.3 discusses whether the NHS is actually able to exploit this potential. This takes the discussion into the area of public procurement – the ability of the NHS to exploit its power depends on how it organizes its purchasing policy. However, the buyer must obviously transact with the sellers, in this case the manufacturers of THR prostheses. Industrial Organisation Theory provides a framework in which to explore how a (potentially) powerful buyer may interact with potential powerful sellers. This is discussed in

section 6.4. Section 6.5 then turns to the supplying industry and establishes a statistical description of the nature of its market structure – in the terminology of IO, a near duopoly market dominated by two large Multinational Firms, Depuy and Stryker. Section 6.6 considers the implications for relative market power, and this forms the basis for the econometric estimation of the next chapter. Section 6.7 concludes.

6.2 Patient Choice under the NHS

6.2.1 Brief history of recent policy on patient choice

In the present context, the main concern is with the choice (by patients) as to which prosthesis to have implanted and in which hospital. 'Patient Choice' has been increasingly an issue in the policy and political debate, but, as will be seen, currently is really only concerned with choice of hospital.

Patient Choice was first emphasised by the Conservative government as part of their 'internal market' reforms in 1989. They aimed to provide patients with better health care and a wider choice of services regardless of where they lived within the UK. However, subsequent research [122] found that, contrary to this intention, in practice, the choice of in which hospital the patient is treated lay largely with the GP who would make the choices on behalf of the patient.

Patient choice and the introduction of competition to the health care sector was not initially a focus for the Labour government which came into power in 1997. Their main focus was still more on choice over the date and time of hospital appointments rather than which hospital. In 2002, it published a 'progress report' called' ***Delivering the NHS Plan***, this was their first step in offering patients on waiting lists, the opportunity to choose an 'alternative' hospital provider with a shorter waiting time. However, patient choice was expanded in January 2006, when patients were offered a choice of four or five providers at the point of GP referral; by May of the same year, this had further expanded to include a choice of providers across England including Foundation Trusts (FTs) and Independent Sector (IS) Treatment Centres and, in August 2006, IS hospitals as well. Patients now made their choice through the 'choose and book' facility and choice was firmly established in the **NHS Constitution** (2008/9) so that all choices should be made from a national position. The Kings Fund has carried out some work to identify whether in

practice this choice is taking place at the patient level, they cite that 44% of patients referred for treatment in May 2007 could recall being offered a choice by their GP, an increase from 30% in May 2006. However, they point out that this still compares poorly to the DoH's 80% target, set in April 2007 [122].

Extending policy on patient choice beyond just choosing their provider of treatment, i.e. also to which GP, choice of treatment or pharmaceutical, is still undeveloped [122]. However, the recent annual report by NHS choices [123] states a continued commitment to patient choice by the Coalition government. This is set out clearly in the current consultation document: *Liberating the NHS: Greater choice and control. A consultation on proposals[124]*

In the present context, the choice of hip prosthesis to be implanted is mediated as follows: a GP refers the patient to a specific surgeon when a need for THR is identified. At the point of referral, the patient is given a choice of hospital and thus some choice in surgeon. When the patient goes for a consultation with the surgeon, the choice of which prosthesis to implant is made.

6.2.2 Theoretical perspective

There is a very limited literature on patient choice between different interventions, and this is not surprising as patients are very often poorly informed about the alternatives. This is a classic example of what is known in economic theory as **asymmetric information** "*a situation in which one side of an economic relationship has better information than the other*" p.17,[125] . Thus, we need to be able to define 'who' is making the choice of prosthesis to implant, bearing in mind that the patient has no information from previous experience (as this is the first time they are undergoing this type of surgery) and they have not usually sought information, which if they had, would not be easily accessible or comprehensible to the lay person. This is the **principal-agent problem**, which refers to an economic relationship in which one party, the principal, hires a second party, the agent, to perform some task on the party's behalf [125]. There is vast Economic literature on the Principal-Agent problem[126]. Within the context of this chapter, there are two dimensions to the problem: firstly the patient is the principal and the surgeon or the hospital are the agent, making the choice of which prosthesis to implant on behalf of the patient[127, 128]. Secondly, the surgeon is also effectively the agent for the funder, the NHS, choosing what prosthesis to purchase/implant.

Thus, while current policy has focused on reducing the asymmetry of information between patient and GP (where the GP is acting as the patient's agent and also of the third party payer) by providing patients with greater information on the performance of hospitals[129], the success of such measures is still not clear. Either way, this has no obvious implications for the choice between services/pharmaceuticals/medical devices. Here, the principal-agent problem is still very apparent. Surgeons hold technical medical knowledge which the patient lacks. This means that the patient demands two different services from their GP or surgeon, they require both information and then the intervention itself which relies on the surgeon, acting as the agent, recommending what is in their best interest.

6.3 Procurement in the NHS

From now on, the assumption made is that the decision about which prosthesis to implant is taken largely or exclusively by the surgeon/hospital, on behalf of the patient. While it would be expected that this decision will be based on the patient's characteristics, in terms of their need and ability to benefit from a specific prosthesis[130], it is also likely that surgeons have their own preferences for a specific brand of prosthesis due to their experience and training together with their own clinical results, and that they may, or may not, bear in mind different budgetary constraints.

Turning then to the second dimension of the principal-agent problem – the surgeon/hospital as principal of the NHS, more information is needed on the choices made by the surgeon or hospital, and how this relates to the purchasing decisions of hospitals. This moves into the area of Public Procurement - the purchase of goods and services by the public sector. This is of key policy interest, since public procurement accounts for a large proportion of public sector expenditure and demand in the UK economy – ranging from 11 to 18% of GDP[6]. Public procurement is increasingly attracting the interest of competition authorities[131]. This requires further theoretical examination and also in depth research on the procurement processes within the NHS.

Broadly speaking, the demand for health care services in the UK is dominated by a single buyer, the NHS. It follows that any analysis of a particular market in this sector should start with the recognition that this is, in principle, a monopsony market. The NHS is a large national

institution, which accounts in this case for the large majority of purchases of hip replacement prostheses. Economic theory informs that, if the NHS was to act as a profit maximizing monopsonist, with considerable buying power, it should be able to force down price. Where there is a large number of suppliers, a monopsony buyer should be able to negotiate with its suppliers to pay no more than the perfectly competitive price. Where there are relatively few suppliers, then the situation becomes oligopolists selling to a monopsonist, and theory provides no clear-cut predictions – it depends on the relative bargaining strengths of the two sides, although the threat to switch to alternative suppliers may still be credible. In general, more bidders (suppliers) in a procurement setting should equal more competition, although the design of the public procurement process can also affect the likelihood of firm collusion. Where the number of bidders increases, there is less opportunity for collusion (where firms "coordinate on their jointly preferred equilibrium")[37], which is also the case if the there is transparency in bidding process[131].

Amongst other things, the outcome will depend on their 'outside options' – can each side threaten not to transact at all because they can easily find alternative buyers/sellers? On the demand side at least, there are no other feasible large buyers – if a manufacturer wants to sell at all in England and Wales, it must sell to the NHS in order to achieve sizeable sales[26]. In other words, the NHS is potentially able to exercise considerable buying power if it acts as a single entity in its procurement policies.

When applying these theories in the present context, it should be recognised that public sector buyer power may be quite different to that of the private sector - for legal and regulatory reasons. It may differ significantly from the private sector because the public sector may be more risk averse about new patterns of purchasing or suppliers. As with other comparisons between private and public sector, public sector procurement decisions may not driven by a desire to maximize profits, although the public sector should be just as intent on minimising costs.

Turning from the theory to the practice, it must be recognised that, for many purposes, the NHS is not a single monolithic entity - it is dispersed according to location and specialty, and

[26] As will be seen below, in this market (similar to pharmaceuticals), most of the suppliers in this market are diversified multinational firms, for which the UK THR market accounts for only a small share of their worldwide sales. However, withdrawing from the UK, will have considerable ramifications in other markets. So in principle, although they could threaten to withdraw from the UK market, this seems unlikely except in extreme circumstances.

individual hospitals having different degrees of market power. Purchasing in the NHS (termed Commissioning) has undergone considerable reform in the past two decades.

As part of the ‘new Labour’ NHS package on public procurement, the 2005 'Framework for Health Reform in England' established the following major reforms with implications for NHS purchasing:

- Demand side reforms giving PCTs the responsibility for commissioning,
- Supply side reforms, establishing NHS Foundation Trusts (FTs) (discussed in chapter 6) as autonomous entities, not performance managed by the SHAs, and
- Transaction reforms, moving from negotiated contracts with hospitals i.e. block and cost and volume contracts to the system of PbR where hospitals are paid on a 'per case basis' and the prices are fixed nationally.

A further key recommendation for how the NHS purchases was made by the The Office of Fair Trading (OFT) in 2007, regarding the purchasing of pharmaceuticals[132]. They recommended that the government reform the Pharmaceutical Price Regulation Scheme (PPRS) of profit and price controls and replace it with a value-based approach to pricing. The OFT report suggested that this new approach would both benefit the patient and encourage innovation in the drug industry.

The change in government from Labour to the ‘Coalition’ (Conservative and Liberal Democrat parties) in the summer of 2010 has led to proposals for further reforms of the NHS with particular reference to NHS commissioning. The Coalition health reforms entitled: ‘Liberating the NHS’ (July 2010)[133] sets out to move budgets from PCTs to new ‘GP consortias’ who will be statutory bodies, supported by and accountable to the new NHS Commissioning board. These reforms are still yet to be confirmed at the time of writing, and they have undergone considerable criticism including a debate by the British Medical Association (BMA) on whether to have a vote of no confidence in the proposed bill. Most recently, these criticisms have led to reported likely concessions by the current health secretary[134].

Alongside the proposed reforms of commissioning, two high profile reports on procurement in the public sector have been published since the establishment of the Coalition. The first, a report by Sir Phillip Green and commissioned by the Coalition government on efficiency at the Central

Government level[43], does not refer specifically to the NHS but nevertheless, identified large scale inefficiencies at the Central government level. The second was published by the National Audit Office (NAO) in February 2011, titled "*The procurement of consumables by NHS Acute and Foundation Trusts*"[42]. This pivotal report details current procurement practices within the NHS set alongside the backdrop of the necessity to make £15-20 billion of annual savings by 2014-15. The report includes medical and surgical devices within its definition of 'consumables', which account for 49% of the nearly £4.6 billion (see figure 5, in [135]) spent on consumables per year, thus reiterating the theme of this thesis of ensuring value for money when purchasing hip prostheses by the NHS. The report summarises that, under the recent policy reforms of devolved responsibility (introduction of FTs), the majority of NHS Trusts are outside the DoH's direct control, meaning that hospital trusts have complete freedom to make decisions about which consumables to purchase and how they go about doing so. Within this, they can choose to involve regional collaborative procurement hubs and the NHS Supply Chain (national supplies and distribution organisation). However, this collaborative involvement is not compulsory and there is no mechanism to secure commitment by the separate trusts to purchase in a 'collective' manner. The report suggests that this means that "*significant economies of scale are being lost across the NHS*". As a consequence of fragmented purchasing, understandably, the report identifies widely varying prices for the same items taking place under small purchasing orders. Thus, the implication of the NAO report for this thesis is that the NHS is not able or willing to act as a single purchaser of consumables and exploit the buyer power resulting from its potential monopsony role in the market. This leaves it firmly open to the possibility that the main prostheses manufacturers may be able to exploit their potential for market power.

The report interestingly provides a case study (shown in figure 6.1): an example of 'savings from the pan-London framework for replacement hip and knee joints').

Figure 6.1, NAO Case Study - pan London framework for hip and knee prostheses

Savings from pan-London framework for replacement hip and knee joints in July 2008, the London Procurement

Programme, in conjunction with Epsom and St. Helier NHS Trust and the Elective Orthopaedic Centre established a pan-London framework for orthopaedic hip and knee implants. The aim of the new framework was to ensure that the prices trusts paid for implants would be low enough to enable trusts to recover the cost of procedures, for which they are paid standard (NHS tariff) rates. London Procurement Programme gathered evidence to show that trusts were paying too much for implants and were undertaking procedures at a loss, to influence trusts to join the contract. In 2009-10 the contract generated savings of £1.9 million on purchases of almost 6,000 hip and knee implants – a total spend of £11.5 million after savings. In early 2010, 18 trusts of the 24 London trusts which carry out orthopaedic surgery had joined the contract. *Source: London Procurement Programme*

Source: Figure 16, p.32, [70]

It explains that this particular framework has ensured that the prices trusts paid for implants would be low enough to enable the trusts to recover the costs of the procedures for which they are paid the standard NHS tariff rate (under PbR). The programme found that trusts were paying too much for prostheses and they used this to encourage trusts to join their programme. This example is evidence of how, by operating as a single united buyer with associated increased buyer power, enables the trust to negotiate better purchasing contracts. Future work could usefully extend this case-study to a more comprehensive assessment at the national level[27]Chapter 7 (amongst other things), builds on the NAO findings, determining how specialisation of individual hospitals compares with the NHS as a whole and whether there is evidence that smaller hospitals tend to be more or less specialised in their choice of prostheses than large ones.

Given the findings of the NAO report, there are some indications that the NHS is not acting as a monopsonist and thus not exploiting its buyer power. However, as mentioned above, this may not matter if the supply side i.e. the manufacturers of prostheses, is a perfectly competitive market. This motivates the further examination of the supply side industry below.

[27] This type of agreement could be viewed as collusive oligopsony and thus introduces the possibility that this type or agreement could breach competition laws. However, there does not appear to have been any evidence that arrangements of this type having been considered within the practice of competition laws in Britain. It is anticipated that it would not breach competition laws as the individual hospitals all form part of the same organisation in Britain: the NHS.

6.4 The Supply side

6.4.1 Previous literature

Very little literature currently exists on competition and purchasing within the hip prostheses market. In fact there is very little on purchasing and competition in medical devices in general (of say, diagnostics equipment, implants such as coronary stents and so on). Thus, it may be helpful to look first for parallels and contrasts with the Pharmaceutical industry which is more widely documented and researched within both Health Economics and Industrial Organisation. As will be shown below, the UK market for THR is dominated by just a few large firms, and the same is true in most parts of the Pharmaceutical Industry: for most drugs, there are only a few suppliers (the top ten firms accounting for 58.8% of the entire sector share in 2002[136]). Much of the competition in the drug industry takes place through innovation and the development of new drugs[137], rather than price. Uncertainty is also a key feature in pharmaceuticals – in terms of whether Research and Development (R&D) will generate a useable product, and also about the drug's long-term effectiveness. As discussed in earlier chapters, uncertainty is also a major issue with hip prostheses, with the long-term survival of hip prostheses unknown for 10 to 20 years post primary surgery.

The European Commission Competition website[138] includes many cases where firms (including drug companies) with market power have abused their position in a variety of ways, preventing or delaying competition[138]: Recent examples include investigations of AstraZeneca and other firms, who, it was believed, may have acted individually or jointly, notably to delay generic entry for a particular medicine[139]. AstraZeneca were also investigated by the Commission in 2005, for abusing the patent system by "delaying competition to a blockbuster drug from generic and parallel imported pharmaceuticals" for which they were fined 60 million Euros[140]. A similar case was recently investigated by the UK OFT on Reckitt Benckiser, who were alleged to have restricted competition for *Gaviscon* by de-listing the generic version from the NHS prescription channel[141].

Another area where pharmaceuticals companies have acted in an anticompetitive way is the use of 'kick-backs' to prescribing Doctors. This issue has received much attention in the US, even leading to the establishment of the website 'no free lunch'[142], reportedly run by health care

providers who '*believe pharmaceutical production should not guide clinical practice*'. The issue has also been highlighted by the BMJ in a recent article by Melanie Newman[143] who discusses the repeated kickbacks by drug companies in the US such as Pfizer and AstraZeneca. However, the literature is not confined to the US, in 1986, the then editor of the BMJ, Dr Richard Smith, highlighted the very same issue[144].

A third area where drug companies' have been frequently investigated in competition policy is merger cases. Although these do not necessarily raise a 'problem' for competition, the reason they are so frequently investigated is that they occur in highly oligopolistic markets with the potential for market power. A recent high profile example is the merger of Bayer and Schering 2006, investigated by the EU Commission[145].

In contrast, there are very few examples of abuse of market power in the medical devices sector, and more specifically, hip prostheses. The European Commission (EC) provides a web page on medical devices, stating that: "*The EU's involvement concerns mainly the regulatory framework for market access, international trade relations and regulatory convergence, all aiming to ensure the highest level of patient safety while promoting the innovation and the competitiveness of this sector*."[47]. However, this web page does not include information on issues of competition in the medical devices field. In 2009 the EC also published a report: "*Exploratory Process on the Future of the Medical Devices*"[146], part of which focused on' competitiveness and innovation of the medical devices industry'; this investigated the global innovation and competitiveness challenges faced by the industry including R&D, emerging technologies and green economy. This report recommended that focus be on, along with two other areas: supporting the competitiveness of the EU medical devices sector with an emphasis on supporting Small and Medium Enterprises (SMEs). It highlighted procurement procedures and reimbursement, suggesting that the duration of the supply contracts can inhibit competition by creating barriers to entry for other manufacturers; it also identified a tendency to centralise tenders with increased size (buyer power) suggesting that this might reduce competition and block the uptake of innovation. However, this reduction in competition should be less the case where commissioning takes place at the local level, as encouraged by both recent policies by the coalition government and previously by New Labour (section 6.3 above). Although this report includes no evidence or case history, it does demonstrate an increasing interest in competition issues and raises the

profile of economic issues related to medical devices. This is echoed in the health economics literature, for example volume 4 of the 'Value in Health' journal had a particular focus on whether economic evaluation of medical devices should be the same or different to those methods used for drugs[147].

There is also some evidence of kickbacks in medical devices: the website 'Pharmalot' has posted an example including: Johnson and Johnson (J&J) Depuy settling a kickback charge in 2007, and more recently, a general discussion of the 'Undisclosed conflicts among Docs and device makers' was posted. [148, 149]

The European Commission has also investigated two merger cases in THR: Johnson and Johnson and Depuy in 1998, which was not opposed by the Commission, and Smith and Nephew's attempt to acquire Centerpulse in 2003, which again was not opposed. However, in the second case, Zimmer made a more aggressive take-over bid later in the year to acquire CentrePulse. These reports are used below in the discussion of market definition.

The fact that there are few documented cases of anti-competitive practices in medical devices may simply mean that these rarely occur. But it might also be because medical devices tend not to be as high profile as pharamaceuticals. In the rest of this chapter and the next one, I will explore the possibility that anti-competitive practices might exist in the context of hip prostheses, by applying the theory and methods of Industrial Organisation to the THR supplying industry.

6.4.2 An Industrial Organisation Theory perspective

The theory of Industrial Organisation (IO) is the basis of competition policy. This section draws on two recent advanced textbooks in this area: Motta, [37] for an applied policy perspective, and Belleflame and Peitz[150] for the underlying theory (see also section 1.4.2 in chapter 1).

One of the key concepts in IO is that of market power. Market Power can be defined as '*the ability of firms to set prices above the marginal costs*' (Motta p. 39). As well as raising price to the consumer, the existence of market power can have various other negative impacts on competition, by lowering quality, restricting choice and slowing innovation, introducing barriers to entry of new firms[151], and not minimizing costs. Market Power depends on many things, including the structure of the industry – concentration (as a measure of the degree of oligopoly),

barriers to entry, product differentiation, buyer power. Monopoly is the most extreme form of dominance and market power[37]. However, it is always stressed that where there is only one or a few, big firms in the market, this does not necessarily lead to a welfare loss, there are cases where it would not be in the consumers' best interest to keep less efficient firms active in the market: "*competition policy is concerned with defending market competition in order to increase welfare, not defending competitors*"[152]. It may sometimes be that firms are monopolies or have few competitors simply because their products are superior to those of any other firms.
In the real world, pure monopoly is rare. In practice, most markets have at least a small number of different firms, i.e. oligopoly. Under oligopoly, market power can still exist however. In some cases, it may be that firms do not feel it necessary to compete actively on price and quality, but in other cases, firms may either explicitly or tacitly agree not to compete, sometimes referred to as collective dominance[151].

In the case where firms have formed a formal collusive agreement, this is known as a cartel, where the cartel itself will maximise profits if it acts as a monopoly, with all the harmful effects associated with monopoly, i.e. higher prices, lowering quality, restricting choice, slowing innovation, introducing barriers to entry of new firms and not minimizing costs. In most countries in the world, including the UK, the EU, and the US, cartels are illegal and thus viewed as not in the consumer's best interests. Most of what is known about how cartels behave is based on reports on illegal cartels discovered by the competition authorities (e.g. the European Commission or the OFT). The standard theory of cartels is based on a repeated game between a small number of players, which shows that it is possible to identify a set of market characteristics which make the existence of cartels more likely (chapter 14,[150], chapter 4,[37]). These are described at length in the above references, but briefly include, amongst other things:

- **High concentration, with a small number of similar firms**. This is because agreement between them is easier to form and sustain when there are fewer parties, while the gains from cheating on agreement are relatively higher when there are more firms. Agreement is also easier when firms are similar (e.g. have similar market shares). This is because all firms have similar incentives, and individual firms are less tempted to undercut the others because they realize that their rivals can easily match their lower prices.

- There are **barriers to the entry of new firms**. This is because it is harder for the cartel to maintain higher prices if there is the potential for new firms to enter the market and undercut them.
- There is **no buyer power**. This is because large strong buyers are more able to bargain down high prices and threaten to use alternative suppliers, as well as to detect coordination amongst suppliers.

From empirical studies based on the reports of competition authorities on prosecuted cartels [38, 153] it is known that members of cartels agree to: (i) share information on prices and quantities, and to (ii) fix prices and/or (iii) share the market. In markets where contracts are the subject of auctions or tendering, cartels rig the bidding so as to share contracts out. Market sharing can take one or more of three forms: (i) **territorial market sharing**, where the firms share out amongst themselves the right to be monopolies in different national or regional markets; (ii) **customer allocation**, where firms agree not to compete for the custom of each other's customers, and (iii) **quota market sharing**, where the cartel sets out specific stable quotas for each firm's market share, which they then agree to not exceed.

However, it is usually accepted that explicit cartels are relatively rare, although of course we only know about those ones which are actually detected by the authorities. Nevertheless, even where there is not explicit illegal collusion, it is argued that firms may sometimes 'tacitly' collude, so as to stay within the law. 'Tacit collusion' is said to occur where firms behave in much the same way as a cartel, but without explicitly meeting to share information and to agree on price and/or market sharing. In practice, this may take the forms of carefully watching other's prices, ensuring they move in parallel ways, avoiding price wars and/or aggressive behaviour towards each other. This is discussed and described alongside cartels by both Belleflume and Peitz ([150]chapter 14), and Motta([37] chapter 4). Again, the theory of tacit collusion is based on the theory of repeated games. This predicts that tacit collusion is more likely under similar conditions as listed above[154]. The results of tacit collusion are similar to those of cartels (high price, lower quality, less choice), and while there may be no explicit sharing of the market, it may still be reflected by firms not competing in each other's territories and implicitly accepting not to compete for each other's customers.

6.5 The Hip Replacement Supplying Industry

In order to establish whether market power of the sort described above side may be a concern in this case, I provide a detailed description of the supply side industry for hip prostheses.

6.5.1 Market definition

The initial step in assessing the degree of dominance and market power in any industry is to first define the market. With the market defined, the structure and the potential for power in that market is assessed, by taking into account various features of the supplying firms (e.g. their market shares), their buyers and the nature of the product. There are two standard dimensions to defining the market: Product definition and Geographic definition. In this case, these were both defined by the EC in two key merger cases [68, 69], which I will use to guide my own empirical work.

The Product Market

The EC defined the relevant product market as the market for hip prostheses for primary total hip replacement surgery. It identified the two broad types of hip prosthesis, as described in chapter 2: cemented (where the prosthesis is fixed with cement) and cementless (where the prosthesis is secured through biologic fixation i.e. bone grows into and through the pores in the prosthesis). However, there is a high degree of substitutability between the two and therefore both should be included in the market as one broad product market EC[68]).

The Geographic Market

In Europe, the responsibility for making a judgment on whether to allow any merger rests with the EC if the merger affects the market in more than one member state. This was the case in J&J and Depuy; for example table 6.1 reproduces the market shares of the merging firms in those countries where they would have a combined market share of 40%[28][68]. As is clear from both EC merger reports, most of the manufacturers are multi-national firms, present in multiple international markets, although this table illustrates that some manufacturers have a much higher presence in some countries than others: an extreme example of this is Biomet, who had a market

[28] This report is the most recent source of geographic market shares by manufacturers located in a literature search. Company annual reports report sales not market shares.

share of 25-35% in Portugal, but only 1-10% in the UK and no market presence at all in Ireland. The EC judged that the relevant geographic definition of the market is the individual member state as opposed to a single European market. It justified this on the grounds that prices and market shares of the main manufacturers differ from country to country, and argued that the extent of the manufacturer's presence in the market, in terms of training and assistance with the prosthesis, is an important factor to the purchaser (hospital/surgeon). The differing reimbursement systems across countries will also impact on the differences in the international markets - for example, England moved from a block contract system pre. 2004/5 to a DRG type reimbursement system (PbR), bringing it more in line with some other EU countries (Germany and France).

Thus, for the purpose of the rest of this thesis, the market is defined as hip prostheses in England.

Manufacturer	Ireland	UK	Portugal
J&J *	75 to 85%	40 to 50%	40 to 50%
Stryker Howmedica Osteonics**	5 to 15%	15 to 25%	1 to 10%
Zimmer***	1 to 10%	2 to 20%	5-15%
Biomet	-	1 to 10%	25 to 35%

Source: ***[68]***
* combined market share of J&J and Depuy
** referred to as Howmedica in the report
*** Zimmer (includes Sulzer market shares)

Table 6.1 - Market shares of 4 main manufacturers 1998

6.5.2 The Leading Firms

The EC's merger report on J&J and Depuy, identified 6 major players in Europe in 1998: Depuy (previously owned by Roche), J&J, Howmedica (previously Pfizer and since acquired by Stryker Howmedica Osteonices (Stryker)), Zimmer, Sulzer (acquired by Zimmer in 2003) and Biomet. It found that a merger between Depuy and J&J would give it a combined market share of 40 to 50% in the UK compared to 15 to25% for Stryker-Howmedica (the second largest player).

In Table 6.2, I use the NJR database to identify the main players in England and Wales, 2003-8. The NJR data used in this and the next chapter are compiled from data supplied to me by Northgate Information Systems in March 2009, explained in more detail in Appendix 2, including a discussion of the variables available on the NJR. This data includes observations up to the end of the financial year 2008/9. NJR data for 2009/10 has since become available which

has been used in the aggregated form in chapter 3. In total there are 25 manufacturers, but only 5 (Stryker, Depuy, Zimmer, JRI and Biomet) have a market share consistently over 5%. Of these, Stryker, Depuy, Zimmer and Biomet are all American based multi-nationals and are all conglomerates (i.e. hip prostheses are only part of their overall product range). It follows that hip prostheses in England form only a small part of their world wide activities. JRI is the exception: it is owned wholly by the British Furlong Research Charitable Foundation, and appears to only be present in the UK market (based on a small web based search and search of a sample of 4 other national joint registries) producing orthopaedic implants and surgical instrumentation only.

	2003	2004	2005	2006	2007	2008
Stryker	31	31	35	35	36	35
Depuy	38	37	34	33	34	34
Zimmer	6	7	7	8	8	7
JRI	8	9	9	8	7	7
Biomet	6	5	5	6	6	5
Smith and Nephew	2	3	3	3	3	3
Endoplus (UK) Limited	1	1	1	1	1	2
Corin	1	1	1	1	1	1
B Braun/Aesculap	1	1	1	1	1	1
Waldemar Link	1	1	1	1	1	0
Wright Medical UK	1	1	1	1	1	1
Other	4	4	3	4	4	4

* Market shares measured by volumes of prostheses, displayed as percentages
** There are 14 other manufacturers, none of which has a market share of more than 0.25% in any one year
Source: the authors calculations, based on primary NJR data

Table 6.2 The Leading Suppliers to the market in England and Wales: market shares* (%) 2003-8

Table 6.3 shows the world's largest manufacturers. Where appropriate, figures have been converted into UK sterling and inflated into 2011 prices[16]. All four US manufacturers are based in Indiana, where most early US innovation in this industry took place, their UK headquarters are all in the South/South-West with the exception of Depuy, who are based in the North of England, Leeds.

Manufacturer	Headquarters	UK office	Annual worldwide sales (£million)*
Stryker [44]	Warsaw, Indiana, USA	Newbury, Berkshire	£4,666
(J&J) Depuy [45]	Warsaw, Indiana, USA	Leeds	£3,728
Zimmer[46]	Warsaw, Indiana, USA	Swindon	£2,842
Biomet Orthopaedics[47]	Warsaw, Indiana, USA	South Wales & Swindon	£1,492†
Smith & Nephew[48]	London, UK	London	£1,155†*
B Braun Medical Limited+[49]	Sheffield, UK	Sheffield, UK	£1,300†*
Smaller UK firms			
JRI Ltd [155]	Sheffield, UK	Sheffield	£86†*
Corin[50]	Cirencester, Gloucester	Cirencester, Gloucester	£43‡

* converted to UK £ from US $, price year 2009 inflated to price year 2011[16]; † converted to UK £ from US $; Price year 2006 inflated to price year 2011[16]; ‡ price year 2009 inflated to price year 2011; †* price year 2008 inflated to price year 2011; +part of B Braun Aesculup

Table 6.3, The world's top manufacturers

Depuy is part of the wider diversified Johnson & Johnson group (J & J), while Stryker is more specialised in hips, knees, spine and trauma, which accounts for approximately 61% of their overall sales (of which hips account for approximately 15%). Stryker has the world's largest single market share for hip prostheses (Depuy 13%, Zimmer 10%, Smith and Nephew 8%, Biomet 6% and others 28%)[156]. Zimmer and Biomet are also U.S based multinational firms who both report the majority of their sales taking place in the US[157, 158]. The other four manufacturers shown in table 6.3 are based in the UK, although Smith & Nephew and B Braun Medical Limited have a presence in the US market[159], with Smith & Nephew, basing the Orthopaedic part of the company (which includes hip prostheses) in Memphis Tennessee in the US[160]. For comparative purposes, the table shows the two much smaller UK firms, Corin and JRI, who are relatively large in the UK (see table 6.2). Corin is also present in European markets, with only a 5% presence in the U.S market[161]. Corin is based in Cirencester, England and JRI is based in Sheffield, England.

6.5.3 Concentration of sellers

A key feature of market structure is the level of concentration. This is roughly a measure of the degree of oligopoly. Table 5.2 shows that there is a duopoly of two large firms, Stryker and Depuy, who between them, account for 69% of the market. There are three other large firms, but with much smaller shares of 5-7% (Zimmer, Biomet and JRI), and then a fringe of very small

players who together account for little more than 10%. As mentioned above, Depuy companies (formerly part of Roche) was acquired by the Johnson & Johnson group in 1998, following clearance by the EC. There have been a number of other smaller acquisitions during the time period used in this chapter, most notably: Zimmer acquired Centerpulse in October 2003, increasing their international market share by approximately 3%, and Smith and Nephew acquired Medical Technologies in March 2004, increasing their international market share by 0.01% (Midland Medical Technologies is more focused on the hip resurfacing market).

In IO and Competition Economics, the concentration of sellers is measured by a variety of different statistical indices. The two most common are as follows.

HHI: The Herfindahl-Hirschmann Index of concentration, defined as the sum of squared market shares of all firms[162].

$$HHI = \Sigma s_i^2 \text{ for } i=1\ldots..N \qquad (1)$$

The HHI can vary from 1/N to 1: the lower limit occurs if there are N different manufacturers, each with an identical market share (1/N each); the upper limit occurs if there is only a single manufacturer in the market (i.e. a monopoly). For presentational purposes, the index is sometimes expressed in a reciprocal numbers equivalent form:

$$N(HHI) = 1/HHI \qquad (2)$$

This translates the distribution of firms into a hypothetical number of equal sized firms – the number of equal sized firms who would record that value of HHI if they had equal shares (see below for examples)

The other most common concentration measure is the concentration ratio[163], which records the combined market shares of the largest firms. Here, we use:

CR2: the two firm concentration ratio which measures the combined share of the two largest firms in the market, and

CR5: The five firm concentration ratio, which shows the combined shares of the top five suppliers.

Table 6.4 reports the values of each index on a yearly basis. As can be seen, the HHI has remained fairly constant throughout, at around 0.26, indicating a consistently highly oligopolistic market: in terms of its numbers equivalent, equivalent to a market of just four (equalised) firms. CR2 is also consistently 68-70%, and CR5 shows that the top five manufacturers account for approximately 90% of the market.

Table 6.4: Seller Concentration in England and Wales, 2003-8

	2003	**2004**	**2005**	**2006**	**2007**	**2008**
HHI	0.26	0.25	0.26	0.25	0.26	0.26
CR2	69%	68%	69%	68%	69%	70%
CR5	90%	89%	89%	90%	90%	89%

Source: Authors calculations, based on primary NJR data

Table 6.4: Seller Concentration in England and Wales, 2003-8

6.5.4 Firms' Product Portfolios: the Prostheses

Table 6.5 lists the leading brands for cups and stems. These include 6 stems and 10 cups. The 'best-selling' brand is the Exeter V40 (Stryker), which is a cemented stem and accounts for approximately 36% of all stems. For reference, Appendix 1 reports the combinations of different cups and stems which are most often used in England Wales – usually the cups and stems are produced from the same manufacturer which may reflect regulatory approval, technical compatibility and complementary marketing of the products (only 5 out of the 26 combinations listed use cups and stems from different manufacturers). The most common cup and stem combination (by almost 4,000) procedures is the Pinnacle cementless cup with the Corail cementless stem, both manufactured by Depuy.

Brand	Manufacturer	Prosthesis Type	2003	2004	2005	2006	2007	2008	Total 2003-8
Exeter V40	**Stryker**	Stem	31%	34%	38%	37%	36%	35%	36%
Corail	**Depuy**	Stem	3%	6%	8%	11%	16%	21%	21%
Contemporary	**Stryker**	Cup	7%	7%	11%	9%	11%	12%	8%
Pinnacle	**Depuy**	Cup	5%	3%	7%	11%	15%	19%	10%
Charnley stem	**Depuy**	Stem	17%	13%	10%	7%	5%	3%	8%
Elite Plus Ogee	**Depuy**	Cup	10%	9%	9%	8%	8%	7%	8%
Trident	**Stryker**	Cup	2%	4%	6%	9%	11%	12%	8%
Trilogy	**Zimmer**	Cup	8%	7%	7%	8%	7%	7%	7%
CSF	**JRI**	Cup	6%	7%	7%	7%	6%	4%	6%
C-Stem	**Depuy**	Stem	6%	7%	8%	7%	5%	3%	6%
Furlong	**JRI**	Stem	6%	7%	7%	8%	6%	6%	7%
Charnley cup	**Depuy**	Cup	10%	8%	6%	4%	3%	2%	5%
Charnley Ogee	**Depuy**	Cup	9%	7%	5%	3%	3%	2%	4%
CPT	**Zimmer**	Stem	4%	4%	4%	5%	6%	5%	5%
Elite Plus	**Depuy**	Cup	4%	4%	4%	4%	5%	4%	4%
Exeter Duration	**Stryker**	Cup	7%	6%	5%	5%	4%	3%	5%
Other			69%	66%	58%	53%	53%	55%	58%

Source: Authors calculations, based on primary NJR data
Note: all market shares represent shares of the total for each component type, e.g. the Exeter V40 has a 36% share of al stems. For this reason, each column sums to approximately 200%.

Table 6.5, Market shares for the top 16 brands, 2003-8

The HHI index of concentration for prosthesis brands is 0.161 for stems and 0.057 for cups which is equivalent to approximately 6 equal sized stems and 17 equal sized cups. This suggests a fair amount of diversification of brand choices within the NHS on average, although there is clearly more specialisation/concentration higher in stems. In the next chapter, it is convenient to refer to a single measure of specialisation for each hospital, and I will use the average of the index for cups and stems. At the national level, the average of the above HHIs for stems and cups is 0.109, which translates into roughly 10 equal sized brands.

Figures 6.2(a-d) depict the volumes implanted for these main brands.

Figures 6.2a-d

Figure 6.2 a - Number of prostheses implanted for the main 6 brands of cemented cup

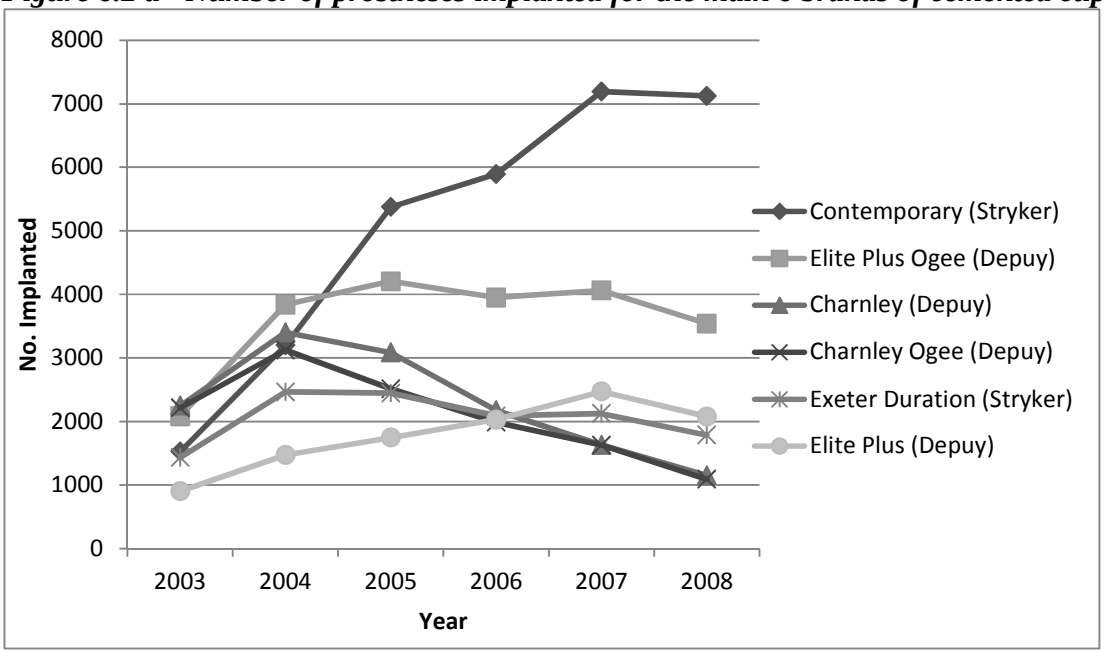


Figure 6.1 b - Number of prostheses implanted for the main 6 brands of cemented stem

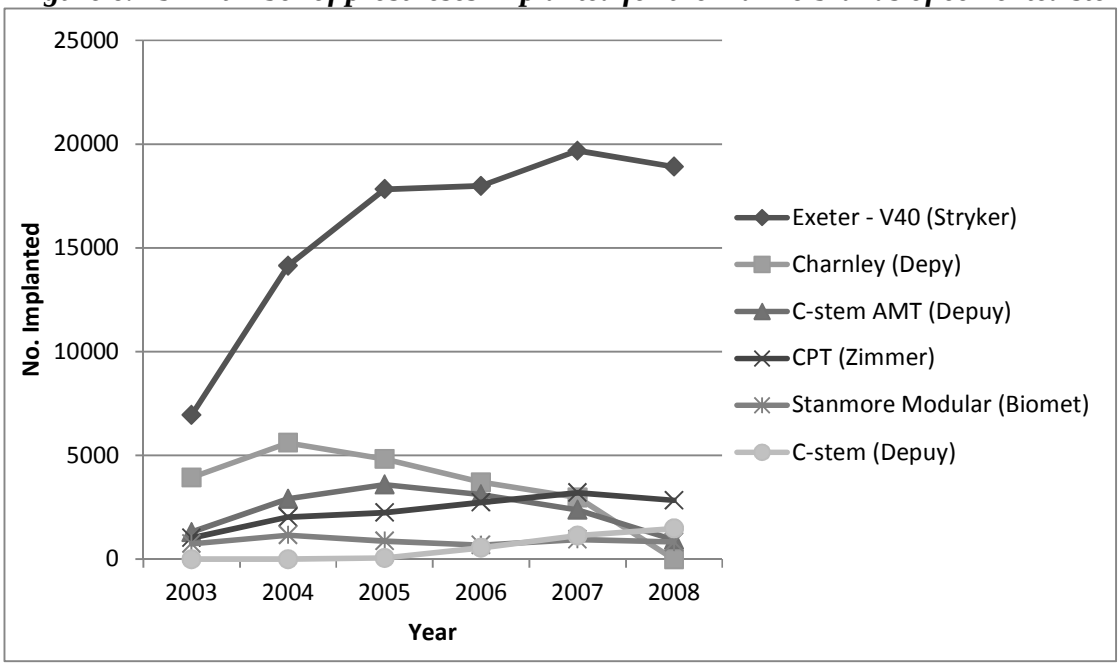


Figure 6.1 c - Number of prostheses implanted for the main 6 brands of cementless cup

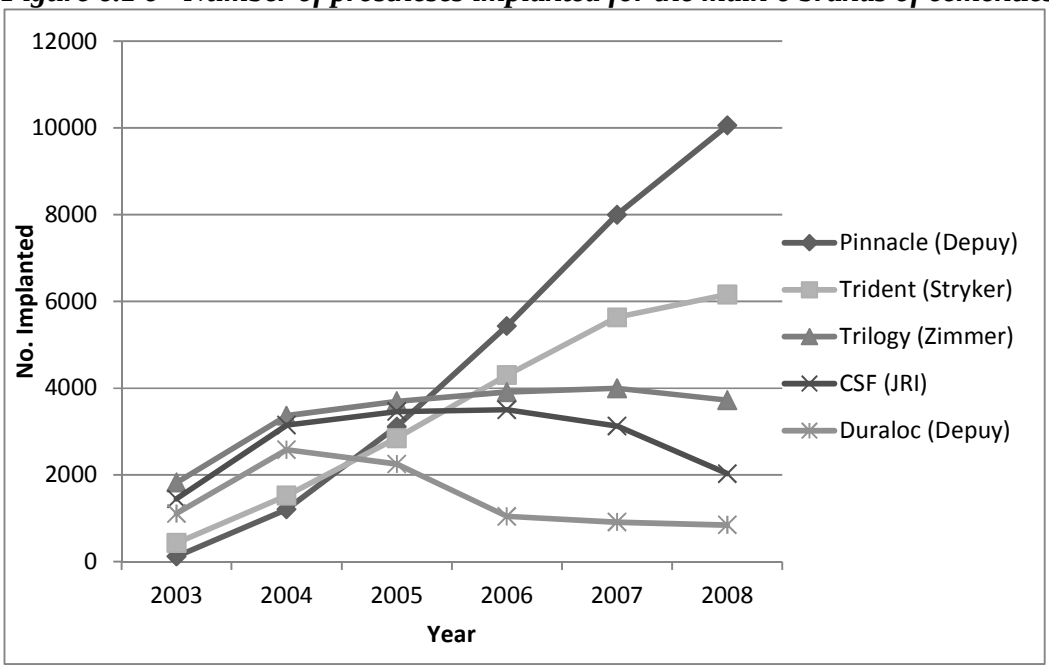
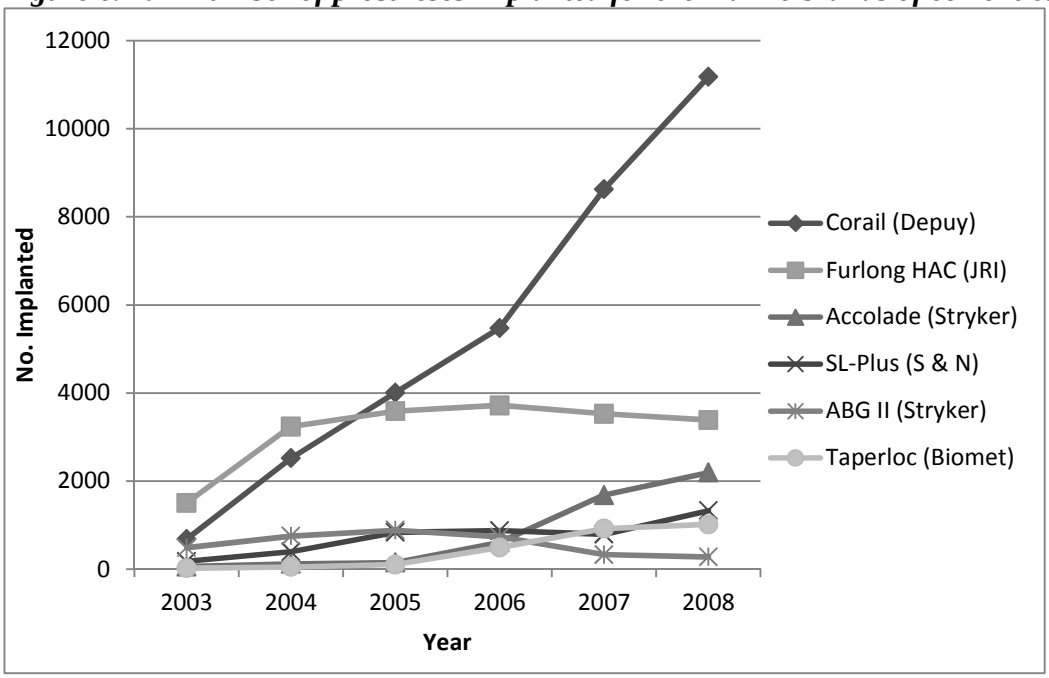


Figure 6.1 d - Number of prostheses implanted for the main 6 brands of cementless stem



What now becomes clear is that Stryker appears to be dominant in the cemented market, particularly with the Exeter V40, which has grown its market share from 40 to 60% of stems, but

also in cemented cups, its Contemporary model has increased its share from 10 to 30%. Stryker's growth in the cemented market, means that it has been able to maintain a 35% overall share (Table 6.2), despite the fact there is the general trend to implant cementless prostheses, in which it has a much smaller presence. However its cementless cup (Trident) has also grown in market share from 6 to 21% by 2008.

In contrast, Depuy is clearly dominant in the cementless sector with a rise in their cementless stem, Corail. This is clearly at the expense of the JRI's Furlong which starts off fairly dominant with 38% of the market share compared to Corail's 17%. However, by 2008, Corail has risen to 46% while Furlong has fallen off to only 18%. In cementless cups, Depuy has gained some market share with the Pinnacle growing from 2 to 33%, alongside Stryker's Trident, both of which appear to have been at the expense of Zimmer's Trilogy which drops from 23 to 12% by 2008.

Overall, it appears that Stryker and Depuy are both dominant, but separately in the cemented and cementless sectors respectively. JRI still has an important presence in the cementless sector both for cups and stems, but its market share has reduced year on year in the face of the rises by Stryker and Depuy in this sector. Other manufacturers such as Zimmer, Biomet and Smith and Nephew have some prostheses in each sector which have stayed fairly constant around the 2-3% of market share, with the Exceed cementless cup (Biomet) rising from 1 to 4% - the only real rise for any manufacturer other than Stryker and Depuy.

6.6 Implications for market power

These values of the concentration indices indicate a highly concentrated market. For example, all of the leading competition authorities in the UK, EC and UK designate a market as 'concentrated' if HHI>0.10, and 'highly concentrated' if HHI>0.20[164]

When identifying whether individual firms are dominant, there does not appear to be one clear rule: a market share of 50% or more would usually be interpreted as evidence of dominance, while the OFT in their "Assessment of Market Power" guidelines state that below 40% it is unlikely that a firm is considered dominant [37]. In this case, neither Depuy nor Stryker would

be considered to be singly dominant, but a combined market share for the top two of nearly 70% might be interpreted as potentially 'collectively dominant', especially bearing in mind Depuy's strength in cementless and Stryker's in cemented.

As discussed earlier (section 6.4.2), high concentration is considered to be a necessary condition for firms to be considered dominant, and a market to be considered as uncompetitive, or in extreme cases, open to collusion. However, high concentration alone is not sufficient for this conclusion, other evidence is required. Even very large firms may not be able to exploit their strong market shares if entry into the market by new firms was potentially easy, or if a strong buyer could use its buyer power to resist high prices/poor quality or to threaten switch to competing suppliers. The two EU Commission merger reports mentioned above [68, 69] found that in the European market for medical devices, there had been expansion by the current manufacturers into new products and countries, and an absence of significant entry barriers (establishment and transport costs, distribution networks, patents and R&D etc.). However, this may not be the case at the national market level, where in England, as we have seen, there are only two main suppliers (perhaps jointly dominant), few small firms and very little evidence of new entrants in the period covered by the NJR.

Of course, clear evidence would be necessary before concluding that firms in a given market were colluding. This would require evidence of price fixing and/or some form of market sharing. In this case information on prices is not available (as acknowledged in the NAO report on procurement in the NHS[135]). However, the NJR data will allow at least some preliminary analysis to identify any indirect evidence of market sharing. Recalling the discussion in 6.4.2, market sharing could be observed through one or more of the following: territorial market sharing at the international or regional levels - for the purpose of this thesis, this might be observed at the regional level within England; allocative market sharing - certain hospitals purchasing only from specific manufacturers; and general quota market sharing - the manufacturers agree on retaining a stable quota division of market shares between them. One of the purposes of the next chapter is to look for any indirect evidence of this.

6.7 Conclusions

The descriptive statistics presented in this chapter indicate that potentially a dominant duopoly exists in the UK market (Stryker and Depuy) and these two manufacturers appear to be maintaining their grip over time. They have achieved this by consolidating the shares of their established brands and by growing the market shares of their newer brands (e.g. Corail and Pinnacle by Depuy).

As already explained, large or very large market shares are not necessarily a cause for concern. The strength of the leading brands (and therefore firms) may simply reveal that they are superior technically, and that surgeons and hospitals in the NHS are recognizing this, for example the clinical evidence for Stryker's Exeter hip and Depuy's Corail hip show that both are ODEP 10A rated (these two components are virtually the only femoral components that have 15 year survival outcomes data available from generalisable registry data and the peer-reviewed literature[165, 166]). Although there is not much evidence of significant entry of new firms into this market, it must be remembered that if there were a dominant buyer in the UK - the NHS - then this might be sufficient to constrain the potential market power of the two main manufacturers in this market. However, if as indicated by this chapter and the findings of the NAO report, the NHS is faced with sellers with considerable potential selling power, then the question is how should the NHS conduct its procurement with respect to these suppliers? The next chapter turns to a deeper analysis of the choices made within the NHS, at the level of the individual hospital, and attempts to identify and explain any patterns which emerge from the disaggregated data.

Chapter 7, Hospital choice of hip prostheses

7.1 Introduction

Chapters 2 to 6 highlighted the large number of hip prostheses on the market. Using NJR data, the previous chapter (6) identified that the prosthesis industry as a duopolistic or concentrated oligopolistic market, suggesting that the NHS could be faced with sellers with significant selling power. As outlined previously, if the NHS purchases as a single homogenous entity then it may be able to exploit its buyer power to attempt to overcome the seller power and achieve efficient purchasing of joint prostheses. However, as discussed in the previous chapter, there is some evidence to suggest that the NHS is currently purchasing at a disaggregated, local level and thus has the potential to suffer inefficiencies.

This chapter builds on the preceding chapters, empirically testing choice of hip prostheses at two levels, first in terms of patient choice: how well the characteristics of the patient explain choice and beyond this, choice at the hospital level to determine whether the NHS acts as a homogenous powerful entity, or whether the manufacturers of prostheses are able to exploit their potential seller power. It is organised in five sections: Section 7.2.1 sets out the hypotheses which will be tested empirically. Section 7.2.2 discusses the theoretical model, Section 7.3 describes the data set. Section 7. describes the methods of analysis – the specification of dependent variables and estimating equations and the econometric estimators. Section 7.5 reports the main results and section 7.6 re-visits the original hypotheses to provide a discussion of the main results. Finally, section 7.7 summarises the main conclusions and sets out an agenda for future work.

7.2.1 Hypotheses

The analysis is structured around the following hypotheses:

1. Choice of prosthesis is largely determined by the characteristics of the patient.

As discussed in the previous chapter, the agent is the surgeon/hospital, choosing which prosthesis to implant on behalf of the patient (the principal). If it is found that it is mainly patient characteristics that determine which prosthesis they receive, then this implies little or no discretion on the part of the agent, and it is as if the agent's (surgeon) role is merely to identify

the ‘correct’ prosthesis for that patient, i.e. the one the patient would have chosen herself if she were fully informed. This is tested at the aggregate level, i.e. just between broad prosthesis type (cemented or cementless) – if it is not confirmed at that level, it would be pointless to conduct further tests at the level of individual brands of prosthesis.

If this hypothesis was accepted, then there would be little need for any further analysis. However, if the hypothesis is rejected (which it overwhelmingly is), then the subsequent stages in the analysis are devoted to explaining prosthesis choice, having controlled for patient characteristics, in terms of the following hypotheses.

2. The NHS is a homogenous entity in its prosthesis choices – the extent of specialisation is similar across all hospitals

If the NHS is a homogenous entity then it might be expected that individual hospitals replicate the same degree of specialisation as the aggregate national for England and Wales observed in the previous chapter. In that case, individual hospitals would be like ‘mini-clones’ of the NHS as a whole.

However, it is very unlikely that this could apply to smaller hospitals[29], simply because of the low number of patients implanted – previous chapters have shown that within the NHS as a whole there have been large numbers of different brands implanted, and it would be impossible for a hospital with, say, only 50 patients to employ such a large number. Therefore, a more plausible secondary hypothesis is that:

3. Larger hospitals are less specialised than smaller hospitals, and, as hospital size grows, specialisation tends, at the limit, to the national level

Thus, evidence is sought that, although smaller hospitals are inevitably more specialised in their choice of prostheses than large ones, this effect declines as activity increases. This might be expected if there are scale purchasing discounts, and/or a 'learning curve effect' by consultants (so that it is best for them to use only a single prosthesis when they only undertake a small number of operations.) On the other hand, there may be an offsetting effect because the biggest

[29] Throughout this chapter, I equate hospital activity or size with the number of patients implanted.

hospitals present the suppliers with the largest and most profitable potential sales opportunities, and suppliers might attempt to tie the biggest hospitals into exclusive purchasing deals.

4. There are predictable differences between various 'segments' within the NHS

While it may be that there is considerable variability within the NHS as a whole, it may be that a large part of this can be accounted for by (i) systematic regional differences or by (ii) the type/status of hospitals (identified here by strategic health authorities (SHA) regions); and hospital status (here I distinguish, FTs from NHS Trusts, IS and ISTC's because their different financial structure may affect their purchasing policies.)

5. Part of the observed variation in choice is the result of systematic behaviour by the manufacturers.

It was shown in the previous chapter that the supplying industry is dominated by two large suppliers, DePuy and Stryker. It was suggested that this might lead to 'tacit collusion' between them. One of the effects of tacit collusion is that price will tend to be higher. This cannot be tested with the current data-set since it does not include any information on the prices paid by individual hospitals. This is one of the intended areas of future research.

However, another potential effect of collusion is for firms to share out the market in some systematic way. In extreme forms of collusion (cartels) firms may even make formal agreements in which each of them is allocated certain territories or customers (see previous chapter.) There is no reason to think that a cartel exists in this market, but it may be that the suppliers informally accept that some regions or hospitals are 'their territory', while others are the territory of other suppliers. At a superficial level at least, it is worth recalling Table 6.3 and noting that Stryker is located in Berkshire (30 miles or so from Central London), while Depuy is located in Leeds (in Yorkshire, North England). Examination of whether the 'market shares' of Depuy and Stryker are systematically different by hospital status, size, and the region of the hospital (SHA) will provide indirect evidence of this[30]. If there is evidence of this sort, then future research will make

[30] Although it should be noted that Depuy is located in Leeds, having moved to the headquarters of the Chas F Thackeray company which it acquired in 1990 and Stryker are based in the Thames valley after it acquired Howmedica. Both manufacturers provided training to surgeons in their original implants, with training focussed initially in the area of their regional base.

a deeper analysis for the potential evidence that the manufacturers are exploiting their seller power in the form of market sharing either by region or hospital.

6. Finally, when examining each of these hypotheses, the potential effects of the introduction of PbR in 2006/7 will also be considered.

Expectations of the likely effects are not clear-cut. At this stage, the main purpose on this is to establish whether the introduction of PbR has led to any significant changes in the purchasing behaviour of individual hospitals. The causes of any such changes, if any, will be the subject of future research.

7.2.2 The economic model

In order to test the above hypotheses, this chapter applies a series of econometric tests based on the following theoretical model.

The model draws on random utility theory, using a discrete choice model, first developed in the analysis of consumer behaviour[31]. Discrete choice models of consumer behaviour describe the decision makers' choices between all the alternatives. If the decision maker is assumed to be the consumer, then this specific type of model establishes the link between consumer preferences and the aggregate demand function.

In the context of this thesis, unlike in standard consumer theory, there is a principal-agent dimension to the model, where the patient is the principal and the agent is the hospital/surgeon (at this stage, the identity of the agent is not specified).

Therefore, I proceed in two steps - first specifying the principal's utility and then the agent's. I assume that actual choice is based on the latter. In a general form the following describes the

---

[31] For a deeper discussion of random utility theory with discrete choice models, see: 167. Nevo, A., *Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry.* The RAND Journal of Economics, 2000. **31**(3): p. 395-421, 168. Berry, S.T., *Estimating Discrete-Choice Models of Product Differentiation.* The RAND Journal of Economics, 1994. **25**(2): p. 242-262.

utility patient i would derive from prosthesis j. For expositional convenience, the potential subscript t, for date of implant is suppressed.

$$U_{ij} = \boldsymbol{\beta}_{ij}\boldsymbol{X}_i + \mu_{ij} \qquad \{1\}$$

where:

U_i is principal i's utility

X is a vector of patient characteristics (some of which maybe unobserved by the econometrician), and β_{ij} and μ_{ij} are effectively 'taste' parameters. Thus $\beta_{ij}X_i$ refers to the utility patient i would derive from brand j given its vector of characteristics X_i. Note that β_{ij} may vary across patients and prostheses.

The prosthesis which is most appropriate is the one which maximizes the principal's utility. However, as described in section 6.2.2, in this context, it cannot be assumed that this is necessarily the prosthesis which is actually implanted, because the 'principal' is not fully informed and effectively it is the agent that makes the choice.

In making this choice I assume that the agent will take account of cost, and therefore the price of the treatment should appear positively in what is now the net utility function for agent k:

$$U_{ijk} = \boldsymbol{\beta}_{ijk}\boldsymbol{X}_J - \alpha_k P_j + \boldsymbol{\gamma}_{jk}\boldsymbol{Z}_{k\_+} \mu_{ijk} \qquad \{2\}$$

where:

Z_k is a vector of the characteristics of agent k

P_j is the price of the prosthesis

β_{ijk} is the agents perception of the principal's 'taste'

It is assumed that the agent chooses the prosthesis which maxmises his/her utility function. This is the equation which implicitly underlies all the following empirical work. In the first part of the empirics, it informs the equations identifying the choice between cemented and cementless prostheses in terms of the patient characteristics. In the later parts it supports the chosen mix of

prostheses at the hospital level which is implicitly based on an aggregation of the choices at the individual patient level.

To specify the variables in the patient characteristics vector, I draw on the list of characteristics often included in previous studies of THR.

Principal (patient) characteristics:

*Age – It is expected that younger patients are more likely to receive a cementless prosthesis, as reported in section 2.2.4.

*Gender – It is expected that males are more likely to receive a cementless prosthesis, as reported in section 2.2.4

BMI

Height

Weight

*Pre-operative health status

Socio-economic status

Marital status

Dependents

*Employment

Side of surgery

Bilateral surgery

Position during surgery

In the present context it is not obvious that all of these variables will have an impact on the *choice of prosthesis*. The variables which potentially are most important are starred in the above list with my prior expectations of their likely effect. In the other cases I have no strong priors.

Clearly some assumptions need to be made regarding who the 'agent' is. However, for the purpose of this thesis, I must remain agnostic as to the identity of the agent pending future work on decision making in hospitals for hip prostheses. At this stage, the agent will be equated to the

hospital[32], although it is accepted that the hospital is the aggregation of all surgeons within the hospital, and ideally, some recognition would be made of heterogeneity between sugeons within individual hospitals.

Agent (hospital) characteristics:

Hospital status (NHS trust, Foundation trust and so on)
Supplier (manufacturer) preference
Hospital teaching status
Regional location of hospital
Socio-economic status of the region of hospital location
Size of the hospital (with respect to scale of hospital and volume purchases)

I have no strong priors as to how these characteristics will impact on choices. The purpose here is largely descriptive - to identify what the data reveal.

A key limitation for the empirical model is that data on a number of these variables is not available, including price of prosthesis and the identity of the surgeons who are working within a hospital. Addressing these omissions will be an area of future work.

7.3 Data

The analysis utilizes the individual patient level data received[33] from the NJR and HES (hospital episodes, collected from all hospital admissions in England). These data cover the period 2003-2008. The data are used in two forms for the analysis of this chapter:

1. At the individual patient level (NJR and HES linked data-sets). This is used for the initial stage focusing on the role of patient characteristics

[32] Further work will aim to disentangle this issue.
[33] Note that this does not include 2009 because this year was unavailable at the time (March, 2009). The analysis in the earlier chapter 2 does include 2009 in aggregate form because that chapter was written after the publication of the 2009/10 NJR Annual Report.

2. Hospital level data (using only just the NJR, aggregated to the hospital level for each year). This will be used to test the hypotheses relating to hospital specialisation and purchases from the two leading manufacturers.

The cleaning and merging of the NJR and HES linked databases is explained in detail in Appendix 2. However, in brief, it unavoidably involved dropping a large number of patient level observations due to differences in the sizes and make-ups of the data-sets provided: I was provided with 199,457 individual patient level observations from HES, and 350,238 individual patient level observations from the NJR, thus the merging process inevitably involved losing some of the NJR observations. Patients in Wales and the Independent Sector patients were dropped because they are included in the NJR database, but they do not appear in HES, this accounted for the majority (n=167,502) of the 'lost' observations. These lost observations mean that any results are confined only to a population of patients in England who are NHS funded. It would be inappropriate to generalise any findings to the population of all patients in England and Wales, without further investigation of whether these exclusions might lead to selection bias (for example, in terms of patient mix). The final linked HES and NJR data-set contains 145,870 patient observations. The strength of the linked HES data-set is that it contains more patient characteristics variables than the NJR. However, as will be shown, patient characteristics explain very little of the choice of prosthesis, and consequently the decision was taken to use the NJR data-set only for most of the analysis since this covered far more patients: 278,063 individual patient level observations[34], including patient characteristic variables for only: gender, age and side of surgery. These NJR data were then aggregated up to the 'hospital-year' level, for example: the number of Exeter V40 prostheses implanted in a given hospital in a given year. When aggregated to the hospital level, the dimensions of the resulting panel are as in table 7.1. below:

	2003	**2004**	**2005**	**2006**	**2007**	**2008**	**Total**
No. Hospitals	306	341	341	350	344	336	1948
No. Patients	22061	41512	46042	46214	51152	51088	258069

Table 7.1, Dimensions of the hospital-level data-set

[34] 71,175 individual patient level observations were dropped from the original 350,238 NJR data-set for the following reasons: they were revision or resurfacing procedures or they were incorrectly coded as knee procedures.

7.3.1 Dependent variables

As explained below, the analysis is structured into three stages. In the first stage, the dependent variable refers to the identity of the prosthesis implanted in the individual patient data-set; the second stage analyses the extent of specialisation in choice of prostheses by the individual hospital; and the third stage examines what determines the extent to which the hospital concentrates its purchases on the two main suppliers (Depuy and Stryker). The precise specification and measurement of these dependent variables is described in section 7.4.

7.3.2 Covariates

The explanatory variables to be employed include a vector of patient characteristics and a vector of hospital characteristics

Patient characteristics

For each patient, I have data on 9 characteristics: age, gender, side of surgery (left or right), patient position during surgery[169], ethnicity, whether bilateral surgery was carried out, primary diagnosis of osteoarthritis, whether Minimally Invasive Surgery was used[170], and whether Image Guided surgery was used[171][35]. The descriptive statistics are shown in Table 7.2(i).

[35] Minimally invasive surgery (MIS) and Image Guided Surgery (IGS) are surgical techniques. MIS uses a single small incision to avoid damage to muscles and tendons, using specially designed retractors and customised instruments to expose the hip joint[160]. The NJR annual report 2010, reports that less than 5% of all hip procedures in 2009/2010 used MIS[8]. Image guided surgery is the use of sophisticated computer technology for the optimization of surgical performance[161].

	Mean
Mean patient age (years)	69
Proportion of Female Patients	62%
Proportion of patients receiving right side prosthesis	55%
Proportion of patients positioned laterally	85%
Proportion of patients classed as 'white'	99%
Proportion of patients undergoing bilateral surgery	0.2%
Proportion of patients receiving minimal invasive surgery (MIS)	5%
Proportion of patients receiving image guided surgery (IGS)	0.4%
Proportion of patients with a primary diagnosis of Osteoarthritis	87%

Table 7.2(i) Descriptive statistics of explanatory variables in the individual patient level panel (n=145,870)

	Mean or %
Mean patient age (years)	70
Proportion of Female Patients	63%
Proportion of patients receiving right side prosthesis	45%
Hospital Size (number of implants)	336
NHS Trust	78%
NHS Foundation Trust	11%.
Independent Sector	6%
NHS Trust Treatment Centre and 'Other'	3%
Independent Treatment Centre	3%
South East	11%
East	13%
East Midlands	7%
North West	17%
London	11%
Yorkshire	9%
South West	9%
South Central	8%
West Midlands	10%
North East	5%

Table 7.2(ii) Descriptive statistics of explanatory variables in hospital panel (n=1948)

In the hospital-panel form, these variables are measured as the averages or proportions of all patients receiving implants in each hospital in a given year. Unfortunately, only three of the 11 above characteristics (age gender and side of surgery) can be employed as these are the only characteristics reported in the NJR data-set – see Appendix 2[36]. There is also a small number (70) of missing observations on patient age and gender in the NJR, and these have been dropped from the hospital panel. The descriptive statistics of the remaining 1948 hospital panel observations are shown in Table 7.2(ii).

Hospital characteristics

Hospital size (THR activity)

Hospital 'size' here is measured by the number of prostheses implanted by a hospital in a given year; therefore this should be interpreted as a measure of the extent of a hospital's hip replacement activity, rather than as a measure of its more general size. Table 7.3 shows the size distribution of hospitals according to this measure

It is important to reiterate here that 2003 was an incomplete year for data coverage in the NJR, hence the pattern for hospital activity differs somewhat from the subsequent years. Apart from this, the mean number of implants per hospital has remained fairly steady across the years. Most hospitals implant between 50 and 300 prostheses in a given year, but there is also a few very active hospitals (>300), while roughly half are fairly small (<100).

	<50	**50-99**	**100-199**	**200-299**	**300-399**	**400+**	**Average no. patients per hospital**
2003	145	96	48	12	2	3	306
2004	98	90	93	41	10	9	339
2005	69	87	114	51	9	11	341
2006	87	99	95	46	7	16	350
2007	75	95	87	52	19	16	344
2008	63	87	101	57	13	15	336

Source: NJR data[9, 58, 64-67, 172]

Table 7.3, Size distribution of hospitals: size measured by number of implants

[36] The NJR includes data on patient height, weight and BMI. However, there were considerable missing observations for these variables and would have significantly reduced the data-set size had they been included in the regressions.

Status of hospitals

Table 7.4 distinguishes five broad types of hospitals[37]: About 80% are NHS Trust hospitals, 11% are FT, and 6% are Independent Sector (IS) hospitals. As can be seen, there is a continuing rise in the number of FTs, after their introduction in 2003/4. The final row shows the number of hospitals who have changed their status from NHS Trust to FT in each year - most of the switches take place in 2006/7 which is consistent with the data provided on FT by Monitor (the regulator for FT) [173].

Status	2003	2004	2005	2006	2007	2008	Total
Trust	278	302	286	279	241	223	1,609
Foundation	2	14	23	41	74	77	231
Independent	18	24	22	17	17	25	123
Other provider & NHS Treatment Centre	7	8	12	9	8	8	52
Independent Treatment Centre	3	4	8	12	13	16	56
TOTAL	308	352	351	358	353	349	2,071
Switches	1	12	10	22	32	9	86

Source: NJR data[9]

Table 7.4, Hospitals by type

Geographical location of hospitals

Dummy variables are also constructed for the region where the hospital is located, and these correspond to the SHAs[174][38].

7.4 Structure of analysis, specification and estimators

The analysis is structured in three stages, to follow the sequence of hypotheses in section 7.2.

[37] Information on provider type is included in the HES data-set, but not in the NJR. Fortunately, it was possible to map provider type into the panel using information on linked patients. But this was not possible in all cases. This is one reason for the loss of some patients from the NJR in the panel. These are patients for whom it was impossible to identify hospital type. In order to avoid categories with too few observations, NHS treatment centre's and others are combined.

[38] SHAs were introduced in 2002 to manage the NHS locally and provide a link with the DoH, in 2006 these were refined from 28 regions to 10[164].

Stage 1 Choice of prosthesis type for the individual patient

This stage is designed to establish how far the choice of prosthesis for the individual patient is explained purely in terms of the patient's characteristics (hypothesis 1 above). If there is more or less uniform national decision-making, it would be expected that patient characteristics are the dominant factor, but if there are important differences between hospitals, then this establishes that further analysis at the hospital level is needed. In principle, this stage could be conducted across all the many different brands of prostheses. However, it turns out to be sufficient to consider only the very basic choice between the broad cemented and cementless types. As will be seen, patient characteristics provide only a very limited explanation of even this most aggregate of choices, and there is therefore nothing to be gained by disaggregating down to the level of the individual prostheses within these two broad types.

The model to be tested is[39]:

$$P(y_i = 1) = \Phi\ (\beta_0 + \beta_1 X_i + \beta_2 t_i) \qquad i - 1,...,n \qquad (1)$$

where

$\mathbf{y}_i$ =1 if patient i receieved a cementless prosthesis; 0 otherwise$\mathbf{X}_i$ is a vector characteristics of patient i (including age, gender, etc.), as above (age, gender etc**)**

$\mathbf{t_i}$ is a set of dummy variables indicating the year in which patient i received the prosthesis.

Φ (.) is the standard cumulative distribution function.

In this equation, and all following ones, t is measured using 5 year dummy variables (2004-8) with the omitted default 2003. This is to control for changes over time, which are of particular interest with respect to the impact of the introduction of PbR and the national tariff in 2005/6. Time is included in the form of year dummies, rather than as a continuous time trend, to allow for the possibility that trends may not be smooth or even monotonic.

This equation is fitted initially to the individual patient level data (145,860 patient observations); the dependent variable is measured in binary form, taking the value 1 if patient i receives a

[39] This is the probabilistic equivalent to the principals utility function described in section 7.2.2.

cementless prosthesis at time t, or 0 if cemented. Standard probit and bi-variate probit models are used to estimate the equation[175, 176][40]. Since more than one choice is being estimated, the bivariate probit model maybe more appropriate because it allows for the choices to be correlated due to unobserved characteristics. –

For comparability, the model is also tested on the hospital-panel form of the database[41]. In this case, the dependent variable becomes the proportion of patients implanted with cementless prostheses in hospital i at time t, and the X vector now denotes, for each characteristic, the average (e.g. for age), or the proportion (e.g. of females) across all patients in i at t. This is estimated using the random effects panel Tobit model[42], as the dependent variable is confined to the range 0 and 1 with a fairly large number of observations at the two bounds (see the next section for more discussion) [175, 176] .

All statistical analysis in this chapter is carried out using STATA SE version 11.

Stage 2 Hospital Specialisation

This stage explores the determinants of the extent to which individual hospitals specialize in their choices across brands and manufacturers. This is of relevance to heterogeneity within the NHS (hypothesis 2 above), whether specialisation decreases with the size of the hospital (hypothesis 3), whether there are systematic differences by region and hospital status (hypothesis 4),.and whether there are significant changes over time, especially with respect to the introduction of PbR (hypothesis 6).

The model fitted is:

$$y_{it} = \beta_o + \beta_1 \boldsymbol{X}_{it} + \beta_2 \boldsymbol{Z}_{it} + \beta_3 \boldsymbol{T} + u_{it} + v_i \qquad (2)$$

[40] Alternatively, the Logit model might equally be used. As explained by Cameron and Trivedi[165] , Probit and Logit invariably give qualitatively very similar results, and this is the case here. My preference is for the Probit, as it is based on a preference for the normal, as opposed to logistic assumption: the Central Limit theorem provides a strong reason for assuming normality, while the logistic assumption is more *ad hoc*[166].
[41] Here this is based on the probability equivlalent of the agents utlity function.
[42] The option to use fixed effects Tobit is unavailable in STATA.

where $\mathbf{X}_{it}$ is the vector of average patient characteristics as above (age, gender, side of surgery)
$\mathbf{Z}_{it}$ is a vector of hospital characteristics (hospital size, hospital type, region of hospital)
T is the vector of time dummies (2003 to 2008)
u_{it} is a conventional idiosyncratic disturbance term
v_i is a random time invariant hospital effect

The dependent variable, y_{it}, measures the extent to which hospital i at t specializes its purchases on a small number of, first, brands, and then manufacturers. Various different indices might be used to measure specialisation, but for consistency with how concentration is measured in chapter 6, the HHI is again used. In the context of brand specialisation, it is now referred to as SPB and is defined by: the sum of squared shares of the hospital's total purchases accounted for by brand j:

$$SPB_{it} = \sum_{j=1}^{N} s_{ijt}^2 \qquad (3)$$

where s_{ij} is the share of hospital i's total purchases accounted for by brand j where j=1…N, and N is the total number of possible brands. Higher values of SPB indicate more specialisation by the hospital. The upper limit is 1, which occurs if the hospital purchases only a single brand, the lower limit is (1/N), where it purchases equal amounts of all N different brands. Because each patient will receive both a cup and a stem, the value of the index for each hospital is calculated as the average of the cup and stem index values; in other words, SPB=1 would indicate that the hospital purchases only one brand of stem and one brand of cup.

The equivalent index of hospital specialisation across manufacturers is SPM_{it} , defined as in (3), but with j now denoting a manufacturer. This is estimated using the hospital panel data, again using the random effects panel Tobit model[43], as the dependent variable is confined to the range 0 and 1.

[43] The option to use fixed effects Tobit is unavailable in STATA.

Stage 3, Hospital market shares of the leading manufacturers: Stryker and Depuy

The third stage focuses on the extent to which the hospital concentrates its purchases on the two leading suppliers (the above specialisation indexes merely indicate the extent of its specialisation, regardless of the identities of the firms it buys from.). In this case, the dependent variables are the proportion of the hospital's purchases that are from Stryker, and the proportion of the hospital's purchases that are from Depuy, and, as a residual, the proportion of the hospital's purchases that are from all other suppliers. This is relevant to hypothesis 5 above. The previous chapter showed that both firms has about a third of the national market, and we now examine whether these high market shares occur roughly equally across all hospitals or whether it is because some hospitals have a strong preference for one manufacturer, while another hospital prefers another manufacturer.

In this stage, the estimated model can still be described by equation 2 above, but now the dependent variable denotes the share of hospital i's purchases that are from Stryker, Depuy or Others. In addition, these equations will also be estimated at a disaggregated level for each of cemented cup, cemented stem, cementless cup and cementless stem. These disaggregations should provide insights into the sources of the two firms' dominance in the different segments of the market, and this can often be equated with individual leading brands: for example, Stryker's dominance in cemented stems is largely accounted for by the Exeter, and Depuy's dominance in cementless stems, accounted for by the Corail (both of which have a 10A rating from ODEP).

In this stage, the panel Tobit model is again used because the dependent variables are bounded between 0 and 1. Since this stage involves a system of 3 equations in each case, this would seem to suggest using a Seemingly Unrelated regression model (SUR). However, as explained below in section 7.5.3, this turns out to be unnecessary[175].

7.5 Results

This section reports the results of estimating the various equations described in the previous section. To avoid undue repetition, the comments after each equation are relatively brief, emphasising signs and significance of individual coefficients, with a wider discussion of their implications for the research hypotheses of section 7.2 provided in section 7.6.

7.5.1 Stage 1 , The relative importance of patient characteristics

In order to assess the impact of as many characteristics of the patient as possible, it is necessary to first use NJR-HES linked data (HES contains information on more patient characteristics, see Tables 7.2). As explained in section 7.3, this NJR-HES linked data-set covers substantially fewer patients than the NJR panel data-set, due to the merging process, but when assessing patient characteristics, it provides a more comprehensive picture.

Table 7.5(i) reports the results of separate regressions for cup and stem components using the individual patient-level data-set (n=145,870). As explained earlier, the dependent variable is binary, (cemented =1 and cementless = 0) and this is estimated using first the probit model and secondly the bi-variate probit modell. Both equations are estimated with robust standard errors.

Explanatory variables	**Cementless cup**		**Cementless stem**	
Year (reference year 2003)	**Coefficient**	**Marginal effects**	**Coefficients**	**Marginal effects**
2004	0.17956***	0 .0715023	0.19329***	0.06938
2005	0.31951***	0.1269162	0.34702***	0.12676
2006	0.48612***	0.19179	0.48999***	0.18047
2007	0.52390***	0.2065461	0.58879***	0.21556
2008	0.64488***	0.2522949	0.76039***	0.28113
Age	-0.04085***	-0.0162158	-0.03311***	-0.01147
Right side	0.0048	0.0019046	0.01564*	0.00542
Lateral position	0.32105***	0.1247067	-0.06039***	0.02115
Female	-0.11863***	-0.0471263	-0.122176***	0.04265
Non white	-0.01229	-0.0048763	-0.0147	0.00507
Bilateral Indication	0.32086***	0.1272336	0.18317*	0.06626
Osteoarthritis diagnosis	0.06652***	0.0263258	0.22673***	0.07481
MIS used	0.38647***	0.1528167	0.64299***	0.24516
IGS used	-0.03934	-0.0155828	-0.16041**	-0.05307
Constant	2.01302***		1.16425 ***	
Log Likelihood	-90878		-82980	
Pseudo R^2	0.096		0.0825	
Number of observations	145,651	145,651	145,651	145,651

legend: +, significant at the 10% level (p<0.10) *, significant at the 5% level (p<0.05) **, significant at the 1% level (p<0.01) ***, significant at the 0.1% level (p<0.001)

Table 7.5(i) Determinants of the probability that a patient receives a cementless implant - probit model

Explanatory variables	Cementless cup		Cementless stem	
Year (reference year 2003)	**Coefficient**	**Robust standard error**	**Coefficients**	**Robust standard error**
2004	0.17971***	0.018185	0.20763***	0.020126
2005	0.31602***	0.017471	0.37004***	0.019242
2006	0.48355***	0.016979	0.52636***	0.018699
2007	0.52190***	0.016528	0.61610***	0.018233
2008	0.64488***	0.016621	0.78854***	0.018272
Age	-0.03955***	0.00036	-0.03330***	0.000356
Right side	0.00573	0.006876	0.01347*	0.007114
Lateral position	0.32592***	0.010357	-0.0304***	0.010419
Female	-0.12060***	0.007057	-0.12430***	0.007264
Non white	-0.01229	0.037975	-0.00710	0.035989
Bilateral Indication	0.31731***	0.091063	0.19078*	0.0865
Osteoarthritis diagnosis	0.08302***	0.01016	0.17322***	0.010521
MIS used	0.38067***	0.015911	0.63739***	0.015592
IGS used	-0.03080	0.056946	-0.16344***	0.061184
Constant	1.89899***	0.030618	1.17681 ***	0.03134
Number of observations	145,651	145,651	145,651	145,651

legend: +, significant at the 10% level (p<0.10) *, significant at the 5% level (p<0.05) **, significant at the 1% level (p<0.01) ***, significant at the 0.1% level (p<0.001)
Log-likelihood: -151854.89
the log-likelihood ratio test of rho = 0 is significant

Table 7.5(ii) Determinants of the probability that a patient receives a cementless implant - bivariate probit model

A test of the correlation between the error terms of the two equations reveals that it is positively significant (ρ =0.807, significant at the p<0.001 level), indicating that the results for the bivariate probit model in table 7.5(ii) are to be preferred.

For both Cup and Stem, these equations confirm a number of expected results. Cementless prostheses are more likely to be used for younger patients and for males, and they have become more likely in recent years (indicated by the successively increasing coefficients on the time dummies). There are also a number of other significant characteristics – MIS, IGS, bilateral surgery, diagnosis of OA, patient position and patient side (for stems).

Table 7.5 also shows the corresponding marginal effects at the mean. Generally in this chapter, a detailed discussion of the magnitudes of estimated coefficients is unnecessary, but in this case,

this serves to provide useful background. Patients are about 20% more likely to receive a cementless cup and stem in 2008 than 2003; females are 5% less likely to receive cementless than males; and as a patient ages, they are 1% less likely to receive a cementless for every year older they are.

However, the most important result is that this vector of patient characteristics only explains a small part of the overall variance in whether patients receive cemented or cementless prostheses: for both components, less than 10% as measured by the pseudo R squared. Of course, inclusion of all the patient level characteristics listed earlier in the theoretical model may have improved the fit, although I would expect that age and gender (both included) would be the most important.

Therefore, the equation is re-estimated now including fixed effects dummy variables for each individual hospital. This is shown in Table 7.5(iii): the signs and approximate magnitudes of all coefficients are largely unchanged from those in Table 7.5(i) (apart from right side of surgery, non-white and bilateral indication), but most importantly, there is a much improved overall fit of the model (from an R-squared of less than 10% to approximately 35%)[44]. Thus, much more of the overall variance of prosthesis type is explained by the hospitals themselves. This result almost certainly understates the impact of the inclusion of hospital dummies as necessarily, the model drops all patients in those hospitals which implanted only cemented or cementless prostheses (40 hospitals)[45]. Alternative measures of goodness of fit can also be estimated and a comparison of equation 7.5i and 7.5iii[46] using the 'fitstat' option in STATA confirms the ranking reported for the R-squared (for example, equation 7.5iii has a lower AIC and a higher McFadden's R-squared). It should be noted here that an R-squared of 35% as reported for equation 7.5iii, still leaves 65% of the variance unexplained, thus there is still a lot of unexplained variance in the model. This could be better explained were the data-set to include more of the variables specified in the theoretical model earlier in the chapter. Nontheless, an R-squared of 35% is still an acceptable level for probit analysis of this sort of data and so despite the data constraints, it justifies the conclusions and warrants the further analysis now described.

[44] The bi-variate probit model does not report an R-squared, so the R-squared result from the Probit model is reported for comparison.
[45] This is because, in these cases, the identity of the hospital is a perfect predictor of the choice of cementless or cemented.
[46] The fitstat option in STATA will not run on the bivariate probit model, so can only be estimated on equations 7.5i and 7.5iii (probit models).

Explanatory Variables	Cementless cup	Cementless stem
Year (reference year 2003)		
2004	0.22969***	0.24325***
2005	0.37698***	0.46165***
2006	0.61346***	0.73929***
2007	0.71587***	0.89546***
2008	0.88452***	1.10376***
Age	-0.05753***	-0.04617***
Right side	-0.00133	0.00969
Lateral position	0.48359***	0.18196***
Female	-0.19340***	-0.19636***
Non white	-0.16267***	-0.06334
Bilateral Indication	0.17765	0.08571
Osteoarthritis diagnosis	0.02052	0.21368***
MIS used	0.33037***	0.50611***
IGS used	-0.05809	-0.16090*
Constant	2.92923***	0.96042***
Pseudo R^2	-65992.58	-57942.39
Log Likelihood	0.3392	0.3518
Number of observations	145,651	145,651

legend: +, significant at the 10% level (p<0.10) *, significant at the 5% level (p<0.05) **, significant at the 1% level (p<0.01) ***, significant at the 0.1% level (p<0.001)

Table 7.5(iii) Determinants of the probability that a patient receives a cementless implant, the effects of including hospital fixed effects[47]

This establishes that patient characteristics alone provide a very limited explanation of which broad type of prosthesis is implanted for a given patient; it is clear that the choice significantly differs between hospitals, i.e. a given patient would receive a different type of prosthesis depending on the hospital of the surgery. Remembering that this result relates to just the most basic of all choices (cemented versus cementless), there is little to be gained from disaggregating further down to the specific brand level in terms of just patient characteristics. Thus I am able to reject Hypothesis 1 above.

Hospital-panel data-set

For this reason, all of the following analysis in this chapter switches to the hospital-panel dataset in order to focus on differences between hospitals, while controlling for differing patient mixes (in terms of averages for age, gender and side.)

[47] Equation 7.5iii is a probit model because the bivariate probit model will not run with the hospital fixed effects.

Before moving on to the other stages of the research, Table 7.6 reports the results of estimating the equivalent of the equations in Table 7.5 but now using the hospital panel, where the dependent variable is now the proportion of patients in a given hospital-year receiving a cementless (cup or stem) implant in a given year. In the first equation, only age, gender and side and the year dummies are included. The equation is estimated using a panel random effects Tobit model because the dependent variable is bounded between 0 and 1, with a significant number of observations at these bounds (100 observations at 0 and 32 observations at 1). Throughout this chapter the panel model used is random-effects Tobit because STATA has no command for a fixed-effects model[177][48].

The signs of estimated coefficients are generally consistent with those in Table 7.5 (apart from gender which is now not significant): the proportion of cementless increases steadily from year to year, and hospitals with older patients (on average) implant fewer cementless prostheses.

The second equation in Table 7.6 shows the results of adding the vector of hospital characteristics of region of location, hospital type and hospital size. In this, and all subsequent equations in this chapter, the omitted year dummy is 2003, the omitted region is East Anglia and the omitted hospital type is Foundation Trusts. This means that all significance levels on the included dummies refer to differences with respect to these defaults. However, interpretation of results often requires a wider range of hypotheses tests for significance between estimated coefficients (and not just relative to the defaults). This is done by conducting post-estimation Wald tests[49], and these will be discussed when relevant in the text. They are not shown in the Tables.

A log likelihood ratio test (LLR=60.5), confirms a highly significant ($prob>chi^2=0.000$) improvement in the fit of the model by including the hospital characteristics. As can be seen, all of the year dummies are significantly different from 2003, but Wald tests reveal that they are significant (at the 5% level) successive increases year on year in the number of cementless prostheses implanted. Second, patients in the South East (significant at the 10% level) and London (at the 5% level) are significantly more likely to receive a cementless prosthesis, while

[48] "there does not exist a sufficient statistic allowing the fixed effects to be conditioned out of the likelihood"[167] [165] p.631

[49] Throughout the chapter Wald tests are reported at the 5% level of significance

those in the North-West are the least likely. The remaining seven regions (including the default East Anglia) lie in between. Third, the only significant result by hospital type is that patients in NHS trust hospitals are more likely to be implanted with a cementless prosthesis than those in the other types of hospital. Finally, there is no significant tendency for larger hospitals to fit more cementless prostheses.

Variable	First equation	Marginal effects	Second equation	Marginal effects
Year (reference year 2003)				
2004	0.06133***	0.06133	0.05839***	0.04224
2005	0.10297***	0.10297	0.10068***	0.07353
2006	0.15855***	0.15855	0.1580***	0.11633
2007	0.20101***	0.20101	0.20161***	0.14897
2008	0.26900***	0.269	0.27087***	0.20022
Average Age of Patients	-0.01294***	-0.01294	-0.01273***	-0.00915
Proportion of Female Patients	0.03994	0.03994	0.04115	0.02926
Proportion of right sided surgery	0.05605	0.05605	0.05175	0.03679
Region (reference East Anglia)				
South East	-	-	0.08591+	0.06280
East Midlands	-	-	0.06486	0.04724
North West	-	-	-0.13288**	-0.08947
London	-	-	0.14717**	0.10895
Yorkshire	-	-	-0.01362	-0.00963
South West	-	-	0.01501	0.01073
South Central	-	-	-0.05361	-0.03718
West Midlands	-	-	-0.00919	-0.00651
North East	-	-	-0.00927	-0.00658
Hospital Type (reference FT)	-	-		
Trust	-	-	0.03077*	0.02169
Independent Sector	-	-	0.01562	0.01118
NHS Treatment Centre & other	-	-	-0.02627	-0.01843
Independent Treatment Centre	-	-	-0.07073	-0.04841
Hospital size	-	-	0.00832	0.00591
Hospital size quadratic	-	-	0.00012	0.00009
Constant	1.08112***	1.08112	1.0326462***	-
σ u	0.23844***	0.23844	0.22322	-
σ e	0.12160***	0.1216	0.12099	-
Log Likelihood	615.568	-	645.5922	-
Number of observations	1948	1948	1948	1948

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 7.6: Explaining cementless as a proportion of implants

7.5.2 Stage 2: Hospital Specialisation

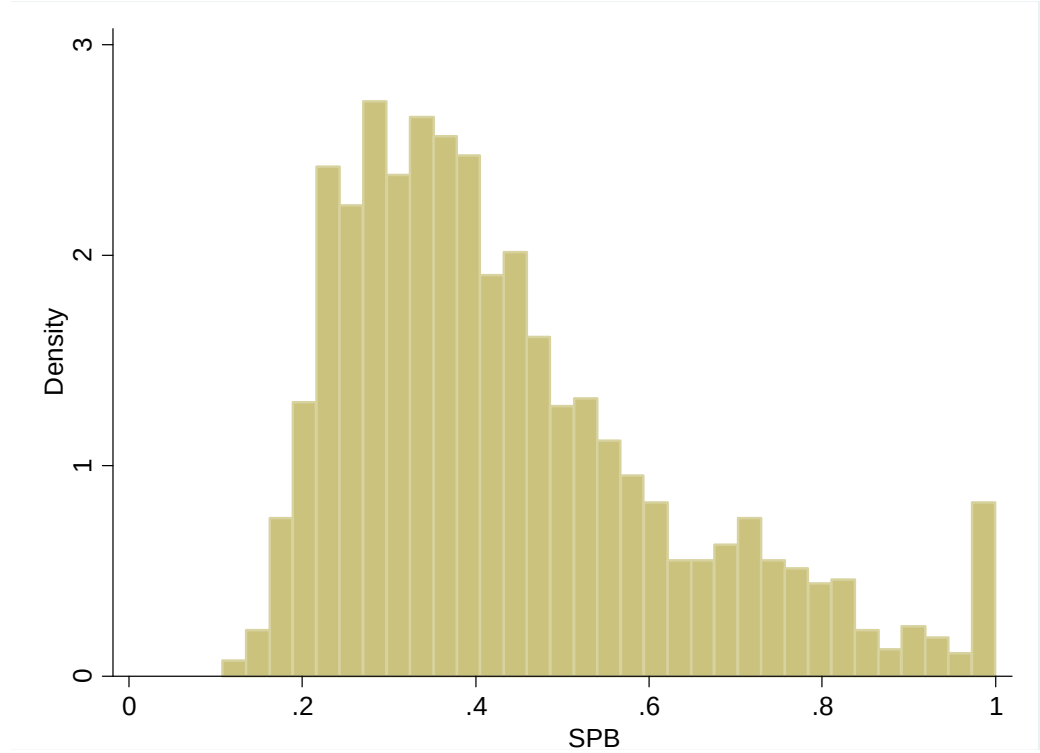
This stage now switches attention to identifying and explaining differences between hospitals in the extent to which they specialise their purchases - first by brand, and then by manufacturer.

Brand level

Figure 7.1, shows the distribution, pooled over 2003-8, of hospital specialisation by brand (SPB). It is roughly bi-modal, with one mode at around 0.25-0.4 and the other, to a lesser degree at 1. Thus, many hospitals implant the equivalent of 3 or 4 equal sized brands, but there are some hospitals (1.5%) which implant one brand only. The mean SPB is 0.44, indicating that the typical hospital implants the equivalent of just 2.27 brands. It should be remembered here that SPB is measured as the average of the separate SPBs for cups and stems. In other words, the typical hospitals implants the equivalent of 2.27 brands of each component.

Recalling Hypothesis 2 above, this hospital-level specialisation can be compared with specialisation at the national level. The dotted line in figure 7.1 shows the mean HHI at the national level for brands, which is 0.109 (from Chapter 6, section 6.5.4), this can now be interpreted as the degree of brand specialisation at the national level. In its numbers equivalent, NSPB=10, this is equivalent to 10 equal-sized brands. If each individual hospital was a small replica of the national market, then this would be the mean hospital SPB. However, as we see, the mean hospital level value (2.27) is much lower. In that sense, the typical hospital is 4 times more specialised by brand of prosthesis than the NHS as a whole, and as can be seen from the figure, virtually all hospitals are more specialised than the national level.. Thus I can reject Hypothesis 2.

Figure 7.1: SPB hospital specialisation at the brand level



* line depicts the national value of SPB (referred to as HHIB in chapter6, section 6.5.4)

Figure 7.2 provides a first test of hypothesis 3 - that smaller hospitals will tend to concentrate on implanting just a few brands, and will therefore be more specialised than larger hospitals,. From visual inspection of the figure, there does not appear to be a strong relationship between specialisation and hospital size: many smaller hospitals have a low SPB of between 0.2 and 0.4, while some of the larger hospitals are highly specialised. However, the hypothesis also raised the possibility that suppliers might target larger hospitals for exclusive contracts – in other words, an opposite effect. To assess whether there is any evidence on these two conflicting effects, a quadratic line is fitted to the data as follows:

$SPB_{it} = 0.504 - 0.0706^{***} SIZE_{it} + 0.00742^{***} SIZE_{it}^{2}$ R2 = 0.0143 (overall)

where $SIZE_{it}$ is the number of implants in hospital i at t.

Both hospital size and hospital size squared are significant at the 1% level, and the estimated coefficients imply a U shape with a turning point (minimum) at 475 patients. In fact there is only a small handful of hospitals larger than 475 (see Table 7.3), and so only the downward

sloping part of the U really applies. In other words, this is best summarised by concluding that there is a general significant tendency for brand specialisation to decrease as the size of hospital increases, but at a diminishing rate. However, the large scatter in this figure, and the low R^2, show that most of the variance is unexplained by hospital size. Thus, in general, Hypothesis 3 is weakly confirmed, but the relationship is weak. Nevertheless, in all remaining equations in this chapter, hospital size is included in quadratic form to examine whether this is also apparent in multivariate analysis.

Figure 7.2 Specialisation at the brand level and hospital scale

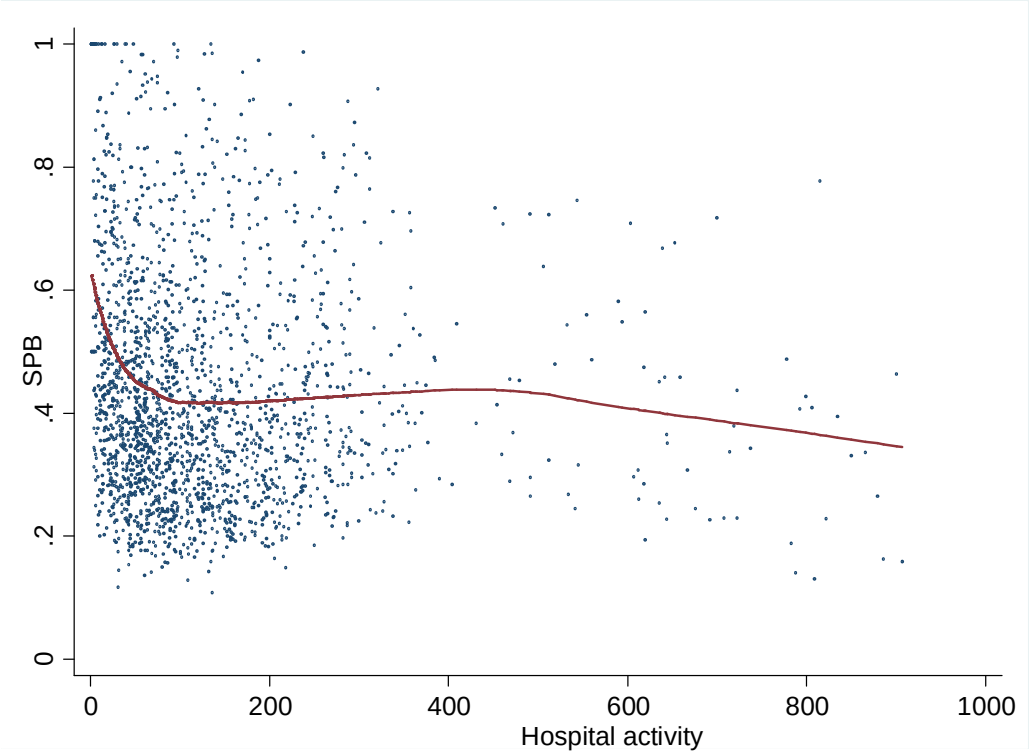


Table 7.7 shows that, on average, specialisation increased steadily 2004-7/8. The higher initial value for 2003 seems out of line with this trend, but if may reflect the fact that the NJR had much lower coverage in its first year (see section 2.2.1).

Year	Observations	Mean	Std Dev	Min	Max
2003	306	0.465	0.201	0.176	1
2004	341	0.423	0.192	0.128	1
2005	341	0.424	0.191	0.108	1
2006	350	0.432	0.19	0.154	1
2007	344	0.443	0.192	0.117	1
2008	336	0.440	0.188	0.130	1

Table 7.7 Hospital specialisation at the brand level: descriptive statistics by year

Building on these descriptive results, Table 7.8 reports the results of multivariate analysis of SPB. The choice of appropriate estimator requires some attention. By definition, the dependent variable is bounded between 0 and 1, and this suggests that a Tobit model might be most appropriate, although as shown in figure 7.1, there are no observations at the lower bound and only a small number (1.5%) at 1. Further investigation shows that SPB has a slight positive skew[50][178], which disappears if the variable is logged. As a consequence, three alternative forms of the model are reported in table 7.8: Model 1 reports the results of a panel 2-limit Tobit equation where specialisation is measured without logging. Model 2 reports a panel Tobit regression where SPB is logged (here there is only an upper limit of lnSPB=0); and model 3 is a standard random effects model (i.e. not Tobit), and in this case with robust standard errors. (Unfortunately, there is no option in STATA to estimate a panel Tobit model with robust standard errors.[179]) In principle, model 2 is the preferred form because it is most appropriate to use a Tobit model when the dependent variable is symmetric and broadly normal. However, model 3 benefits from having robust standard errors, and the non-use of the Tobit may not be too inappropriate given the small number of observations at the upper bound.

In fact, the results of the three models are very similar in terms of signs and magnitudes of the estimated coefficients: the only exception is the sign on the 'proportion of female patients', but this is insignificant in all three forms of the model.

[50] The Stata statistic for skewness is 1.01 when SPB is not logged and 0.12 when it is logged.
A normal distribution has a skewness of 0 and an indicator of a normal curve requires a skew between +2 and –2[168].
The Stata statistic for kurtosis is 3.51 for non-logged and 2.5 for the logged dependent variable. The reference value for normally distributed data is 3[165].

Explanatory variables	Model 1 Panel Tobit, SPB	Model 2 Panel Tobit, lnSPB	Model 3 Panel lnSPB
Year (Reference year 2003)			
2004	-0.01566	-0.04428*	-0.04489*
2005	-0.00575	-0.02711	-0.02822
2006	-0.00350	-0.01602	-0.01685
2007	0.01225	0.01769	0.01707
2008	0.00987	0.01743	0.01641
Average Age of Patients	0.00142	0.00242	0.00252
Proportion of Female Patients	-0.00637	0.00850	0.01406
Proportion of right sided surgery	0.02722	0.01035	0.00829
Region (reference East Anglia)			
South East	0.01992	0.06306	0.06271
East Midlands	0.04153	0.11372	0.11299
North West	0.00407	0.02700	0.02920
London	-0.06932*	-0.14772*	-0.14421+
Yorkshire	-0.00669	-0.01287	-0.01544
South West	0.05754	0.14914+	0.14744+
South Central	0.05408	0.15643+	0.15810+
West Midlands	0.02219	0.07535	0.07617
North East	-0.01386	-0.01586	-0.01269
Hospital Type (reference FT)			
Trust	-0.01808	-0.06136*	-0.06016
Independent Sector	-0.03911	-0.10246+	-.010692*
NHS Treatment Centre & other	-0.05127	-0.13426+	-0.13250
Independent Treatment Centre	0.04110	0.06251	0.06242
Hospital size	-0.07315***	-0.15143***	-0.14611***
Hospital size quadratic	0.00726***	0.01457***	0.01399***
Constant	0.40788***	-0.92132***	-0.93863***
σ u	0.14942***	0.33263***	-
σ e	0.11556***	0.25022***	-
Log Likelihood	956.03538	-528.46832	-
R-squared	-	-	0.05020
Number of observations	1948	1948	1948
Right censored observations	29	29	n.a
Left censored observations	0	0	n.a

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 7.8 Hospital specialisation by brand: panel equations

In terms of significance of coefficients, Model 1 performs worse than Models 2 and 3, suggesting that the logged form of SPB is to be preferred, but Models 2 and 3 provide almost identical results, suggesting that little has been gained from using the Tobit. For that reason, the following comments refer to the results of models 2 and 3.

The coefficients on the dummy time variables increase year on year, apart from the anomalous 2003. Putting 2003 aside as unexplained, I have conducted Wald tests for significant differences between the other years[51]. In particular, a Wald test identifies a significant difference (at the 7% level) between the three earlier years (2004-2006) on the one hand, and the two later years (2007/8)[52], indicating a structural break in specialisation at the brand level, with the increase coinciding with the introduction of PbR in 2006/7. None of the three variables controlling for average patient characteristics are significant, indicating that there is no effect of average patient characteristics on the hospital's specialisation (as one would suspect). The significant hospital size coefficients confirm the above result of a quadratic effect, indicating that larger hospitals tend to be less specialised in how many brands they implant up to some level, albeit at a declining rate[53]. Amongst the regions, Wald tests identify three significantly different groups. Five regions show no significant differences in specialisation relative to the reference region of East Anglia, and these six therefore form one group. However, London hospitals are significantly less specialised, while the South West and South Central are significantly (at the 10% level) more than the others. All hospital types have negative coefficients other than ISTCs, indicating that they are less specialised than FTs, but depending on which model is selected, there are differing levels of significance, although the IS hospitals are significant in both model2 and model3. A Wald test on NHS Trusts and ISTCs confirms a significant difference, although it should be noted that there are very few observations for ISTCs.

Manufacturer level (SPM)

The above analysis is now replicated for specialisation at the manufacturer level. Figure 7.3 shows that the distribution of SPM is also roughly bi-modal with modes at around 0.4/0.5 and at

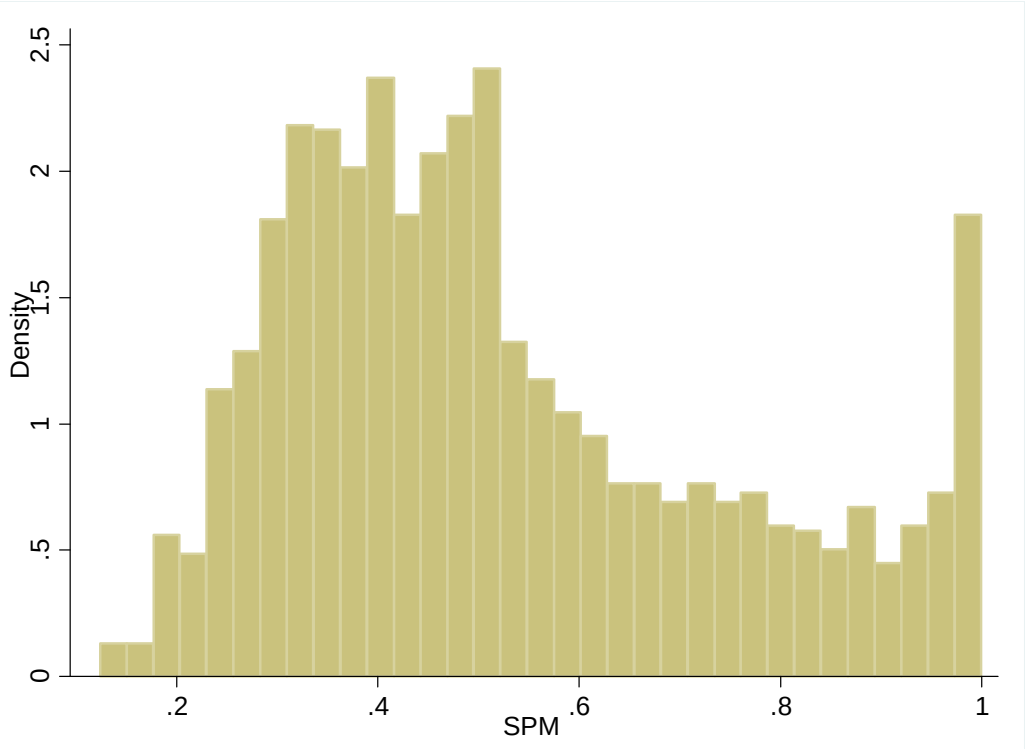
[51] Since the default year is 2003, the results of the significance tests shown in the table refer to differences between each later year and 2003, which is less interesting, given that the higher levels of specialisation in 2003 might be misleading given the much less complete coverage in 2003.

[52] A Wald test for a significant difference between 2006 and 2007 is significant at the 5% level (Chi^2= 5.37, Prob>Chi^2=0.02)

[53] The values of the coefficients in model 2 identify a turning point at hospital size = 0.151/(2*0.0146) = 520 which is virtually beyond the range of observed sample values.

1, and mean of 0.52. Thus, there is a minority of hospitals (3%) that buy exclusively from just one supplier, but more typically hospitals purchase from two or three different manufacturers. The dotted line in the figure recalls the HHI (re-defined as SPM here) at the national level of 0.26 (see Chapter 6, Table 6.4) - in its numbers equivalent form, NSPM=4, this indicates that at the national level the NHS buys from the equivalent of 4 equal sized firms. This figure would also apply at the individual hospital level if each hospital was a small replica of the national market, but in fact the typical hospital (with SPM=0.5) is twice as specialised as the NHS as a whole, in terms of the number of manufacturers it buys from.

Figure 7.3 Hospital specialisation at the manufacturer level



*Dotted line depicts the national SPM

Figure 7.4 plots the scatter between hospital size (activity) and specialisation at the manufacturer level. The pattern is very similar to that observed above for brands – a wide scatter but nevertheless with a significant U shaped quadratic regression line. In this case, the turning point is at Size = 408. So again specialisation declines again as hospital size increases. In other words, there is a tendency for larger hospitals to buy from more different manufacturers (possibly because larger hospitals have a greater number of surgeons operating which increases the preferences for different prostheses) but at a diminishing rate, and as there are 16 hospitals of

size greater than 400, there is some limited evidence that the very largest hospitals may be slightly more specialised. But again the fit is very low, and size alone leaves most of the variance unexplained.

Figure 7.4 Specialisation at the manufacturer level and hospital scale

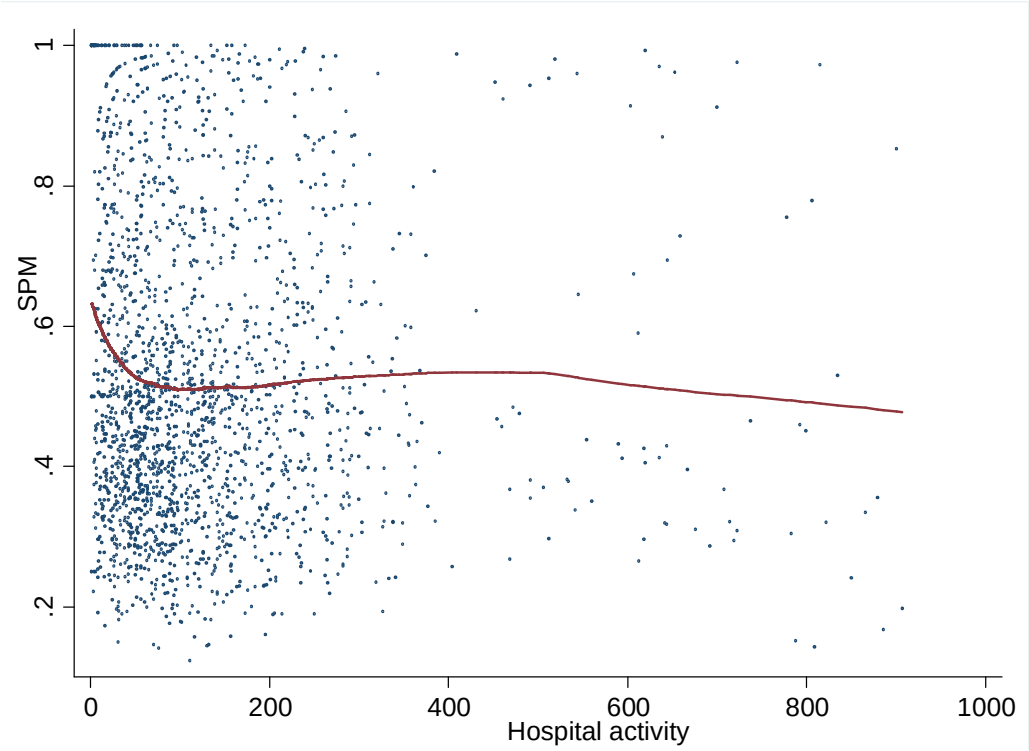


Table 7.9 shows that, on average, manufacturer specialisation also increased steadily 2004-7/8. Again, a higher value is observed for 2003 but this may again reflect the fact that the NJR had much lower coverage in its first year.

	Observations	**Mean**	**Std Dev**	**Min**	**Max**
2003	306	0.535	0.219	0.14	1
2004	341	0.509	0.213	0.157	1
2005	341	0.52	0.216	0.123	1
2006	350	0.514	0.214	0.144	1
2007	344	0.53	0.223	0.15	1
2008	336	0.536	0.222	0.142	1

Table 7.9 Hospital specialisation at the manufacturer level: descriptive statistics by year

Explanatory variables	Model 1 Panel Tobit, SPM	Model 2 Panel Tobit, lnSPM	Model 3 Panel lnSPM
Year (Reference year 2003)			
2004	-0.00611	-0.01705	-0.01870
2005	0.00895	0.00864	0.00576
2006	0.00371	0.00333	0.00037
2007	0.02187+	0.03244	0.03185
2008	0.02799*	0.04723*	0.04576
Average Age of Patients	-0.00026	-0.00143	-0.00168
Proportion of Female Patients	0.02014	0.06424	0.07175
Proportion of right sided surgery	-0.05584	-0.09185	-0.07610
Region (reference East Anglia)			
South East	0.08118*	0.16982*	0.16681*
East Midlands	0.06619	0.13581	0.13509
North West	0.06734+	0.14396*	0.14468*
London	-0.08402*	-0.16858*	-0.16175*
Yorkshire	0.11514**	0.23049**	0.22566**
South West	0.01174	0.02656	0.02774
South Central	0.00058	0.05092	0.05505
West Midlands	0.04699	0.12171	0.12245+
North East	0.03714	0.07148	0.07576
Hospital Type (reference FT)			
Trust	-0.01752**	-0.04554+	-0.04551
Independent Sector	-0.07934	-0.16199**	-0.16562**
NHS Treatment Centre & other	-0.04617	-0.09440	-0.09161
Independent Treatment Centre	0.03376***	0.02636	0.02184
Hospital size	-0.04999***	-0.08638***	-0.07981***
Hospital size quadratic	0.00530***	0.00809**	0.00747*
Constant	0.57764	-0.58681**	-0.59005*
σ u	0.17109	0.33089	-
σ e	0.12860	0.24413	-
Log Likelihood	714.003	-502.058	-
R-squared	-	-	0.0608
Number of observations	1948	1948	1948
Right censored observations	49	49	n.a
Left censored observations	0	0	n.a

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 7.10 Hospital specialisation by manufacturer: panel equations

Table 7.10 shows the results from fitting the same three models for SPM. In this case, there is less evidence of non-normality of the dependent variable[54], but for comparability I also report the results for the logged dependent variable. Again, by using the non-Tobit panel model I can report robust standard errors. Broadly speaking all three versions of the model report the same results and for that reason again, I will focus on models two and three. The year 2008 is significant at the 5% level for model1 and model2, with 2007 at the 6% level in model1, suggesting that there may again be an effect of PbR, but later and less pronounced. None of the patient characteristics variables are significant. In terms of the regional dummies, hospitals in the South East, North West and Yorkshire are significantly more specialised than all of the others, while London is significantly less specialised, confirmed by Wald tests for a significant difference. NHS Trusts and IS hospitals are less specialised than FTs with differing levels of significance depending on which model is chosen, as with SPB, ISTCs are more specialised and significant in model1. Hospital size and hospital size quadratic are again negatively and positively highly significant, confirming a U shaped effect.

7.5.3 Stage 3: Hospital market shares of Stryker and Depuy

The previous stage examined how far hospitals are specialised in their choice of manufacturers and brands; this stage goes further by examining how far this specialisation involved the two dominant manufacturers, Stryker and Depuy. The previous chapter showed that their national market shares are both roughly one third, and here I examine whether and how these shares vary between hospitals and over time.

	Stryker	**Depuy**	**Other**
Stryker	1		
Depuy	-0.447*	1	
Other	-0.529*	-0.521*	1

Table 7.11 Pairwise correlations between manufacturers' market shares
* significant at the p<0.001 level

[54] Skewness = 0.68 non-logged, -0.08 logged. Kurtosis=2.51 non-logged and 2.46 logged.

Preliminary descriptive analysis of the market shares data shows two key points. First the three pairwise correlations between Stryker's, Depuy's and the others' shares are reported in table 7.11. This shows that all three are substitutes for each other, for example, where a hospital uses relatively more Stryker prostheses, it uses relatively fewer Depuy, and relatively fewer Others. These negative correlations are not surprising - if the share of one manufacturer increases, the joint shares of all the others must decrease by the same amount. However, the sizes of the three correlations are almost identical, suggesting that increases in Stryker's shares tend to affect both Depuy and Others roughly equally, and vice versa. In this sense, all three seem to be equal substitutes for each other.

Second, Table 7.12 shows how their market shares vary across hospitals. Thus, for example, 26% of hospitals do not buy Stryker at all, and another 8% buy very few (less than 5%) of their prostheses from Stryker; on the other hand, 30% buy more than half of their prostheses from Stryker. The proportions are similar for Depuy. The figures for the residual show that in 50% of hospitals 'Others' account for fewer than 50% - and therefore Stryker and Depuy combined account for more than 50% of prostheses.

	Percentage of hospitals in which share (s) is:		
Market shares (%)	**Stryker**	**Depuy**	**Others**
s = 0	26	19	4
0 <s <5	8	10	16
5 ≤s< 10	4	6	8
10 ≤s< 20	9	9	13
20 ≤s <30	9	9	10
30 ≤s <40	9	9	10
40 ≤s <50	8	9	7
50 ≤s <60	10	9	7
60 ≤s <70	5	6	5
70 ≤s <80	4	4	4
80 ≤s <90	4	5	4
90 ≤s <100	4	4	5
s=100	2	2	6

Table 7.12 Manufacturers' shares in individual hospitals

Against this background, Table 7.13 shows the results of fitting a panel 2-limit Tobit model, in order to identify how far these observed variations between hospitals can be explained by the

hospitals' characteristics. In this case, normality tests of the dependent variables suggest non-significant skewness or deviations from normal kurtosis. However, a relatively large number of observations lie at the lower bound (see table 7.12). Thus, I report only the model for non-logged panel 2-level Tobit. At first sight it would seem appropriate to estimate these equations as a system, using SUR (seemingly unrelated regression). This is because (unobserved) variables which impact on one firm's market share should also affect (in an opposite direction) the other(s). In that case disturbances will be correlated across the three equations and statistical efficiency could be improved by incorporating that into the estimator. However, as explained by Cameron and Trivedi [176], where the set of explanatory variables is identical in all equations (as they are here) there is no gain in efficiency and the results using SUR are identical to those not using SUR , thus table 7.13 reports the results of fitting 2-limit Tobit models

Stryker

From Table 7.13, there is a clear year by year trend for Stryker's market share to be higher between 2005 and 2007 as shown by the positive and significant year dummy coefficients. Wald tests on 2004 and 2005,6 and 7 confirm this as a significant difference. There are also interesting regional differences: patients from the East Midlands are significantly less likely to receive a Stryker prosthesis, while patients in both the South West and South Central are more likely, confirmed by Wald tests of a significant difference. This result is particularly interesting when related to table 6.3, where I reported that Stryker's headquarters are in Berkshire, South England.On the other hand, there does not appear to be a hospital type effect. The coefficient for both hospital size and hospital size squared are significant at the 5 and 10% level, indicating a significant inverted U shape - as hospital size increases, generally they implant relatively more Strykers, but this effect tails off and then reverses once size exceeds 418[55], i.e. very large hospitals tend to become less reliant on Stryker than large hospitals. This is the opposite to the earlier finding (tables 7.8 and 7.10) that larger hospitals tend to be less specialised, except at very high levels.

Depuy

A number of the results in the Depuy equation are in direct contrast to the results for Stryker. The second column in Table 7.13 shows a significant decline in Depuy's market share in the

[55] 418 is the turning point in the quadratic with the coefficients here for size and size squared (4.18 hundreds.)

years between 2005 and 2008 compared to 2003 and 2004, and a Wald test points to a significant break between 2004 and 2005. Regionally, Wald tests show that patients in the North West, Yorkshire and the West Midlands are significantly more likely, and those from London less likely, to receive a Depuy prosthesis than those living in the other regions. As found above for Stryker, these regions of strength for Depuy are geographically close to the firm's headquarters – in this case, Leeds in Yorkshire (see Table 6.3.). By type of hospital, the only significant difference is that patients treated in an NHS Trust hospital are less likely to receive a Depuy prosthesis than those treated in FTs and other hospitals. Finally, unlike for Stryker, there is no evidence of a hospital size effect.

Others

The third column in the Table, for Others, is included mainly for completeness, but it does include some additional findings of interest. First, there is no evidence of a time trend in all other firms' combined market share, which implies that the gains for Stryker in the early years are largely accounted for by the losses for Depuy. Here, most of the regional coefficients are negatively significant, but these are all relative to the default, East Anglia. The three exceptions are the South East, East Midlands and London. The most appropriate way to interpret this, which is confirmed as significant by Wald tests, is that the 'other' manufacturers record significantly higher market shares in the three regions in the East of England: East Anglia, East Midland and the South East, compared to the rest of England; and that their share is even higher in London. The potential causes of this strong geographical pattern deserve further investigation in future research. Amongst the hospital types, patients treated in an NHS trust or IS hospitals are more likely to receive an 'other' prosthesis than those treated in an FT or other hospitals.

Explanatory variables	Model 1 Stryker	Model 2 Depuy	Model 3 Other
Year (Reference year 2003)			
2004	0.00508	-0.04024	0.02260+
2005	0.07509*	-0.07622*	0.00909
2006	0.08064**	-0.10340***	0.02022
2007	0.09997**	-0.10492***	0.01328
2008	0.08748**	-0.08638**	0.00899
Average Age of Patients	0.00599	-0.00172	-0.00178
Proportion of Female Patients	-0.00362	0.02288	-0.01593
Proportion of right sided surgery	-0.27474*	0.03348	0.10853*
Region (reference East Anglia)			
South East	0.07407	-0.09532	0.02265
East Midlands	-0.27882	0.21236	0.00142
North West	0.01216	0.52842***	-0.21027***
London	-0.12525	-0.26650*	0.14790**
Yorkshire	-0.08246	0.47754***	-0.16993**
South West	0.37289*	0.16110	-0.17039**
South Central	0.55038***	-0.04059	-0.19065**
West Midlands	0.12420	0.44374***	-0.20241***
North East	0.20365	0.10179	-0.11272
Hospital Type (reference FT)			
Trust	0.00870	-0.10520**	0.03825*
Independent Sector	-0.02834	-0.08834	0.07580*
NHS Treatment Centre & other	-0.00074	-0.05386	0.00702
Independent Treatment Centre	-0.09641	-0.14277	0.05059
Hospital size	0.05611*	0.01581	-0.00340
Hospital size quadratic	-0.00696*	-0.00001	-0.00008***
Constant	0.02206	0.65118*	0.48611
σ u	0.66859	0.58378	0.25899
σ e	0.32315	0.31923	0.14895
Log Likelihood	-1160.49	-1136.08	315.12573
Number of observations	1948	1585	1758
Right censored observations	18	17	103
Left censored observations	468	346	87

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 7.13 Manufacturer's hospital share: panel equations

Stryker's and Depuy's hospital market shares, disaggregated by segment

Finally, the above market share equations are re-estimated separately for each of the four broad types: cemented cup and stem and cementless cup and stem. In each of these four segments, the two major firms tend to have one or perhaps two leading brands, and so these equations can almost be interpreted as an analysis of the leading brand shares. For example, Stryker's dominance in cemented stems is largely accounted for by the Exeter V40. Again, all equations are estimated using Panel Tobit with random effects.

To avoid unnecessary duplication, the following discussion focuses only on the most important findings, referenced to Appendix 8, which reports the equations in full. Figures 6.2 from the previous chapter also help to illustrate these results.

Cemented cup (table 1, Appendix 8)

Main brands: Stryker/ Contemporary[56]; Depuy/Elite and Charnley; 'Other'/CPT

In cemented cups, there is a steady significant yearly increase in Stryker's market share and a corresponding negative time trend in Depuy's cemented cup shares. Patients treated in hospitals in the North West, Yorkshire and the West Midlands are more likely to have a Depuy and this seems to be mainly at the expense of 'other' manufacturers, but the opposite is true for London. This regional result is confirmed by significant Wald tests for the two regional groups for Depuy. The only significant result by hospital types is that NHS trust hospital patients are significantly less likely to receive a prosthesis from Depuy. There are no significant regional or type differences for Stryker, which suggests that the increasing dominance of its Contemporary prosthesis is widespread across all regions and types of hospital.

Cemented stem (table 2, Appendix 8)

Main brands: Stryker/ Exeter V40; Depuy/Charnley

As with cemented cups, there is a steady significant yearly increase in Stryker's market share of cemented stems at the expense of Depuy. There is also a regional pattern, with patients in the

[56] In each category the firm's leading brand in that category is shown.

South West and South Central being more likely to receive a Stryker and patients in the North West, Yorkshire and the West Midlands, more likely to have a Depuy[57] (although this clustering may also occur for historical reasons, as discussed in footnote. This significant difference between regional groupings is confirmed by a Wald test. This is particularly interesting when referring back to table 6.3, which shows that Stryker is based in the South and Depuy in Yorkshire. However, patients treated in an NHS Trust hospital and an ISTC are the least likely to receive a Stryker. No hospital type coefficients were significant for Depuy cemented stem. These results capture the very dominant role of Stryker's Exeter prosthesis, which has gained mainly at the expense of Depuy's Charnley and C-Stem.

Cementless Cup (table 3, Appendix 8)

Main brands: Stryker/Trident; Depuy/Pinnacle; 'other'/CSF

In the cementless components, the time trends move in the opposite direction. There is a positive time trend in Depuy's share of cementless cups accounted from 2005 onwards, and this seems to have been at the expense of other manufacturers, which shows a significant negative time trend for this period. Again, hospitals in the South East and South Central are significantly more likely to use Stryker and hospitals in the North West and Yorkshire and more likely to use Depuy, again confirmed by Wald tests. Interestingly, hospitals with older patients are less likely to use Depuy cementless cups. There is a weak indication that larger hospitals are more likely to implant a Depuy cementless cup (hospital size is weakly significant at the 7% level). These results reflect the increasingly dominant role of Depuy's Pinnacle, mainly at the expense of JRI's CSF, rather than Stryker's brands.

Cementless stem (table 4, Appendix 8)

Main brands: Stryker no one leading brand; Depuy/Corail; 'other'/Furlong

There is a similar positive time trend in the share of Depuy's cementless stems but in this case a significant negative trend in Stryker's share; Others' share also declines significantly particularly after 2006. Again, patients living in the South East and South Central are more likely to receive a Stryker cementless stem, in addition to patients treated in the West Midlands, confirmed by a

[57] This clustering may occur due to the historical reasons discussed in the previous footnote number 30.

Wald test. Patients living in the North West and South West are most likely to receive a Depuy, but those in South Central significantly less likely. These results are again confirmed by a Wald test. Patients in NHS Trusts and ISTCs are more likely to receive a Stryker cementless stem than in other hospital types. The significant negative results for Depuy in trusts is superficially misleading – remembering that Foundation Trusts are the default dummy, this should be interpreted as evidence of a particularly strong Depuy share in Foundation Trusts. The positive and weakly significant coefficient for hospital size in the Depuy equation suggests a weak trend for larger hospitals to implant a Depuy, but the negative quadratic coefficient, suggests that this tails off at very large sizes. Again, these results largely reflect the rapid growth in Depuy's Corail prosthesis. This is largely at the expense of JRI's Furlong. However, in this sector, there are important differences between regions, so that, although Stryker is mainly in second place, its share is stronger in regions close to its main base.

7.6 Discussion

I now discuss the implications of these results for the original hypotheses in section 7.2:

1. The choice of prosthesis is largely determined by the characteristics of the patient - rejected.

At the individual patient level, patient characteristics play only a small part in explaining even the most basic of choices – whether to fit a cementless or cemented prosthesis. It appears that a much larger role is identified by including hospital dummies, shown by the increased Pseudo R-squared increase from approximately 10% to 35% with the inclusion of hospital fixed effects. This justifies the decision to focus the subsequent analysis on explaining differences between hospitals. There is some evidence that the age of the patient determines which 'type' of prosthesis is implanted i.e. older patients are more likely to receive a cemented prosthesis (see age coefficient, tables 7.4 and 7.6), but this will not account for much of the differences between hospitals unless their patient mixes vary dramatically. Of course, it is worth noting that there may be some case mix differences by region, for example, there tends to be more elderly patients in rural or coastal regions, although the main evidence from the analysis is that there are large differences between hospitals.

2. The NHS is a homogenous entity- rejected

The data reveal considerable differences between hospitals in the extent of their specialisation, and the extent to which they purchase from one or the other of the main suppliers. Some hospitals buy most if not all of their prostheses from one or two suppliers, and only a few show diversification of purchases similar to the national market shares (figures 7.1 and 7.3.) In other words the NHS is not made up of identical clone hospitals – much seems to vary from hospital to hospital.

3. Larger hospitals are less specialised – weakly accepted, subject to a qualification

There is a significant broad tendency for larger hospitals to be less specialised – both in the number of different prostheses and the number of suppliers it uses. Further work is required to establish how far this might be explained by the fact that larger hospitals employ more surgeons, and that surgeons differ in their choices, even within the same hospital. This is important for establishing whether choices are made at the hospital level, or at the individual surgeon within the hospital level. However, in both cases, this is reversed at high values of hospital sizes – the very largest hospitals are more specialised (see below.)

4. There are predictable differences within the NHS between broad segments-largely rejected

Table 7.6 reports that patients in NHS Trusts are more likely to receive a cementless prosthesis than those treated in a FT, which could indicate that the NHS reforms (discussed earlier) have had a direct impact on hospitals with financial autonomy (FTs) in that they are choosing to implant the cheaper type of prosthesis. The regression on specialisation of manufacturers at the hospital level does not provide strong results, other than that IS hospitals appear to be less specialised in their purchasing. The main finding on the market share regressions (aggregated and disaggregated - Appendix 8, tables 1-12) is that, in general, NHS Trusts are more likely to implant a Stryker prosthesis, particularly a cemented and cementless stem, compared to patients in FTs. The reasons why deserve further research.

5. Part of the observed variation in choice is the result of systematic behaviour of the manufacturers – some suggestive evidence.

There are two main results that might point to market sharing understanding (even if only implicit) between Depuy and Stryker. The first is the general tendency for Stryker to become increasingly dominant in the cemented category and Depuy in the cementless. This became clear in the previous chapter, but it also shows through in Tables 7.13-15 (and Appendix 8) of this chapter. The second is the results on the regional dummy variables. There are some significant regional differences in terms of specialisation (Tables 7.8 and 7.10): hospitals in the South-East and Yorkshire are more specialised than average in their choice of manufacturers, while hospitals in London are less specialised in their choices of manufacturer and brands. But the most interesting results relate to the aggregated and disaggregated market share regressions (Tables 7.13-7.15 and the Appendix). These indicate that Stryker is particularly strong in the Southern regions (South-West and South Central) where it also has its headquarters in Berkshire (table 6.3), and this is at the expense of other manufacturers. Depuy enjoys greater market dominance in the Northern regions where it has its headquarters in Leeds (table 6.3) (Yorkshire, the North-West and also West Midlands), in each case at the expense of the other manufacturers. On the other hand, 'other' manufacturers have larger shares in London, at the expense of both Depuy and Stryker. These results are also supported by the disaggregated regressions in Appendix 8. Despite the fact that the general use of cemented prostheses is declining in favour of cementless (chapter 6), Stryker remain a dominant presence in the aggregate market (cemented plus cementless), particularly because they are increasing their market share in cemented so quickly, with the Exeter V40. In contrast, Depuy's presence in the cemented market has declined year on year, but they have increased their dominance in the cementless sector.

6. A potential impact of PbR – some evidence

The inclusion of year dummies picks up any simple shifts over time. Most interesting would be any apparent shifts in or around 2007 which would imply that the PbR policy as had some effect on choice. In fact, there do seem to have been various changes over time. First, Tables 7.4 and 7.6 both confirm the general trend over all the years towards increasing use of cementless prostheses (not related to PbR.) More relevant is the results in Tables 7.10 of a positive significant difference between 2007 and 2008 and the earlier years. This indicates that hospitals became more specialised in their choices of manufacturer from 2007 onwards, and this could possibly be a consequence of the introduction of the PbR policy i.e. leading to more

specialisation in purchasing at the hospital level. From Table 7.13 there also appears to have been a shift from 2006 onwards in the use of Stryker's prostheses, and a shift in 2005-8 away from Depuy. In turn, this appears to have been due mainly towards implanting more Stryker *cemented* prostheses, detected in the disaggregated hospital market shares regression. (shown by the tables in Appendix 8). It remains for future research to establish whether these changes are the result of PbR.

7.7 Conclusions

The determinants of choice of hip prostheses have been explored at both the patient and the hospital level. The results from this chapter and the previous one, reveal a purchaser with potentially significant buyer power (NHS) which is not currently being exploited and a supply side oligopoly. On the supply side two dominant manufacturers are identified: Stryker and Depuy: they account for two thirds of the aggregate market. There is also evidence of increasing dominance by Stryker in the cemented and Depuy in the cementless market, both of which ***could*** indicate some form of mutual forbearance between the two main manufacturers. However, caution should be taken here, as these results could also be simply a result of the manufacturers responding competitively - for example, Stryker may have been more successful in innovating in cemented prostheses, but Depuy more successful in cementless with both promoting their prostheses to the NHS on the basis of long-term outcome evidence.

More of the prosthesis choice appears to be explained by hospital characteristics as opposed to patient characteristics, contrary to what one might have anticipated and hypothesized in section 7.2. There is also some evidence of increasing specialisation at the hospital level over time, particularly in 2007/8, which is around the time one might expect to see some evidence of the impact of the introduction of PbR. There is also an indication that NHS Trusts are more likely to implant the more expensive cementless prostheses than FTs which could be a consequence of the financial autonomy that FTs enjoy and their response to managing their own budget. The results also provide some evidence of larger hospitals being more diversified in their choices.

The most intriguing result is the indication of a split in manufacturers strengths in the different parts of the country (Stryker in the South and South West, Depuy in the North West), and a

similar split in each broad type of prosthesis used in these parts of the country (Stryker are the main supplier of cemented type and Depuy, supplying cementless prostheses), this is also consistent with market sharing on behalf of the manufacturers . However, the association between the location of the supplier's innovating hospital and the accompanying regional surgical training rotation[58] may also account for some of this observed regional market sharing.

The analysis in this chapter might be further enhanced by further disaggregation of explanatory variables: within the hospital type variables, an indication of whether the NHS Trust or FT are teaching hospitals (see [180]) and the regional dummies disaggregated further beyond the Strategic Health Authority categories to allow for further variations. It might also be worth considering any other potential patient characteristic explanatory variables, although the constraints of the current data-set restrict further analysis at this level.

There are still a number of outstanding areas that require further work. Specifically, consideration of the relationship between the number of consultants carrying out THR surgery per hospital. For example, some smaller hospitals may only employ one surgeon to carry out all hip surgery, and that surgeon may only implant one or possibly two brands of cup and stem, given the learning curve associated with surgery. On the other hand, if a hospital employs a large number of surgeons to carry out hip surgery, this may lead to diversification of purchasing prosthesis brands and from manufacturers (i.e. differing surgical preferences and historical preference). Analysis at this level would allow for further exploration of the impact of hospital size on purchasing patterns.

Further analysis of the impact of the PbR policy on prosthesis purchasing is also warranted. This has not been explored in detail in this chapter and there are only hints from the time trends that there may have been some change in purchasing behaviour. Including a variable indicating where a switch from NHS Trust to FT has occurred and when, would provide further opportunity of analysis at this level (this information is available from the Monitor website [173]). Beyond this, the impact of clinical guidelines on the choice of prosthesis implanted also warrants greater consideration, such as the NICE guidance of 2000. This has not been pursued in this thesis because data is not available for the time period immediately following the NICE guidance[181]

[58] For example: the Exeter hip is centered on Exeter, Devon and the Charnley hip is centered on the Wrightington Hospital, Lancashire.

and other relevant clinical guidelines have not been identified which are directly relevant to this issue.

The findings from this and the previous chapter report what choices are currently being made and have identified that choice of prosthesis is not explained by the characteristics of the patients alone. Convention may lead us to believe that the NHS is a homogenous single buyer in the health care market, however, the reverse appears to be the case. In the case of hip prostheses (and other consumables, see the NAO report[135]), the NHS is making choices and purchasing at the disaggregated hospital level, thus severely restricting the potential to exploit its buyer power (supported by the findings in the NAO report). Given the lack of homogeneity in purchasing patterns in the NHS, the finding of a highly concentrated, oligopolistic seller side of the market, suggests concern if the manufacturers are able to exploit their seller power. Potential evidence of this has been shown empirically in this chapter, particularly in the potential regional market sharing between the two main players: Stryker and Depuy[59].

.

The findings in this and the previous chapter regarding what choice of hip prostheses are being made, may indicate inefficiencies in the current NHS policy on purchasing of hip prostheses, it also warrants deeper investigation of the supply side in order to establish whether true anti-competitive behavior exits. However, these findings are speculative at this stage and warrant a deeper understanding regarding purchasing behaviour i.e. who is making the choices and why they are making them. Future stages in the work will be to explore the nature of the decisions being made by collection of further, perhaps more qualitative data, to provide greater understanding of the nature of this transaction.

[59] Although it is noted that there are other explanatory factors which exist and may explain regional clustering of hospitals purchases

Chapter 8, Conclusions

This thesis has explored the choice between alternative THR prostheses from two different perspectives:

a) The value of survival curve analysis in guiding how choices between prostheses *should* be made, bearing in mind that information is scarce on the long-term survival rates of different prostheses. This was explored in chapters 2 to 5.

b). The choices that are *actually* being made, and investigating what this reveals about the buyer-seller relationship between the NHS and the large multinational firms who manufacture the prostheses. This was explored in chapter 6 and 7.

8.1 Main findings

Chapter 2 provided a background perspective for the thesis as a whole describing the technological development of hip prostheses and the evolution of the supplying industry over the last half century and, more recently, the pivotal establishment of the NJR for England and Wales in 2003.

Chapters 3 to 5 explored the first of the thesis objectives – the analysis of survival rates for different prostheses. Chapter 3 began by using the published NJR data to establish what conclusions can be drawn on the relatively short-term survival rates now observable from the NJR annual reports. It confirmed and updated the main finding of the only other paper to date[54] which has used NJR data to explore survival rates - cemented prostheses outperform cementless in terms of 5 year as well as 3 year survival rates, despite the fact that cemented prostheses have lost market shares to cementless. The depth of the analysis was then extended by examining revision rates for individual brands of prosthesis. The main findings were as follows: there are quite large variations in revision rates between different prostheses within each of the broad types; the three year revision rate is an imperfect predictor of the 5 year rate; there is no apparent tendency for the prostheses with the lowest revision rates to be the most commonly

implanted in the NHS; and the ODEP classifications do not appear to be closely related to the emerging evidence on revision rates from the NJR.

Chapter 4 reviewed the published economic evaluation literature on the alternative prostheses. This assessed the completeness of the existing evidence base for resource use, costs and cost-effectiveness, in order to determine whether the available evidence can inform current resource allocation decisions in the UK NHS. It clearly established that there is very limited data on the long term survival of hip prostheses in the public domain, concluding that more clinical trials, including head to head comparisons of hip prostheses with long term follow up are required. Relating these findings back to the first thesis objective, it is clear that lack of long-term survival rates of different prostheses is a key barrier to making informed choices between alternative hip prostheses. Chapter 5 then attempted to address this gap by exploring for a well known case study, whether extrapolating survival curves over a lifetime horizon using short term data can be used to compensate for the lack of long term data. Unfortunately, the findings were that prosthesis survival rates estimated using a short time series data cannot be reliably extrapolated. This suggests that reliable cost-effectiveness decisions deciding between alternative hip prostheses may be difficult – they cannot be made relying exclusively on currently observed published evidence on survival over a relatively short time period, or by methods which extrapolate short term survival rates into the future. However, and more positively, as the NJR accumulates a longer time series of data on survival rates, it will become an increasingly valuable resource for enabling robust decisions on the cost-effectiveness of alternative prostheses.

Chapters 6 and 7 turned to the second main objective of the thesis, exploring what choices are *actually* being made, specifically, how the structure of the main buyer (the NHS) interacts with the structure and nature of competition within the supplying industry (the manufacturers of THR) to influence choice at the hospital level.

Chapter 6 introduced, and in some instances measured the theoretical concepts used in Industrial Organisation, to provide the background and hypotheses for the econometric estimation of the various models in chapter 7. The descriptive statistics indicated that potentially a dominant duopoly exists in the UK market (Stryker and Depuy). These two manufacturers appear to be maintaining their share of the market over time by consolidating the shares of their established

brands (which report strong clinical evidence) and by growing the market shares of their newer brands (e.g. Corail and Pinnacle by Depuy). The chapter revealed a highly concentrated, oligopolistic supply side market alongside an NHS which is purchasing at the disaggregated hospital level.

Finally, Chapter 7 empirically tested various hypotheses about the choice of hip prostheses in terms of the characteristics of the patient and beyond this at the hospital level, to determine whether the NHS is a homogenous entity or whether the manufacturers of prostheses are able to exploit their potential seller power. The empirical work revealed that patient characteristics do not explain much of the variation in prosthesis choice, and that there are large differences in prosthesis choice between different hospitals. It also identified considerable differences between hospitals in the extent of their specialisation (purchasing from a number of manufacturers) with larger hospitals tending to be less specialised, although this is reversed at very high values of hospital size – the very largest hospitals are more specialised. Interestingly, some evidence was found that might be consistent with potential market sharing by the manufacturers in terms of both regional and product markets. Thus Stryker is becoming increasingly dominant in the cemented sector, while Depuy is increasingly dominant in cementless prostheses. There is also evidence that Stryker achieves higher market shares in the part of the country close to its base, while Depuy achieves higher shares in regions near to its base, although the historical reasons for this clustering should also be taken into account. Finally, related to the introduction of the PbR policy, there was some evidence that hospitals have become more specialised in their purchasing (hospitals purchasing from fewer manufacturers) over time, especially around the time of the introduction of PbR.

8.2 Contributions to the literature

In terms of the existing academic literature, this thesis makes four main contributions. First, in the area of economic evaluation, it shows that short term data on prosthesis survival rates may not be sufficient to make decisions regarding the long term cost-effectiveness of alternative prostheses. On the case study used here, it appears that extrapolations based on this short-term data are not robust and thus cannot be used to reliably predict prosthesis survival into the future. Second, although one might expect that the characteristics of the patient should explain much of

the variation in the choice of brand of prosthesis by hospitals, this does not appear to be the case. Instead, much of the variation in prosthesis choice can be explained in terms of the hospital in which the patient undergoes the surgery: this may be a reflection of a principal-agent situation, in which it is the agent who makes the decision, and different agents have different preferences. Third, the thesis has implications for the literature on public procurement. In the context of the NHS, I confirm the findings from the recent NAO report - procurement in the NHS is not at all uniform, but appears to differ significantly between different hospitals at the local level. This might suggest that the suppliers of hip prostheses have the potential to exploit their market power – rather than being faced with a single powerful buyer, they supply to a fragmented set of disaggregated purchasers. Fourthly, the thesis presents what is a rare case study of the nature of competition in a medical devices market. Although the evidence presented here cannot be conclusive – further more detailed analysis is needed in the future – it does provide some evidence that is consistent with potential market power and possible market sharing by the main manufacturers of hip prostheses: the patterns of market shares observed are not inconsistent with the results of the theoretical and empirical literatures on collusion and cartels. This merits further investigation including consideration of other explanatory factors for regional clustering of hospital purchases.

8.3 Main policy implications

The thesis as a whole clearly underlines the value of the NJR, which will become an increasingly useful resource in guiding efficient decision making in this area as more data accumulates. However, it is important to highlight some of problems encountered when using the NJR in this thesis, these include: poor coverage of the NJR in the year 2003; the linkage of HES and NJR data-sets does not include patients from Wales[60] or those patients funded from the independent sector. In that sense, it is not comprehensive. This thesis is also unable to establish how the current ODEP classification scheme reported in the NJR, is helping to inform decision making in the NHS - there are a number of prosthesis brands which are currently being implanted but for which there is no, or a relatively poor ODEP rating.

[60] Although the NJR data-set can be separately linked the PEDW data-set (Patient Episode Data Set for Wales)

Second, this thesis serves to reinforce the main message to emerge from the NAO report on procurement in the NHS. It does not appear that the NHS is exploiting its potential position as a buyer with considerable buyer power. There are large variations in the purchasing patterns of prostheses across hospitals, and much of the variance appears to be unexplainable. This raises the thought that procurement within the NHS deserves close attention and fact finding, and potentially a better informed overhaul. This is clearly highly relevant to current reforms being proposed by the current government.

Third, the implications of chapter 6 and 7, suggest that the UK competition agencies might take a close look at the supplying industry of hip prostheses to the NHS. Although the adjacent pharmaceuticals sector has often been the subject of investigations by the competition authorities, investigations in the medical devices sector are almost unheard of. Given the duopolistic nature of this sector, and the patterns of purchasing observed in this thesis, this deserves some attention. Similarly, it suggests that the NHS itself might widen its awareness of the potential for anti-competitive behaviour amongst its suppliers – it is not just in pharmaceuticals that the NHS is faced with a market which is dominated by a few large multi-national firms.

8.4 Areas for future work

The results of this thesis inevitably leave open a number of areas where further work is required. Some of these are as follows.

To further assess the impact of the introduction of the PbR on the choice of alternative prostheses, and in turn on the behaviour of the manufacturing industry. The thesis has already established some apparent structural breaks with respect to specialisation but further, more detailed analysis is required to identify whether PbR has encouraged or discouraged price and quality competition in the prosthesis industry.

To further examine the market structure and nature of competition in the supplying industry. This should include a deeper analysis of what has been the impact of mergers, barriers to entry,

exclusive selling behaviour and the likely effect on the toughness (or otherwise) of price and quality competition amongst the suppliers.

In order for economic evaluations of alternative prostheses to be conducted, more information is clearly needed on the wider costs of procedures using different prostheses. Ideally this should not just include the direct medical costs, but also the indirect costs to society, the patient and their family (including productivity losses, informal care costs and out-of-pocket expenses). The process of acquiring further information on costs should also involve an investigation of the determinants of the cost of THR at the individual patient level in terms of: patient characteristics; hospital specialisation, economies of scale and prosthesis costs. This will help to identify whether there is a growing gap between costs and reimbursement in the NHS.

A major issue identified by the thesis and now warranting further investigation is the issue of predicting long term prosthesis survival for use in economic evaluation models. Further work might usefully explore the use of complex economic models which incorporate multiple sources of evidence, such as from the literature and various international joint registries. This is necessary in order to provide more robust estimates of long term prosthesis survival into the future. It is also clear that the NJR should be exploited to its maximum potential to provide up to date information to patients, surgeons and the health service community on prosthesis survival rates as they become available.

To contribute to the growing debate about regulation of medical devices (as very recently discussed on the Channel 4 programme ‘Dispatches'[182]). This will involve investigation of whether the same approaches for regulation of pharmaceuticals can also be used for medical devices, specifically hip prostheses. This will require careful consideration, including balancing the promotion of innovation and improving access to new hip prostheses with the need to control costs by restricting market power amongst the suppliers.

An unavoidable gap in the data used in this thesis is the purchase price paid by the NHS to the manufacturers of hip prostheses. Such data is not available in the NJR, or more generally in the public domain. This is not surprising because price is negotiated at the micro level (i.e hospital or PCT) and is not routinely revealed. However, further work is essential to provide a fuller

picture of prices in particular, and procurement in general. Data on prices are required in virtually all of the areas identified above. Thus, analysis of the nature of competition requires evidence of the prices set by manufacturers. To investigate the purchasing behaviour of the NHS price information is also required, for example, to identify whether incentives or scale discounts are part of the procurement process. Finally, of course, price information is essential to inform economic evaluations of alternative prostheses. This will almost certainly require detailed survey research, particularly on the nature of the surgeon-supplier relationship and the procurement procedures taking place at the hospital level.

I have been fortunate to secure an Arthritis Research UK post-doctorate Foundation Fellowship which funds three years further study, and this will provide me with the opportunity to pursue some of these areas in more detail.

Appendices

Appendix 1, Cup and Stem mix for prostheses implanted in 2008/9 [9]

This table reports the combinations of brands of cups and stems most commonly used in England and Wales in 2008/9. This provides the evidence for the statement made in 2.2.3 that most cup and stem combinations are from the same manufacturer: with the exception of 5 out of the 26 combinations listed. The most common cup and stem combination by almost 4,000 procedures is the Pinnacle cementless cup with the Corail cementless stem, both manufactured by Depuy.

Cup	Manufacturer Cup	Stem	Manufacturer Stem	No. procedures (%)	Mix and match
Pinnacle	Depuy	**Corail**	Depuy	10,429 (18)	No
Contemporary	Stryker	**Exeter V40**	Stryker	6,985 (12)	No
Trident	Stryker	**Exeter V40**	Stryker	4,225 (7)	No
Trident	Stryker	**Accolade**	Stryker	2,875 (5)	No
CSF Plus	JRI	**Furlong HAC**	JRI	2,400 (4)	No
Elite Plus Ogee	Depuy	**Exeter V40**	Stryker	1,941 (3)	Yes
Exeter Duration	Stryker	**Exeter V40**	Stryker	1,482 (3)	No
Exceed	Biomet	**Taperloc cementless**	Biomet	1,300 (2)	No
Trilogy	Zimmer	**Exeter V40**	Stryker	1,293 (2)	Yes
Trilogy	Zimmer	**CPT**	Zimmer	1,135 (2)	No
CSF	JRI	**Furlong HAC**	JRI	1,038 (2)	No
ZCA	Zimmer	**CPT**	Zimmer	869 (2)	No
EPF Plus	Smith & Nephew	**SL Plus**	Smith & Nephew	805 (1)	No
Elite Plus cemented cup	Depuy	**Exeter V40**	Stryker	765 (1)	Yes
Charnley	Depuy	**Charnley**	Depuy	689 (1)	No
Charnley Ogee	Depuy	**Charnley**	Depuy	676 (1)	No
Trilogy	Zimmer	**Corail**	Depuy	496 (1)	Yes
Duraloc cementless cup	Depuy	**Corail**	Depuy	493 (1)	No
Stanmore Arcom	Biomet	**Stanmore Modular**	Biomet	442 (1)	No
Elite Plus cemented cup	Depuy	**Corail**	Depuy	399 (1)	No
Procotyl	Wright Medical UK Ltd	**Profemur cementless**	Wright Medical UK Ltd	387 (1)	No
Allofit	Zimmer	**CLS cementless**	Zimmer	370 (1)	No
Low Profile Muller	Zimmer	**MS-30**	Zimmer	340 (1)	No
Marathon	Depuy	**Corail**	Depuy	334 (1)	No
Pinnacle	Depuy	**Exeter V40**	Stryker	331 (1)	Yes
Reflection	Smith & Nephew	**Synergy**	Smith & Nephew	295 (1)	No

Appendix 2, The NJR data

This Appendix describes the process of accessing, merging and cleaning the NJR (and HES) data-sets used in this thesis.

Northgate Solutions manages the NJR and HES data-sets on behalf of the Department of Health (DoH). The process of accessing NJR and HES data for the purpose of this thesis has been extensive and involved communications with both Northgate Solutions and the DoH in order to satisfy data security concerns. This was the first data request Northgate and the NJR had received and as such, was a 'learning process'. The much publicised 'loss' of public data from government departments means that obtaining information particularly from HES has become a complex and prolonged process. The full process of securing data access involved completion of the following: NJR data request form; provision of a research protocol; completion of a HES tabulation pack; securing DoH, SCAG (Security and Confidentiality Advisory Group) approval, local Research Governance Committee approval and NRES (ethics approval). The process was initiated in early 2008 and resulted in full data being received by May 2009.

Data was requested on all patients under-going a cemented or cementless primary total hip replacement. Patients undergoing any knee surgery, hip resurfacing and hip revision surgery were excluded from the analysis.

Six NJR files were received in text file format:

- Hips - containing 356,340 observations (Variables including reasons for primary and revision surgery).
- Knees - containing 356,226 observations (Variables including reasons for primary and revision surgery).
- Hip articulation - containing 329,527 observations (Variables including type of cup, head and head size of the prostheses).
- Operations - containing 712,566 observations (Variables include patient characteristics; anaesthesia used, surgical unit number, funding and patient death date).

- Components - containing 3771,172 observations (Variables include patient and procedure identifiers. Prosthesis brand, manufacturer, description, codes and batch numbers).
- Linked primaries - containing 9268 observations, this file provides the patient identifiers to link primary procedures and revision procedures for each patient record.

All files were linkable by a unique patient identifier set up by Northgate Solutions.

Of the six NJR data-sets: Knees, Hip articulation, Hips and Linked primaries were not used for the purpose of this thesis. Table 1 lists the variables included in the raw NJR spreadsheets provided.

Hip	Hip articulation	Component	Operations
Data base	NJR Index No	Data base	Database
NJR Index No	Procedure ID	NJR Index No	NJR Index No
Procedure ID	Cup	Procedure ID	Procedure ID
Previous Procedure	head	Component ID	funding
Osteoarthritis	headsize	Implant Category ID	waiting list initative
ankylosing spondylitis		Implant Batch No	tertiary referral
avascular necrosis		Manufacturer ID	validate overide
congenital dislocation dysplasia of hip		Brand ID	created by
failed hemiarthroplasty		Brand	completed user
failed internal fixation		Details	lead surgeon id
fractured acetabulum		Cat No	consultant id
fractured neck of femur		Manufacturer	general anaesthesia used
other hip trauma		Category Code	epidural anaesthsia used
other inflammatory arthropathy		Category	nerve block anaesthesia used
perthes		Implant Category Group ID	spinal anaesthesia used
previous arthrodesis		Implant Category Group	sedation used
infection		Component Type ID	asa grade
psoriatic arthropathy		Implant Type	consent
seropositive rheumatoid arthritis		Implant Type Group	nhs number
slipped upper femoral epiphysis			nhs number traced
Indication other			nnn
Other indication specific			nnnid
primary procedure type detailed			sex
patient position			weight
incision approach			height
Trochanteric osteotomy			bmi
complex osteotomy			age at operation date
minimally invasive surgery used			patient hospital identifier
incision length			patient death date
image guided surgery used			joint
femoral pulsatile powered lavage used			revision reoperation date
acetabular pulsatile powered lavage used			patient procedure
femoral bone graft used			opcs4
acetabular bone graft used			bilateral indicator
femoral prosthesis cemented			side
gun used			technique1
cement used retrograde			technique2
proximal seal used with gun			technique3

femoral cement mixed			technique4
acetabular prosthesis cemented			technique5
pressuriser used			laminar low theatre
acetabular cement mixed			surgical unit id
aspirin			completed date
chloroquinine			operation date
low dose heparin			created date
low mol wt heparin			patient physical ID
pentasaccharide			patient proced ID
warfarin			
other chemical			
other chemical specify			
foot pump			
intermittent calf compression			
TED stockings			
other mechanical			
other mechanical specify			
No thromboprophylaxis selected			
no uie specified			
calcar crack			
pelvic penetration			
shaft fracture			
shaft penetration			
trochanteric fracture			
other ui event			
other ui event specify			
aseptic loosening stem			
aseptic loosening socket			
implant fracture stem			
implant fracture socket			
implant fracture head			
incorrect sizing head socket mismatch socket			
incorrect sizing head socket mismatch head			
lysis stem			
lysis socket			
malalignment stem			
malalignment socket			
periprosthetic fracture stem			
periprosthetic fracture socket			
dislocation subluxation			

revision infection			
revision pain			
wear of acetabular component			
dissociation of liner			
other indication for revision			
other indication for revision specify			
cemented stem removed			
stem cement removed			
cemented cup removed			
cup cement removed			
uncemented stem removed			
uncemented cup removed			
uncemented cup liner only removed			
femoral head removed			
wound exploration			
open reduction of dislocation			
excision heterotopic bone			
socket augmentation			
orif trochanter			
orif femur			
focal bone graft only femur			
focal bone graft only acetabulum			
other reoperation			
other reoperation specified			
Insert Date			
Ind For Imp MDS3 Previous Infection			
Ind For Imp Trauma Acute Neck Femur			
Ind For Imp Trauma Chronic			
Ind For Imp Previous Hip Surgery			
MDS3 Bone graft Used Femur YN			
MDS3 Bone graft Used Acetabulum YN			
MDS3 Femoral Stem Removed YN			
MDS3 Femoral Stem Removed Brand			
MDS3 Acetabular Cup Removed YN			
MDS3 Acetabular Cup Removed Brand			
Procedure Type ID			

Table 1: Variables included in raw NJR spread sheets

One HES data file was received, containing 398,914 observations on all hip and knee replacement operations carried out between the 1st April, 2003 and 31st March, 2008. Table 3 lists the variables provided by HES.

HES
linkmethod1
linkmethod2
linkmethod3
linkmethod4
linkmethod5
linkmethod6
linkmethod7
hes year
epikey
nhs number
pseudo hes id
discharge date
elected date
epiend
epistart
admission date
admission category
admin category
epiorder
epistat
epitype
main speciality
operation status
spell begin
spell end
treatment speciality
v code indicator
ward at start episode
current ward
local authority district 1998
pcg code
pcg original
pct code
pct nhs
pct historic
post dischare destination

county of residence
resdhsc
government office of residence
sha of residence
local authrority of residence
pct of residence
regional office of residence
sha of residence historic
rural urban indicator
lower super output area
middle super output area
ward in 91
ward in 98
hrg late
hrg late35
hrg nhs
hrg nhs generated code
hrg original
hrg original 35
end age
ethnic code binary
ethnic code
marital status
sex
start age
dhsc treatment
government office of treatment
ha of treatment
pct treat
procode
procode3
procodet
provider type
regional office of treatment
site of treatment
sha of treatment
purchase code
purd hsc
commissioners regional office
commissioners regional sha
commissioner cod

referral orginal
admission method
source of admission
admission status
carer support indicator
category
patient classification
detention category
discharge destination
discharge method
first regular day or night admission
intended management
legal group of patient
legal category
bed year
commissioning serial number
elective duration
waiting time
post operative duration
poast natal stay
pre-operative duration
provider spell number
spell duration
cause code 3
cause code 4
diagnosis 1
diagnosis 2
diagnosis 3
diagnosis 4
diagnosis 5
diagnosis 6
diagnosis 7
diagnosis 8
diagnosis 9
diagnosis 10
diagnosis 11
diagnosis 12
diagnosis 13
diagnosis 14
operation date 1
operation 1

operation date 2
operation 2
operation date 3
operation 3
operation date 4
operation 4
operation date 5
operation 5
operation date 6
operation 6
operation date 7
operation 7
operation date 8
operation 8
operation date 9
operation 9
operation date 10
operation 10
operation date 11
operation 11
operation date 12
operation 12

Table 2, variables provided in raw HES data

All data-sets were imported into STATA, version 11.

These data were used to construct the two data-sets used in chapter 7 as follows:

Data-set 1: Individual patient level data (Primary THR using NJR and HES linked data-sets.

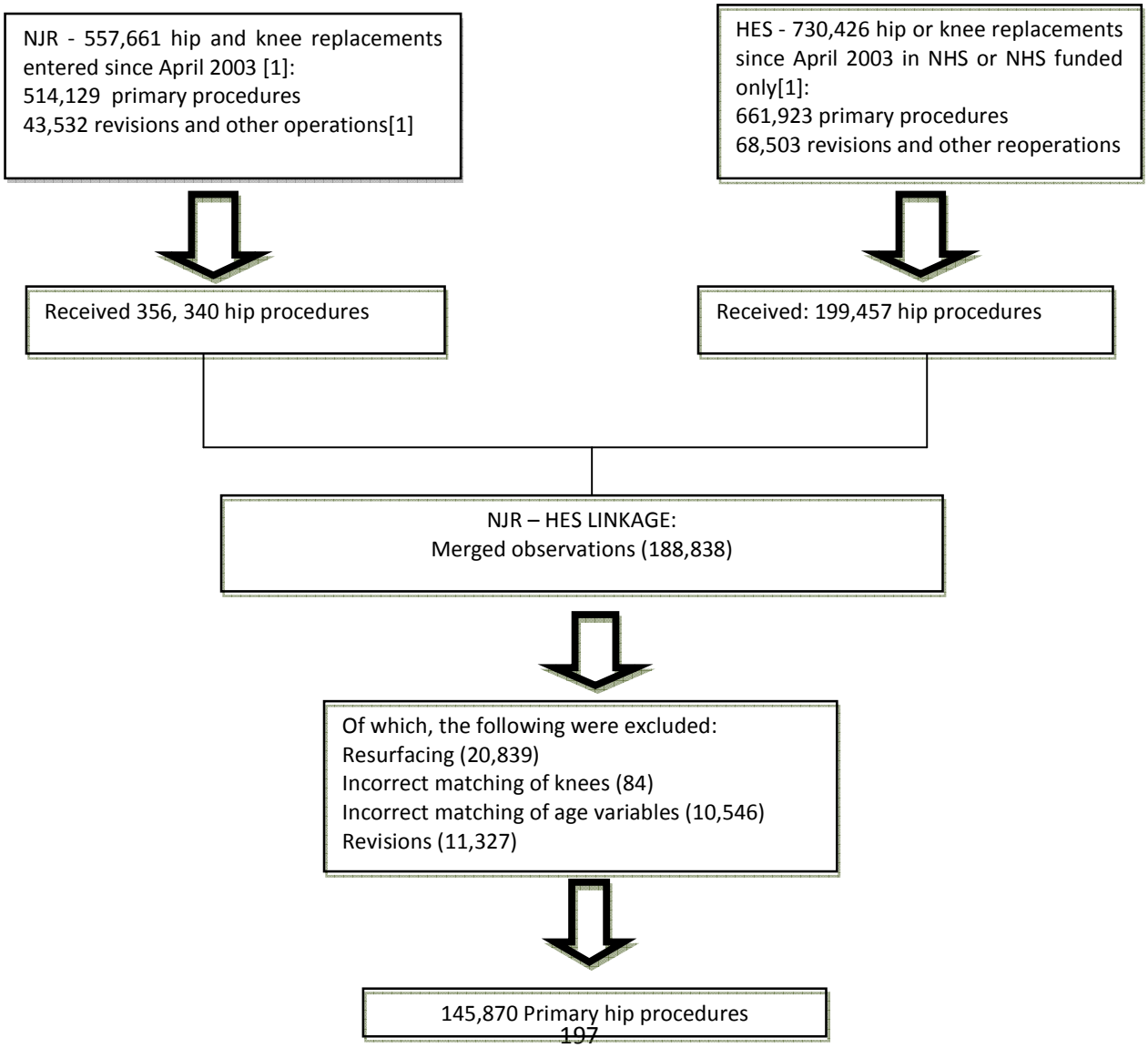
The 'components' file contains multiple observations for each individual patient episode – for one episode (surgery) a patient will receive a cup and a stem prosthesis along with some form of fixation (i.e. cement). A unique patient identifier was created by joining together the NJR index number and the procedure id to give an overall procedure id for each patient. This was then used to reshape the data from long to wide so that for each patient procedure (surgery) the brand and manufacturer of the prosthesis were variables presented alongside patient identifiers.

The same process was used for the remaining NJR data-sets, to establish a unique patient identifier for each episode. The data-sets were then merged together using the unique patient identifier.

The HES data-set was cleaned for any duplicate patient episodes, using the NJR index number (in HES), epi-start (episode start date), epiorder (episode order) and epi-end (episode end). If there was an exact match on these variables, it was established as a duplicate episode and dropped from the data-set. This resulted in 2292 of the 398,913 observations being dropped.

Figure 1 shows a flowchart of the merging process.

Figure 1, Flow chart showing linkage of NJR procedures with HES records - hips[25]



The cleaning and merging process resulted in a final NJR and HES linked data-set of 145,870 patient observations. Table 3 reports the final data-set dimensions. This process involved dropping a large number of the 356,340 individual patient level observations provided from the NJR. A large proportion of observations were 'unmatchable' in the merging of HES and NJR. While the total number of HES observations was 398,914, only approximately half of these observations are hip procedures and it is only these observations which were merged with the 356,340 NJR patient level observations. NJR patient level observations were also dropped for those patients treated in Wales and those funded by the IS because HES only contains data on patients treated in England and funded by the NHS.

As a point of comparison, Sibanda et al were able to successfully link 167,076 of a possible 327,557 primary hip and knee procedures for 2003 to 2006 and of these, 76,576 were primary hip replacements i.e. 23.3% of the 327,557. I was able to link 36% of primary hip procedures out of the possible 398,914 patient observations provided by HES.

Observations	145,870
Variables	263
Years	4/2003 - 12/2008

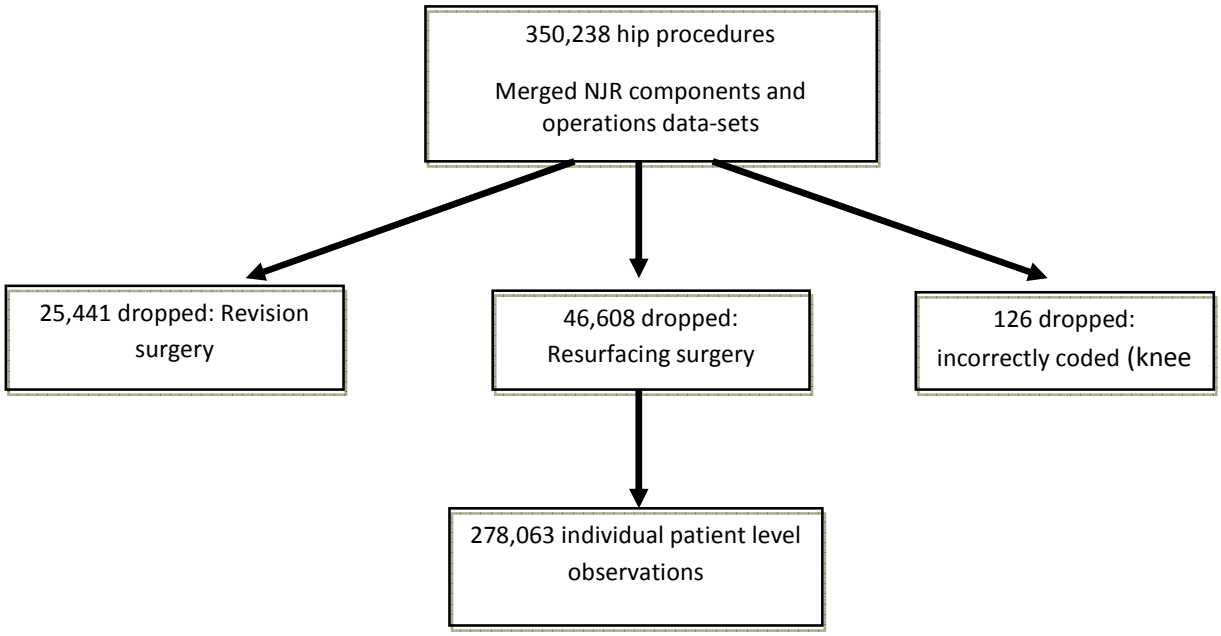
Table 3 - Data dimensions - data-set 1 (NJR & HES data)

Given the loss of such a sizeable proportion of the NJR observations in this process, I decided to also construct a further data-set excluding the HES data-set, but using some of the hospital information from the HES data. This is data-set 2.

Data-set 2: Hospital panel data

The NJR cleaned data-sets 'operations' and reshaped 'components' were merged as in the process described in data-set 1 above. This resulted in a data-set size of 350,238 hip patient level observations. As shown in Figure 2, some of these had to be dropped, resulting in a data-set of 278,050 individual patient level observations shown in table 4.

Figure 2: NJR individual patient level data



Observations	278,063
Variables	501
Years	4/2003 - 12/2008

Table 4 - Data dimensions - NJR individual patient level data

From these NJR individual patient level data, I constructed a hospital level panel for each year 2003 to 2008. The hospital year is the unit of observation. To do this, the 'collapse' command in STATA was used to generate 'mean per hospital year' variables from the individual patient level observation variables e.g. mean number of Exeter V40 prostheses implanted in hospital x in 2003. This resulted in a total of 2281 hospital year observations, accounting for the 278,063 patients.

The NJR does not collect data on hospital characteristics (required for the analysis in chapter 7). However, for this purpose, the linking stage provided a very useful source. The HES data includes hospital characteristics such as: PCT of treatment, provider type and so on. Consequently, in the previous matching stage, I had been able to attach characteristics to most of the hospital identities in the NJR data. In this way, I was able to attach to each NJR patient

observation, the characteristics of the hospital the patient was treated in, so long as any patient in that hospital could be matched in the linking stage.

Hospital, year and an identification number were extracted from the hospital panel data-set and transferred to excel for ease of adding in hospital characteristic variables (bearing in mind that this process involves matching variables from an individual patient level panel to a hospital level panel).

The following algorithm was employed:

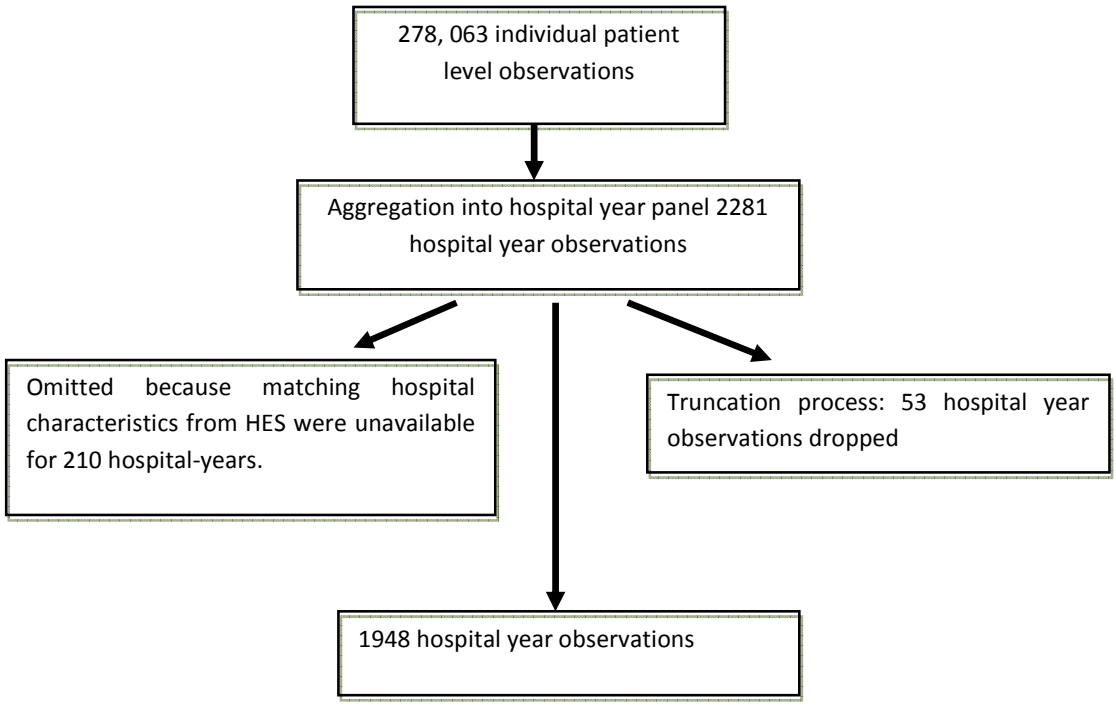
1. Compile a list of all hospital in the hospital panel data-set
2. Identify these hospitals from the hospital panel data-set (in STATA)
3. Extract hospital characteristics for each of the hospitals in the hospital panel data-set[61] from data-set 1: PCT of hospital, and hospital type (NHS trust, Foundation trust, Independent sector and so on).
4. Generate a new variable in excel for PCT and hospital type alongside the hospital identifier.
5. Import the hospital year variable into excel and reformat so that data is available for each hospital where reported
6. Convert data-set 3 from Excel into a STATA file.
7. Merge the existing hospital panel data-set with the new data added from excel.

The merger process resulted in an exact match for 1948 hospital year observations (covering 258,069 of the 278,063 patients). Of these 333 hospital-year observations (19373 patients) were not successfully matched for the following reasons: i. Hospitals in the NJR data-set could not be successfully identified in the NJR and HES matching process, ii. Data were not available for a small number of patient characteristics (age and gender). As explained earlier, these lost observations will include patients treated in Wales (HES only contains data on England) and those treated and funded by the independent sector (HES does contain data on patients treated in the independent sector, but only if they are funded by the NHS). Finally, in order to bring the panel closer to balance, a small number of observations were excluded based on a two stage algorithm. This screening identified 53 hospital year observations (accounting for only 334 patients) which were subsequently omitted. The resulting panel is 374 hospitals, for 265 of which

[61] This is not the full 2071 hospital year observations, but for each hospital

have a full set of observations in each of the 6 years, with an average of 5.5 year observations for each hospital. Table 5 provides an overview of the data dimensions both before and after the exclusion process.

Figure 3 Hospital panel (data-set 3)



	2003	2004	2005	2006	2007	2008
Before Exclusions						
No. Hospitals	308	352	351	358	353	349
No. Patients	22072	41571	46108	46285	51186	51181
After Exclusions						
No. Hospitals	306	341	341	350	344	336
No. Patients	22061	41512	46042	46214	51152	51088

Table 5 - Exclusions from data-set 3 (NJR data)

Limitations of the data-sets:

As discussed earlier in this section, this is the first data release that Northgate has produced on behalf of the NJR linked to HES. Consequently the process of cleaning and linking was no doubt more time consuming than is the case to date. The NJR data required considerable cleaning for errors such as age anomalies. It also required coding, de-stringing and generating new variables and re-shaping of data-sets. The HES data-set also required considerable cleaning.

I have attempted to be as rigouress and transparent about the cleaning and merging process in order to make the process replicable. In order to avoid errors in the data, I have exercised caution with regards to dubious variables, choosing to drop them from the data-set. I have also checked descriptives on each data-set with those reported in the NJR Annual reports as a ‘check’ on the data validity. Inevitably, there may have been some errors in the cleaning and linking/merging process.

Appendix 3, Search strategy for OVID Medline (updated search: May 2010) used in literature review (chapter 4)

Search criteria	Results
total hip replacement.mp.	5164
*Arthroplasty, Replacement, Hip/	9477
total hip arthroplasty.mp.	7298
(hip adj prosthes$).tw.	2776
or/1-4	18040
cost$.mp.	325088
resource use$.mp.	3066
*Cost-Benefit Analysis/	3539
*"Costs and Cost Analysis"/	4832
*Economics/	9922
*Models, Economics/	0
economic evaluation$.mp.	4375
*Economics, Medical/	4996
or/6-13	339390
5 and 14	745

Appendix 4, Example data extraction form (based on Drummond et al checklist [84] for economic evaluations) used in literature review (chapter 4)

Item (Marinelli et al[121])		Yes	No	N/C	N/A	Extract/ comments
Study design.						
1	The research question is stated.	✓				"To establish a framework in which to evaluate the cost-effectiveness of cementless and cemented implants and to analyse how device cost and revision affect the model"
2	The economic importance of the research question is stated.	✓				"Randomised controlled trials are the gold standard for demonstrating the clinical benefits of new technologies. However, detecting small differences in failure rates among implants requires randomizing large numbers of patients and following them for extended periods (15-20 years, or longer). These studies are difficult to perform due to practical considerations of time and cost. In contrast, decision-analysis techniques offer the potential to analyse the performance of a new technology prior to the availability of long-term clinical outcome data. Furthermore, the results from a well-designed decision analysis study can guide further clinical and laboratory research based on the variables that the have the greatest influence on cost-effectiveness. Finally, a cost effectiveness framework can also be readily updated as new information on cost and clinical effectiveness emerges from randomised trials and cohort studies."
3	The viewpoint(s) of the analysis are clearly stated and justified.	✓				"The costs of cementless and cemented THA were estimated from a payer perspective using average hospital costs for prosthetic implants in 2006 Euro's"
4	The rationale for choosing alternative programmes or interventions compared is stated.	✓				Cemented and cementless implants – this reflects current standard practice.
5	The alternatives being compared are clearly described.			✓		"Several different devices (..) are regularly implanted at our Orthopaedics Department. "

	Appendix 4					
6	The form of economic evaluation used is stated.	✓				The authors classify the study as cost-effectiveness analysis. Using the Drummond checklist, it could be described as a cost-utility analysis.
7	The choice of form of economic evaluation is justified in relation to the questions addressed.			✓		"A Markov decision model was used to analyse a theoretical cohort of 70-year patients…."
Data collection.						
8	The source(s) of effectiveness estimates used are stated.	✓				Data on prosthesis revision rates is taken from a prosthesis register (RIPO register). Age-specific probability of death was determined from 2001 United States Life Tables. Published sources were used for other clinical estimates such as peri-operative death and utilities in the model were based on index scored reported in the literature. Methods used to derive the estimates were not explicit.
9	Details of the design and results of effectiveness study are given (if based on a single study).			✓		Further details on the study methodology and greater detail on deriving effectiveness sources is required.
10	Details of the methods of synthesis or meta-analysis of estimates are given (if based on a synthesis of a number of effectiveness studies).			✓		Further details on the methods of synthesis are required.
11	The primary outcome measure(s) for the economic evaluation are clearly stated.	✓				Quality Adjusted Life Years (QALYs) discounted at a yearly rate of 3%. QALYs were estimated using the Markov model.
12	Methods to value benefits are stated.	✓				Utilities were based on quality well-being index scores reported in the literature.
13	Details of the subjects from whom valuations were obtained were given.			✓		Information on utility scores provided but not on subject details other than age.
14	Productivity changes (if included) are reported separately.		✓			Not discussed
15	The relevance of productivity changes to the study question is discussed.		✓			The authors acknowledge that lack of inclusion of all societal costs is a limitation of the study.

	Appendix 4					
16	Quantities of resource use are reported separately from their unit costs.			✓		Resource use not reported in detail or source
17	Methods for the estimation of quantities and unit costs are described.			✓		Yes, but only prosthesis cost. Costs were not broken down.
18	Currency and price data are recorded.	✓				Euro 2006
19	Details of currency of price adjustments for inflation or currency conversion are given.				✓	n.a
20	Details of any model used are given.	✓				Markov model was used. The model structure was provided in a figure.
21	The choice of model used and the key parameters on which it is based are justified.	✓				Appropriate choice of model for this setting.
Analysis and interpretation of results						
22	Time horizon of costs and benefits is stated.			✓		Not made explicit, although it appears to be 5 years.
23	The discount rate(s) is stated.	✓				3% applied to costs and outcomes.
24	The choice of discount rate(s) is justified.	✓				Reference for choice provided.
25	An explanation is given if costs and benefits are not discounted.		✓			n.a
26	Details of statistical tests and confidence intervals are given for stochastic data.			✓		The model is reportedly probabilistic, although details of this in the methodology and results are not provided.
27	The approach to sensitivity analysis is given.			✓		A sensitivity analysis was performed on revision rates, prosthesis costs, preoperative mortality, infection rates and utility values. Details of the sensitivity analysis is not fully reported and thus not fully justified.
28	The choice of variables for sensitivity analysis is justified.			✓		See no. 27
29	The ranges over which the variables are varied are justified.			✓		See no.27
30	Relevant alternatives are compared.	✓				See sections 4 & 5

31	**Appendix 4** Incremental analysis is reported.	✓				Yes
32	Major outcomes are presented in a disaggregated as well as aggregated form.		✓			Outcomes are only reported in aggregated form.
33	The answer to the study question is given.	✓				The authors conclude that the risk of revision is similar between cemented and cementless prosthesis groups, in terms of QALYs and the cost-difference as non-significant.
34	Conclusions follow from the data reported.			✓		Conclusions follow. However, the conclusions are hard to follow due to the limited reporting of methodology, sources and presentation of results.
35	Conclusions are accompanied by the appropriate caveats.			✓		Further clarity could be provided.

Appendix 5, Summary of included studies in chapter 4 (based on the Drummond checklist for economic evaluation studies[84])

Study	Country	Study Design	Interventions	Time horizon	Perspective	Currency & price year	Outcome measure	Modelling	Sponsor	Source of cost & resource use data	Source of effectiveness data
Baxter (99)	UK	CEA	charnley v alternative	20 years	health care system	UK £ 1994	survival rate for prosthesis	Deterministic model: differential life expectancies between prostheses	NHS R &D HTA Programme	Primary study – 2 UK hospitals	Revision rates – published sources
Boardman (97)	USA	cost analysis	hybrid v cementless	-	health care system	US $ 1988	comparison of hospital costs and reimbursement	-	-	Local hospital data (billing sheets), Hospital medical centre-patient records	-
Briggs (04)	UK/ Sweden	CUA	charnley v spectron	60 years	health care system	UK £ 2000/1	QALY/ICER	probabilistic decision model	Part funded by Smith & Nephew	NHS reference costs	Revision rates - Swedish hip register & QALYS - EQ-5D† from local study
Daellenbach (90)	New Zealand	CEA	cemented v cementless	Lifetime horizon	health care system	NZ $ 1985	survival rate for prosthesis	deterministic model: differential life expectancies between prostheses	-	Costing records	Life-tables & Revision rates - published sources
Faulkner (98)	UK	Critical review, CEA	cemented, cementless & hybrid	20 years	health care system	UK £	survival rate for prosthesis	deterministic model: differential life expectancies between prostheses	NHS R &D HTA programme	Primary study – 2 UK hospitals	Mortality data – ONS‡ England and Wales & Revision rates - published sources
Fitzpatrick (98)	UK	Systematic Review, CUA	charnley v new prosthesis	60 years	health care system	UK £	QALY/ICER	probabilistic decision model	NHS R &D HTA Programme	Primary data – local study & published sources	QALY estimates and revision rates – published sources
Gillespie (95)	Sweden/ Australia	CEA	hypothetical: standard v new prosthesis	20 years	health care system	US $	survival rate for prosthesis	deterministic model: differential life expectancies between prostheses	-	Local Health records, Australian bureau of statistics & published sources	Revision rates – primary review study
Givon(98)	Israel	CUA	Cemented//hybrid and cementless with and without HA coating	9 years	Health care system	US $ 1994	Cost/QALY	-	-	DRG used for estimating costs & single study	Single study – secondary care
Marine	Italy	CEA	cemented v	5 years	health care	Euro 2006	QALY/	Probabilistic	-	Local clinical	Mortality rates –

(App 5) **-lli (08)**			cementless		system		ICER	decision model		database	life table, Revision rates - local register & Utilities - published sources
Metz (98)	International/USA	cost analysis	cemented v cementless	-	health care system	US $ 1996	total cost of intervention	-	-	Survey data from surgeons	-
Murray (95)	UK	cost analysis	cemented v cementless	-	health care system	UK £ 1994	price of prosthesis	-	Biomet Ltd	Survey of prosthesis manufacturers	Revision rates from published sources
Pingsmann (98)	Germany	cost analysis	cemented v cementless	-	health care system	DM	total cost of intervention	-	-	Patient records	-
Pynsent(96)	UK	Critical review, CEA	cemented, cementless & hybrid	20 years?	health care system	UK £	survival rate for prosthesis	deterministic model: differential life expectancies between prostheses	-	-	ONS mortality data & revision rates, published sources
Scheerlink (04)	Belgium	cost analysis	Dacup, CPT/Duraloc, Vectra/ZCA & Others	-	health care system	Euro 2001/2	total hospital cost	-	-	Local study – discharge summary	HR-QoL questionnaires pre and post-operative
Spiegelhalter(03)	UK	CUA	charnley v alternative	60 years	health care system	UK £	QALY/ICER	probabilistic decision model	MRC grant	Published sources	Revision rates – Swedish hip register, Mortality rates - UK & QALYs - published sources
Yates (06)	UK	cost analysis	cemented v cementless	-	health care system	UK £ 2003/4	total cost of prosthesis	-	-	Prosthesis prices - manufacturers	-
Unnanuntana (09)	USA	Cost analysis	Cemented/cementless/hybrid	-	Health care system	US $ 2008	Mean cost femoral stem	-	-	3 academic medical centres	-

†, EQ-5D is a standardised instrument for use as a measure of health outcome

‡ ONS – UK Office for National Statistics

Appendix 6, The outcomes for risk of bias in the studies reviewed in chapter 4

Item	Baxter (99)	Boardman (97)	Briggs (04)	Daellenbach (90)	Faulkner (98)	Fitzpatrick (98)	Gillespie (95)	Givon (98)	Marinelli (08)	Metz (98)	Murray (95)	Pingsmann (98)	Pynsent (96)	Scheerlink (04)	Spiegelhalter (03)	Unnanutana (09
Assigned to treatment adequately concealed prior to allocation?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Outcomes of participants who withdrew described and included in analysis?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Outcome assessors blinded to treatment status?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Treatment and control compatible at entry?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Participants blind to assignment status after allocation?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Treatment providers blind to assignment status?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Were care programmes other than trial options, identical?	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1
Inclusion and Exclusion criteria clearly defined?	0	2	0	0	2	2	0	1	0	0	0	0	0	1	0	0
Interventions clearly defined?	2	1	2	1	2	2	0	2	1	0	0	1	0	1	1	1
Outcome measures used clearly defined?	2	1	2	1	2	2	0	2	1	1	1	1	0	1	1	2
Diagnostic tests used in outcome assessment clinically useful?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Surveillance active, and of clinically appropriate duration?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Key: 0 = not defined; 1 = adequate; 2 = clearly defined

Appendix 7, The Swedish data-set

Chapter 5 uses data from SHAR to explore potential difficulties of extrapolating survival curves (survival of prostheses) over a lifetime horizon. The chapter uses a previous well-known study[18] as an example, and extends the original data-set to include an additional 8 years more data in order to assess the accuracy of predictions made in 2000 in the light of a longer time series.

The data for the original study was obtained from SHAR on all patients who received either a Charnley or a Spectron hip prosthesis in the period 1992-2000, where both the cup and stem were from the same manufacturer. For the Spectron, this included all patients receiving an All-Poly Cup and a Spectron EF, or EF primary stem. Table 1 provides an overview of the original data-set dimensions.

Data dimensions	
Total sample size	20,495
Patients receiving a Charnley	18,505
Patients receiving a Spectron	1,990
Mean follow up	4 years, 3 months
Maximum follow up	8 years
Patient years at risk	90000
Failures	574

Table 1 - Data dimensions - Original data from SHAR[18]

For the purpose of this thesis, SHAR was approached for access to an updated form of the data-set, including information on the variables listed in Table 2 for the period 1992 to present.

Swedish Label		English translation
Essential		
GENDER	2	Gender
OPPDAT	3	Date of operation primary procedure
DIAGNOSE	4	Diagnosis
CUP	5	Cup
STAM	6	Stem
OPPAR	35	Date of operation primary procedure
DIAGRP	39	Diagnosis group
AGE	43	Age
AGEGRP	44	Age group
OPRDAT	*45*	*Date of reoperation*
ORSGP	46	Cause of revision by group
REVTIME	48	Time with primary prosthesis
OPRAR	49	Date of reoperation
R	51	Revision indicator
Also if possible		
KLINGRP	36	Clinic group
CUPGRP	37	Cup type group
STAMGRP	38	Stem type group
TLVPCUP	40	Manufacturer cup
TLVPSTAM	41	Manufacturer cup
PROTGRP	42	Prosthesis group
ATGGRP	47	

Table 2 - Data requested from Swedish Hip Arthroplasty Register

The new data-set contains 16 years of data, an additional 8 years from the original data-set. In order to identify the patients in the original data-set (1992-1999), so that they could be followed up in the subsequent 8 years, all patients undergoing primary surgery post-1999 were excluded from the data-set (discussed in chapter 5). Neither the original nor the updated data-set was available with patient identifiers, meaning that data was matched observation by observation in Excel following the steps outlined below:

1. Both data-sets were split into two separate spreadsheets on the basis of whether patients received a Charnley or Spectron prosthesis.

2. They were then further split into yearly spreadsheets

3. Observations were matched ‘observation by observation’ employing the following algorithm:

i. Match on patient age (allowing for 1 year either side)

ii. Exact match on operation date

iii. Exact match on revision date

iv. Count on observations on matched and unmatched observations in both data-sets

v. Exclude all unmatched observations from both data-sets.

vi. Add new variables into old data-set, including ‘revision’ and ‘time until revision surgery’.

98.3% of all the patients in the original data-set were identified in the new data-set. The remaining 1.7% (n=350) non-matched observations were due to occasional minor coding discrepancies. In all such cases, caution was exercised by omitting these patients. Table 3 reports the sample size, patient characteristics and number of revisions in the original and new ‘matched’ datasets. The descriptive statistics show that the two samples are virtually identical in terms of patient characteristics, age and gender.

Despite the rigorous methods employed in matching patients in the two data-sets, there were clearly some unmatched patients, which reduces the sample available for the analysis described in chapter 5. Ideally patient identifiers or some identification number would have been present in both data-sets to enable comprehensive matching. However, as table 3 shows, the two data-sets are very close in dimensions.

	Original data-set		New 'matched' data-set	
	Charnley	**Spectron**	**Charnley**	**Spectron**
Patients	18,505	1990	18,178	1967
Mean age (sd)	72 (9.2)	74 (8.1)	71 (9.2)	74 (8.1)
Age distribution (%)				
<40 years	70 (0.4)	5 (0.3)	66 (0.4)	5 (0.3)
40-50 years	264 (1.4)	16 (0.8)	251 (1.4)	15 (0.8)
50-60 years	1418 (7.7)	60 (3.0)	1,389 (7.7)	60 (3.0)
60-70 years	4836 (26.1)	391 (19.7)	4,753 (26.1)	385 (19.6)
70-80 years	8090 (43.7)	1014 (51.0)	7,945 (43.7)	1,000 (51.0)
80-90 years	3630 (19.6)	481 (24.2)	3,581 (19.6)	479 (24.3)
>90 years	197 (1.1)	23 (1.2)	193 (1.1)	23 (1.2)
Gender (%)				
Female	12337 (66.7)	1472 (74.0)	12,108 (66.7)	1,453 (73.9)
Male	6168 (33.3)	518 (26.0)	6,070 (33.3)	514 (26.1)
Initial diagnosis (%)				
Osteoarthritis	12970 (70.1)	1348 (67.7)	12,826 (79.5)	1329 (69.8)
Fracture	1692 (9.1)	319 (16.0)	1,662 (10.3)	317 (16.6)
Other	3843 (20.8)	323 (16.2)	1,628 (10.1)	258 (13.5)
Revisions (%)				
1992-1999	552 (2.98)	22 (1.10)	528 (2.90)	21 (1.07)
1992-2000	-	-	1,255(6.90)	98(4.98)

Table 3 - Comparison of Samples - Original and new 'matched' data[18]

Appendix 8, Market shares by brand of prostheses (disaggregated)
Table 1 Share of cemented cup

Explanatory variables	Stryker	Depuy	other'
Year (Reference year 2003)			
2004	0.03341	-0.03981*	0.02511
2005	0.11116***	-0.10066***	0.01265
2006	0.12723***	-0.13714***	0.01978
2007	0.16637***	-0.17242***	0.01649
2008	0.18927***	-0.21004***	0.02331
Average Age of Patients	0.00123	0.00206	-0.00259
Proportion of Female Patients	0.06975	-0.03697	-0.00996
Proportion of right sided surgery	-0.17337+	0.08418	0.09012
Region (reference East Anglia)			
South East	0.02315	-0.14854	0.05275
East Midlands	-0.12421	0.07592	0.01162
North West	-0.08078	0.32218***	-0.18575*
London	-0.10005	-0.33429***	0.23740**
Yorkshire	-0.01889	0.27614**	-0.27553**
South West	-0.09908	0.22966*	-0.15835+
South Central	0.16659	0.23339*	-0.37506***
West Midlands	-0.19307	0.33990***	-0.14774+
North East	0.12483	0.07023	-0.15549**
Hospital Type (reference FT)			
Trust	-0.01771	-0.04584+	0.07037
Independent Sector	0.04468	-0.02422	0.04233
NHS Treatment Centre & other	0.07467	-0.05327	-0.05671
Independent Treatment Centre	-0.06092	-0.14919+	0.10924
Hospital size	0.03398	0.02758	0.00676
Hospital size quadratic	-0.00530+	-0.00247	0.00036
Constant	-0.02943	0.16421	0.36701*
σ u	0.5114***	0.43483***	0.37747
σ e	0.2250***	0.20355***	0.20059
Log Likelihood	--669.95	-470.47	-503.94
Number of observations	1892	1892	1892
Left censored observations	534	534	566
Right censored observations	196	196	194

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 2 Share of cemented stem

Explanatory variables	Stryker	Depuy	other'
Year (Reference year 2003)			
2004	0.01548	-0.0285	0.01576
2005	0.07300***	-0.07133***	-0.00497
2006	0.09798***	-0.12907***	-0.00175
2007	0.14272***	-0.17747***	-0.01695
2008	0.16968***	-0.23342***	-0.0211
Average Age of Patients	0.00099	0.0024	-0.0001
Proportion of Female Patients	0.02561	-0.03	-0.01537
Proportion of right sided surgery	-0.05192	-0.0277	0.13524+
Region (reference East Anglia)			
South East	0.01524	-0.0798	0.02968
East Midlands	-0.11117	0.1546	-0.12382
North West	0.00642	0.33032***	-0.30885***
London	-0.09712	-0.20771*	0.13065+
Yorkshire	-0.07406	0.35787***	-0.29638***
South West	0.39425***	-0.1797+	-0.23806**
South Central	0.45293***	-0.1303	-0.38322***
West Midlands	0.17278	0.20088*	-0.31390***
North East	0.08677	0.1233	-0.14185
Hospital Type (reference FT)			
Trust	-0.05045*	-0.02	0.05794*
Independent Sector	-0.03305	0.0314	0.0654
NHS Treatment Centre & other	-0.03155	0.0385	0.02324
Independent Treatment Centre	-0.18990*	0.0033	0.0998
Hospital size	0.01245	0.0263	0.02001
Hospital size quadratic	-0.003	-0.001	-0.00117
Constant	0.2021	0.0157	0.23491
σ u	0.44981***	.40596***	0.33620***
σ e	0.18870***	.19696***	0.1894***
Log Likelihood	-405.36	-466.18	-420.04
Number of observations	1910	1910	1910
Left censored observations	501	725	558
Right censored observations	231	96	151

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 3 Share of cementless cup

Explanatory variables	Stryker	Depuy	other'
Year (reference year 2003)			
2004	0.02824	0.04079	-0.01736
2005	0.03594	0.07495*	-0.04363+
2006	0.03705	0.12856***	-0.06870**
2007	0.02454	0.19422***	-0.09481***
2008	-0.01261	0.25764***	-0.10613***
Average Age of Patients	0.00459	-0.01395***	0.00610*
Proportion of Female Patients	0.02486	0.06963	-0.02845
Proportion of right sided surgery	-0.17157	0.30227*	-0.01443
Region (reference East Anglia)			
South East	0.27451*	0.0466	-0.09283
East Midlands	-0.14797	0.03918	0.07918
North West	0.11633	0.30968**	-0.25482**
London	0.09725	0.0498	0.03979
Yorkshire	0.00154	0.30845*	-0.16349+
South West	0.26158*	0.23548	-0.17198+
South Central	0.43405***	-0.15029	-0.11848
West Midlands	0.22806+	0.25080*	-0.26137**
North East	0.26525+	0.2171	-0.16739
Hospital Type (reference FT)			
Trust	0.05354	-0.07467*	0.02633
Independent Sector	0.03249	-0.07994	0.09165
NHS Treatment Centre & other	0.10253	-0.12511	-0.00413
Independent Treatment Centre	0.0185	-0.02297	0.0498
Hospital size	0.01387	0.04413+	0.02098
Hospital size quadratic	0.00164	-0.00316	-0.00468+
Constant	-0.47509	0.60546*	0.16414
σ u	0.49435***	0.51510***	0.40306***
σ e	0.26632***	0.25744***	0.22703***
Log Likelihood	-80.8	-794.67	-604.54
Number of observations	1850	1850	1850
Left censored	881	870	346
Right censored	120	150	351

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

Table 4 share of Cementless stem

Explanatory variables	Stryker	Depuy	'other'
Year (reference year 2003)			
2004	-0.08468	0.08082	0.03814
2005	-0.06395	0.15407***	-0.0055
2006	-0.14397**	0.21862***	-0.00754
2007	-0.14595**	0.32844***	-0.06604*
2008	-0.18308***	0.40905***	-0.11145***
Average Age of Patients	0.00479	-0.00914	0.00047
Proportion of Female Patients	-0.05845	0.10332	0.00526
Proportion of right sided surgery	-.43023796*	0.61465***	-0.18843
Region (reference East Anglia)			
South East	0.46375**	0.0196	-0.11703
East Midlands	0.03643	-0.00491	0.02792
North West	0.1111	0.36043**	-0.22352*
London	0.25936	-0.08144	0.11209
Yorkshire	0.02862	0.23966	-0.05428
South West	-0.06541	0.37378**	-0.18067+
South Central	0.48807**	-0.28344	-0.03751
West Midlands	0.35916*	0.18577	-0.23176*
North East	0.2835	0.05937	-0.07477
Hospital Type (reference FT)			
Trust	0.10739	-0.15899***	0.06389+
Independent Sector	0.14005	-0.17382	0.11352
NHS Treatment Centre & other	0.22013	-0.25214	0.03352
Independent Treatment Centre	0.27898+	-0.0911	0.02776
Hospital size	0.03944	0.06749*	-0.00048
Hospital size quadratic	-0.00087	-0.0072	-0.00211
Constant	-0.66466	0.11062	0.54013+
σ u	0.66706***	0.56284***	0.43035***
σ e	0.32929***	0.32863***	0.27889***
Log Likelihood	-743.92	-943.49	-843.01
Number of observations	1850	1850	1850
Left censored	881	870	346
Right censored	120	150	351

legend: +p<0.10, * p<0.05; ** p<0.01; *** p<0.001*

References

1. National Joint Registry for England and Wales, *6th Annual Report*. 2008/9.
2. NHS Supply Chain. *ODEP database*. 2011.
3. Browner, B.D., *The Bone and Joint Decade, 2000-2010.* J Bone Joint Surg Am, 1999. **81**(7): p. 903-4.
4. Dandy. D.J and Edwards. D.J, *Essential Orthopaedics and Trauma*. 4th Edition ed. 2003: Churchill Livingston.
5. Horan, F., *EDITORIAL COMMENT.* J Bone Joint Surg Br, 1999. **81-B**(3): p. 377-.
6. Connolly S and Munro A, *Economics of the Public Sector*. 1999, Europe: Prentice Hall.
7. Alexander James MacGregor, *The genetic determinants of osteoarthritis, spinal degenerative disease and musculoskeletal pain - An analysis of data from twins,*. 2008, University of London.
8. Sibanda, N., et al., *Revision Rates after Primary Hip and Knee Replacement in England between 2003 and 2006.* Plos Medicine, 2008. **5**(9): p. 1398-1408.
9. The National Joint Registry., *7th Annual Report*. 2010.
10. Map of Medicine. *Total Hip replacement (arthoplasty)*. 2011; Available from: http://eng.mapofmedicine.com/evidence/map/elective_hip_surgery2.html.
11. Ranawat, C.S., *SURGICAL MANAGEMENT OF THE RHEUMATOID HIP.* Rheumatic Disease Clinics of North America, 1998. **24**(1): p. 129-141.
12. Map of Medicine. *Elective hip surgery*. 2011; Available from: http://eng.mapofmedicine.com/evidence/map/elective_hip_surgery1.html.
13. Hip replacement - Risks. *Risks of hip replacement surgery,*. 2010; Available from: http://www.nhs.uk/Conditions/Hip-replacement/Pages/Complications.aspx.
14. American Academy of Orthopaedic Surgeons. *Joint Revision Surgery - When do I need it?,*. 2007; Available from: http://orthoinfo.aaos.org/topic.cfm?topic=A00510.
15. National Audit Office, *Hip Replacements: Getting it right first time*. 2000, Report by the Comptroller and Auditor General.
16. The Campbell and Cochrane Economic Methods Group and The Evidence for Policy and Practice Information Coordinating Centre. *CCEMG-EPPI Cost Centre Converter*. 2011; Available from: http://eppi.ioe.ac.uk/costconversion/default.aspx.
17. Bordini, B., et al., *Factors affecting aseptic loosening of 4750 total hip arthroplasties: multivariate survival analysis.* BMC Musculoskeletal Disorders, 2007. **8**(1): p. 69.
18. Briggs, A., et al., *Modelling the cost-effectiveness of primary hip replacement: how cost-effective is the Spectron compared to the Charnley prosthesis?* CHE technical paper series, 2003. **28**.
19. Grigoris, P. and D. Hamblen, *THE CONTROL OF NEW PROSTHETIC IMPLANTS.* J Bone Joint Surg Br, 1998. **80-B**(6): p. 941-943.
20. National Institute for Clinical Excellence, *Guidance on the selection of prostheses for primary total hip replacement,*, NICE, Editor. 2000: London.
21. National Joint Registry for England and Wales, *4th Annual Report*. 2007.
22. Briggs, A., et al., *The Use of Probabilistic Decision Models in Technology Assessment: The Case of Total Hip Replacement.* Applied Health Economics and Health Policy, 2004. **3**(2): p. 79-89.
23. National Institute for Clinical Excellence, *Guidance on the selection of prostheses for primary total hip replacement,*, Technology Appraisal Guidance, Editor. 2000.
24. Royal College of Surgeons, *3M Capital Hip System - The lessons learned from an investigation*. 2001.
25. Hospital Episode Statistics. *HES Online - about HES*. 2011; Available from: http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=87.

26. Devlin N J and A. J., *Getting the most out of PROMS*, The Kings Fund & Office of Health Economics, Editor. 2010.
27. Folland. S, Goodman. A.C, and Stano. M, *The Economics of Health and Health Care*. 3rd Edition ed. 2001, New Jersey: Prentice Hall.
28. Rawlins, M.D. and A.J. Culyer, *National Institute for Clinical Excellence and its value judgments.* BMJ, 2004. **329**(7459): p. 224-227.
29. Drummond. M.F, et al., *Methods for the Economic Evaluation of Health Care Programmes*. 3rd Edition ed. 2005, Oxford: Oxford University Press.
30. Briggs A, Claxton K, and Sculpher M., *Decision Modelling for Health Economic Evaluation*, ed. Handbooks in Health Economic Evaluation. 2006: Oxford University Press.
31. Vallejo-Torres, L., et al., *Integrating health economics modeling in the product development cycle of medical devices: A Bayesian approach.* International Journal of Technology Assessment in Health Care, 2008. **24**(04): p. 459-464.
32. Manca, A., *Economic Evaluation of Medical Devices and Drugs—Same or Different?* Value in Health, 2009. **12**(4): p. 401-401.
33. Drummond, M., A. Griffin, and R. Tarricone, *Economic Evaluation for Devices and Drugs—Same or Different?* Value in Health, 2009. **12**(4): p. 402-404.
34. Taylor, R.S. and C.P. Iglesias, *Assessing the Clinical and Cost-Effectiveness of Medical Devices and Drugs: Are They That Different?* Value in Health, 2009. **12**(4): p. 404-406.
35. Church J and W. R., *Industrial Organization - A strategic Approach.* 2000: McGraw-Hill International Editions.
36. Morgan. W, Katz. M, and Rosen. H, *Microeconomics*. 2nd Edition ed. 2009, London: McGraw Hill.
37. Motta M, *Competition Policy - Theory and Practice*. 2004: Cambridge University Press.
38. Levenstein, M.C. and V.Y. Suslow, *What determines cartel success?* Journal of Economic Literature, 2006. **44**(1): p. 43-95.
39. Harrington J.E, J., *Detecting Cartels,*, in *The Handbook of Antitrust Economics,*

. 2008, The MIT Press,.
40. Office of Government Commerce, *An introduction to public procurement*. 2009.
41. The King's Fund. *Commissioning*. 2010; Available from: http://www.kingsfund.org.uk/topics/commissioning/.
42. Department of Health. *Health and Care Reform - Commissioning*. 2011; Available from: http://healthandcare.dh.gov.uk/commissioningmore/.
43. Sir Philip Green, *Efficiency review by Sir Philip Green - Key findings and recommendations,*. 2010.
44. Vale L, et al., *A systematic review of the effectiveness and cost-effectiveness of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease.* Health Technology Assessment, 2002. **6**(15).
45. Burns, A.W.R. and R.B. Bourne, *(vi) Economics of revision total hip arthroplasty.* Current Orthopaedics, 2006. **20**(3): p. 203-207.
46. Anderson J, Neary F, and Pickstone J V, *Surgeons, manufacturers and patients - A transatlantic history of total hip replacement*. Science, Technology and Medicine in Modern History. 2007: Palgrave Macmillan.
47. European Commission - Consumer Affairs. *Medical Devices,*. 2011; Available from: http://ec.europa.eu/consumers/sectors/medical-devices/index_en.htm#top.
48. MHRA, *The CE Mark*, Bulletin, Editor. 2011.
49. Medicine and Healthcare Products Regulatory Agency. *About us*. 2011; Available from: http://www.mhra.gov.uk/Aboutus/index.htm.

50. The Medicines and Healthcare products Regulatory Agency. *UK notified bodies under the medical devices directives,*. 2011; Available from: http://www.mhra.gov.uk/Howweregulate/Devices/NotifiedBodies/UKNotifiedBodiesundertheMedicalDevicesDirectives/index.htm.
51. Taylor R, *What is NICE.* www.evidence-basedmedicine.co.uk, 2002. **3**(4): p. 1-8.
52. NICE, *Guidance on the selection of prostheses for primary total hip replacement.* Online, 2000. **http://www.nice.org.uk/nicemedia/pdf/Guidance_on_the_selection_of_hip_prostheses.pdf**.
53. NHS Supply Chain. *Orthopaedic Data Evaluation Panel*. 2010; Available from: http://www.supplychain.nhs.uk/portal/page/portal/Communities/Orthopaedics/ODEP%20database.
54. Sibanda, N., et al., *Revision Rates after Primary Hip and Knee Replacement in England between 2003 and 2006.* PLoS Med, 2008. **5**(9): p. e179.
55. Faulkner A, et al., *Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model,*, Health Technology Assessment, Editor. 1998.
56. Kolling, C., et al., *Key factors for a successful National Arthroplasty Register.* J Bone Joint Surg Br, 2007. **89-B**(12): p. 1567-1573.
57. Swedish Hip Arthroplasty Registry, *Annual Report 2008*. 2008, Department of Ortopaedics, Sahlgrenska University Hospital.
58. The National Joint Registry, *1st Annual Report*. 2004.
59. The National Joint Registry, *What is the NJR? Aims and objectives*, T.N.J. Registry, Editor.
60. NJR, *4th Annual Report.* The National Joint Registry, 2006/7. **http://www.njrcentre.org.uk/njrcentre/AbouttheNJR/Publicationsandreports/Annualreports/tabid/86/Default.aspx**(Online).
61. The National Joint Registry, *NJR Service Charter,*. 2008.
62. Nuffield Institue for Health and NHS Centre for Reviews and Dissemination, *Total Hip Replacement,.* Effective Health Care, 1996. **2**(7).
63. Dixon, T., et al., *Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need.* Annals of the Rheumatic Diseases, 2004. **63**(7): p. 825-830.
64. Gregg, P., *The National Joint Registry 2nd Annual Report, 2005.* 2005.
65. National joint Registry, *3rd Annual Clinical Report,*. 2005.
66. The National Joint Registry, *5th Annual Report,*. 2008.
67. The National Joint Registry, *6th Annual Report.* 2009.
68. Commission of the European Communities, *Regulation (EEC) No.4064/89 Merger Procedure - Case No IV/M.1286 - Johnson and Johnson/Depuy,*. 1998: Luxembourg.
69. Commission of the European Communities, *Regulation (EEC) No. 4064/89 Merger Procedure - Case No COMP/M.3146 - Smith & Nephew/Centerpulse,*. 2003: Luxembourg,.
70. Audit Office, *Payment by Results Assurance Framework - Independent Sector Pilot Briefing.* 2009.
71. The National Joint Registry, *Prostheses used in hip and knee replacement procedures, 2009*, in *7th Annual Report*. 2009.
72. The National Joint Registry, *Prostheses used in hip and knee replacement procedures, 2008,*. 2008.
73. The National Joint Registry, *Prostheses used in hip and knee replacement procedures.* 2007.
74. Davies, C., et al., *Can choices between alternative hip prostheses be evidence based? a review of the economic evaluation literature.* Cost Eff Resour Alloc. **8**: p. 20.
75. Lorgelly P, et al., *Which prostheses care cost-effective?* The Knee,, 2009. **16**(6): p. 419.
76. Briggs A. Claxton K. and Sculpher M, *Decision Modelling for Health Economic Evaluation*. 2006, Oxford: Oxford University Press.
77. National Institute for Clinical Excellence, *Guide to the Methods of Technology Appraisal*. 2008.

78. Drummond. M.F. Sculpher. M.J. Torrance. G.W. O'Brien. B.J. Stoddart. G.L, *Methods for the Economic Evaluation of Health Care Programmes*. 3rd Edition ed, ed. O.U. Press. 2005, Oxford.
79. Shemilt, I., . Mugford, M. Byford, S. Drummond, M. Eisenstein, E. Knapp, M. et al,, *Incorporating Economics Evidence*, in *Cochrane Handbook for Systematic Reviews of Interventions*. 2008b, John Wiley & Sons: Chichester.
80. Critical Appraisal Skills Programme. *12 questions to help you make sense of a cohort study*. 2004; Available from: http://www.sph.nhs.uk/sph-files/cohort%2012%20questions.pdf.
81. Programme, C.A.S. *10 questions to help you make sense of randomised controlled trials*. 2004; Available from: http://www.sph.nhs.uk/sph-files/rct%20appraisal%20tool.pdf.
82. Philips, Z., et al., *Good practice guidelines for decision-analytic modelling in health technology assessment - A review and consolidation of quality assessment.* Pharmacoeconomics, 2006. **24**(4): p. 355-371.
83. Centre., T.C.a.C.E.M.G.T.E.f.P.a.P.I.a.C. *The Campbell and Cochrane Economics Methods Group & The Evidence for Policy and Practice Information and Coordinating Centre - Cost Converter*. Available from: http://eppi.ioe.ac.uk/costconversion/default.aspx
84. Drummond, M.F. and T.O. Jefferson, *Guidelines for authors and peer reviewers of economic submissions to the BMJ.* BMJ, 1996. **313**(7052): p. 275-283.
85. Pynsent, P.B., S.R. Carter, and C.J.K. Bulstrode, *The total cost of hip-joint replacement; a model for purchasers.* J Public Health, 1996. **18**(2): p. 157-168.
86. Gillespie, W.J., B. Pekarsky, and D.L. O'Connell, *Evaluation of new technologies for total hip replacement. Economic modelling and clinical trials.* J Bone Joint Surg Br, 1995. **77-B**(4): p. 528-533.
87. Daellenbach, H.G., et al., *Economic appraisal of new technology in the absence of survival data - the case of total hip replacement.* Social Science & Medicine, 1990. **31**(12): p. 1287-1293.
88. Baxter, K., *An economic model to estimate the relative costs over 20 years of different hip prostheses.* Journal of Epidemiology & Community Health, 1999. **53**(9): p. 542-7.
89. Faulkner A. Kennedy L G. Baxter K. Donovan J. Wilkinson M. Bevan G, *Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model*, in *Health Technology Assessment - NHS R&D HTA Programme*. 1998.
90. Fitzpatrick R. Shortall E. Sculpher M. Murray D. Morris R. Lodge M. Dawson J. Carr A. Britton A. and Briggs A, *Primary total hip replacement surgery: a systematic reviews of outcomes and modelling of cost-effectiveness associated with different prostheses*, in *Health Technology Assessment - NHS R&D HTA Programme*. 1998.
91. Marinelli M, S.A., Panfoli N, de Palma L.,, *Cost-effectiveness of cemented versus cementless total hip arthroplasty. A Markov decision analysis based on implant cost.* J Orthopaed Traumatol,, 2008. **9**: p. 23-28.
92. Spiegelhalter, D.J. and N.G. Best, *Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling.* Statistics in Medicine, 2003. **22**(23): p. 3687-3709.
93. Givon, U., et al., *Cost-Utility Analysis of Total Hip Arthroplasties: Technology Assessment of Surgical Procedures by Mailed Questionnaires.* International Journal of Technology Assessment in Health Care, 1998. **14**(04): p. 735-742.
94. Unnanuntana, A., et al., *Cementless Femoral Prostheses Cost More to Implant than Cemented Femoral Prostheses.* Clinical Orthopaedics and Related Research, 2009. **467**(6): p. 1546-1551.
95. Yates, P., et al., *The relative cost of cemented and uncemented total hip arthroplasties.* Journal of Arthroplasty, 2006. **21**(1): p. 102-105.
96. Scheerlink, T., W. Duquet, and P.-P. Casteleyn, *Socioeconomic aspects of total hip arthroplasty. A one-year survey in a Belgian university hospital.* Acta Orthop Belg, 2004. **70**(6): p. 525-33.

97. Pingsmann, A., R.T. Muller, and A. Goller, *Cost analysis for total hip arthroplasty by measurement of time and material expenditure.* Archives of Orthopaedic and Trauma Surgery, 1998. **117**(8): p. 421-424.
98. Metz, C.M. and A.A. Freiberg, *An international comparative study of total hip arthroplasty cost and practice patterns.* Journal of Arthroplasty, 1998. **13**(3): p. 296-298.
99. Murray D W, C.A.J., Bulstrode C J, *Which primary total hip replacement?* J Bone Joint Surg, 1995. **77**(Br): p. 520-7.
100. Boardman, D.L., J.R. Lieberman, and B.J. Thomas, *Impact of declining reimbursement and rising hospital costs on the feasibility of total hip arthroplasty.* Journal of Arthroplasty, 1997. **12**(5): p. 526-534.
101. National Institute for Health and Clinical Excellence, *The Guidelines Manual 2009*. 2009. p. Appendix H, Methadology checklist: economic evaluations.
102. Maniadakis, N. and A. Gray, *Health economics and orthopaedics.* J Bone Joint Surg Br, 2000. **82-B**(1): p. 2-8.
103. Black, N., *Why we need observational studies to evaluate the effectiveness of health care.* BMJ, 1996. **312**(7040): p. 1215-1218.
104. Donaldson C. Mugford M. and Vale L, *Evidence-based Health Economics*. 2002, London: BMJ Books.
105. Sculpher, M., M. Drummond, and M. Buxton, *The iterative use of economic evaluation as part of the process of health technology assessment.* J Health Serv Res Policy, 1997. **2**(1): p. 26-30.
106. Palmer, S. and J. Raftery, *Economics notes: Opportunity cost.* BMJ, 1999. **318**(7197): p. 1551-1552.
107. Jameson, S. and M.R. Reed, *Payment by results and coding practice in the National Health Service: The importance for orthopaedic surgeons.* J Bone Joint Surg Br, 2007. **89-B**(11): p. 1427-1430.
108. Department of Health, *National Tariff,*. 2008/9, Department of Health.
109. Briggs, A., *Economic evaluation and clinical trials: size matters.* BMJ, 2000. **321**(7273): p. 1362-1363.
110. Briggs, A., K. Claxton, and M. Sculpher, *Decision Modelling for Health Economic Evaluation*. 2006, Oxford: Oxford University Press.
111. Cleves, M., et al., *An Introduction to Survival Analysis Using Stata*. 2nd Edition ed. 2008, Texas: STATA Press.
112. Briggs, A., et al., *The use of probabilistic decision models in technology assessment: the case of total hip replacement.* Appl Health Econ Health Pol, 2004. **3**(2): p. 79-89.
113. Briggs, A., et al., *Modelling the cost-effectiveness of primary hip replacement: how cost-effective is the Spectron compared to the Charnley prosthesis?* 2003, CHE Technical Paper 28, .
114. National Institute for Clinical Excellence, *Full guidance on the selection of prostheses for primary total hip replacement*. 2000, London, .
115. Swedish Hip Athroplasty Register, *Annual Report 2007*. 2008, Department of Ortopedics, Sahlgrenska University Hospital, .
116. National Joint Registry for England and Wales, *Trend towards cementless hips continues*, in *Joint Approach - the newsletter of the National Joint Registry*. 2009.
117. Yates, P., et al., *The relative cost of cemented and uncemented total hip arthroplasties.* J Arthroplasty, 2006. **21**(1): p. 102-105.
118. Fitzpatrick, R., et al., *Primary total hip replacement surgery: a systematic reviews of outcomes and modelling of cost-effectiveness associated with different prostheses.* Health Technol Assess, 1998. **2**(20): p. 1-64.

119. Daellenbach, H., et al., *Economic appraisal of new technology in the absence of survival data - the case of total hip replacement.* Soc Sci Med, 1990. **31**(12): p. 1287-1293.
120. Baxter, K. and G. Bevan, *An economic model to estimate the relative costs over 20 years of different hip prostheses.* J Epidemiol Community Health, 1999. **53**(9): p. 542 - 547.
121. Marinelli, M., et al., *Cost-effectiveness of cemented versus cementless total hip arthroplasty. A Markov decision analysis based on implant cost.* J Orthopaed Traumatol, 2008. **9**: p. 23-28.
122. Robertson R & Thorlby R, *Patient Choice*, The King's Fund - Brieifing, Editor. 2008.
123. NHS Choices, *NHS Choices Annual Report 2010*. 2010.
124. Department of Health. *Liberating the NHS: Greater choice and control. A consultation on proposals,*. 2011.
125. Morgan. W., K.M., & Rosen. H,, *Microeconomics*. 2nd Edition ed. 2009, London: McGraw Hill.
126. Gravelle H & Rees R, *Microeconomics,*. Third edition ed. 2004: FT Prentice Hall.
127. Greener I, *Are the assumptions underlying patients choice realistic?: a review of the evidence,.* British Medical Bulletin,, 2007: p. 1-10.
128. Smith, P.C., et al., *Principal-agent problems in health care systems: an international perspective.* Health Policy, 1997. **41**(1): p. 37-60.
129. Dixon A, et al., *Patient Choice - How patients choose and how providers respond,*, The King's Fund, Editor. 2010.
130. Pakvis, D., et al., *Is there evidence for a superior method of socket fixation in hip arthroplasty? A systematic review.* International Orthopaedics, 2011: p. 1-10.
131. Office of Fair Trading, *Assessing the impact of public sector procurment on competition,*. 2004.
132. Office of Fair Trading, *The Pharmaceutical Price Regulation Scheme - an OFT market study.* 2007.
133. The Department of Health, *Liberating the NHS: legislative framework and next steps.* 2010.
134. Wright. O. -The Independent - UK. *Coalition may backtrack on NHS reform plans.* 2011.
135. Department of Health, *The procurement of consumables by NHS acute and Foundation trusts,*, National Audit Office, Editor. 2011.
136. Department of Health and Association of the British Pharmaceutical Industry, *PPRS: The study into the extent of competition in the supply of branded medicines to the NHS,*. 2002.
137. Agrawal M, *Global competitiveness in the pharmaceutical industry - The effect of national regulatory, economic, and market factors,*. 1999: Pharmaceutical Products Press - The Haworth Press, Inc.
138. European Commission - Competition. *Pharmaceuticals - Case Law,*. 2011; Available from: http://ec.europa.eu/competition/sectors/pharmaceuticals/case_law.html.
139. The Independent - Business. *AstraZeneca raided in EC competition investigation,*. 2010; Available from: http://www.independent.co.uk/news/business/news/astrazeneca-raided-in-ec-competition-investigation-2151127.html.
140. The European Commission - Competition. *Pharmaceuticals - overview - intellectual property and antitrust,*. 2011; Available from: http://ec.europa.eu/competition/sectors/pharmaceuticals/overview_en.html.
141. Stephen A. *Abuse of Indigestion: Whistleblower in TV Programme Triggers OFT Investigation,*. Compeition Policy Blog 2010.
142. No Free Lunch. 2011; Available from: http://www.nofreelunch.org/.
143. Newman, M., *Bitter pills for drug companies.* BMJ. **341**.
144. Smith R, *Doctors and the drug industry: too close for comfort,.* British Medical Journal,, 1986. **293**(11).
145. Commission of the European Communities, *Regulation (EC) No. 139/2004 Merger Procedure*. 2006: Luxembourg.

146. European Commission - Consumer Affairs, *Exploratory Process on the Future of the Medical Devices - Potential themes for further reflection at the European level and Issues identified by the Members.* 2010.
147. Value in Health, 2009. **12**(4).
148. Silverman E. *J&J's DePuy settles kickback charges,*. 2007.
149. Silverman E. *Undisclosed conflicts among Docs and device makers*. 2010.
150. Belleflame P & Peitz M, *Industrial Organization Markets and Strategies,*. 2010, Cambridge: Cambridge University Press,.
151. Office of Fair Trading, *Assessment of market power - understanding competition law,*. 2004,.
152. Motta M, *Market power and welfare: Introduction,*, in *Competition Policy - Theory and Practice,*, Cambridge University Press, Editor. 2004. p. 39.
153. Harrington J.E, J., *How do Cartels Operate?* Foundations and Trends in Microeconomics 5,, 2006. **2**(1).
154. Ivaldi M, et al., *The economics of tacit collusion,*, in *Final report for DG Competition, European Commission.* 2003.
155. Joint Replacement Instrumentation Orthopaedics. *About us,*. 2011.
156. Stryker Howmedica Osteonics, *Fact book*. 2009.
157. Zimmer, *Annual Report*. 2009.
158. Biomet, *Annual Report - Embracing Change*. 2006.
159. B Braun Medical Limited, *Annual Report*. 2009.
160. Smith & Nephew, *Annual Report,*. 2008.
161. Corin, *Annual Report*. 2009.
162. Motta M, *The assessment of market power,*, in *Competition Policy - Competition and Practice,*. 2004, Cambridge University Press,. p. 124.
163. Khemani R S & Shapiro D M, *Glossary of Industrial Organisation Economics and Competition Law,* , Directorate for Financial Fiscal and Enterprise Affairs - OECD, Editor. 1993.
164. The Competition Commission & The Office of Fair Trading, *Merger Assessment Guidelines,*. 2010.
165. Carrington, N.C., et al., *The Exeter Universal cemented femoral component at 15 to 17 years: AN UPDATE ON THE FIRST 325 HIPS.* J Bone Joint Surg Br, 2009. **91-B**(6): p. 730-737.
166. Hallan, G., et al., *Medium- and long-term performance of 11 516 uncemented primary femoral stems from the Norwegian arthroplasty register.* J Bone Joint Surg Br, 2007. **89-B**(12): p. 1574-1580.
167. Nevo, A., *Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry.* The RAND Journal of Economics, 2000. **31**(3): p. 395-421.
168. Berry, S.T., *Estimating Discrete-Choice Models of Product Differentiation.* The RAND Journal of Economics, 1994. **25**(2): p. 242-262.
169. Widman and Isacson, *Lateral position reduces blood loss in hip replacement surgery: a prospective randomized study of 74 patients.* International Orthopaedics, 2001. **25**(4): p. 226-227.
170. The National Institute of Clinical Excellence, *IPG363 Minimally invasive total hip replacement: guidance.* 2010.
171. Hinsche, A.F. and R.M. Smith, *Image-guided surgery.* Current Orthopaedics, 2001. **15**(4): p. 296-303.
172. National Joint Registry, *4th Annual Report*. 2006.
173. Monitor - Independent Regulator of NHS Foundation Trusts. *NHS Foundation Trust Directory*. [cited 2010; Available from: http://www.monitor-nhsft.gov.uk/home/about-nhs-foundation-trusts/nhs-foundation-trust-directory.

174. NHS Choices *About the NHS.* 2010; Available from: http://www.nhs.uk/NHSEngland/thenhs/about/Pages/authoritiesandtrusts.aspx.
175. Cameron, A., C. and Trivedi, P, K., *Microeconometrics using Stata*. Revised Edition ed. 2009: Stata Press.
176. Cameron, A., C. and Trivedi, P, K., *Microeconometrics: Methods and Practice*. 2005: Cambridge University Press.
177. statalist (STATA). *Re: st: fixed effects tobit model,.* 2003; Available from: http://www.stata.com/statalist/archive/2003-02/msg00270.html.
178. Princeton University. *Descriptive statistics using Excel and STATA,.* 2011; Available from: http://www.princeton.edu/~otorres/Excel/excelstata.htm.
179. FAQs, S.-. *How can I get robust standard errors for tobit?* 2005; Available from: http://www.stata.com/support/faqs/stat/tobit.html.
180. Laudicella M., R.O.K., and Street. A.,, *What explains variation in the costs of treating patients in English obstetrics specialities?* Centre for Health Economics Research Paper,, 2009(49).
181. National Institute for Clinical Excellence. *Selection of prostheses for primary total hip replacement.* 2011; Available from: http://guidance.nice.org.uk/TA2.
182. Channel 4, *The truth about going under the knife*, Dispatches, Editor. 2011.