Outcome Assessment of Total Hip Arthroplasty in The Netherlands and Sweden

Ostendorf, Marieke
Outcome Assessment of Total Hip Arthroplasty in The Netherlands and Sweden

Thesis Utrecht University - with ref.-with summary in Dutch.
ISBN 90-393-3888-4

The cover shows a photograph of the ornament on the hood of a Rolls Royce owned by the late David Cecil (1909-1981), Lord Burghley, 6th Marquess of Exeter, who won the gold medal for the 400 meter hurdles at the Olympic Games in Amsterdam, 1928. In later life he received a Moore prosthesis to treat the hip osteoarthritis he was suffering from. When the prosthesis had to be revised, he had it cleaned, put in silver and had it placed on the hood of his Rolls Royce.

Design & layout Multimedia, UMC Utrecht
Printed by Zuidam & Uithof BV

Financial support for the reproduction of this thesis was generously funded by the following institutions/companies

Arsis Medical
Bauerfeind
Biomet
Dept. Orthopaedics, Institute of Surgical Sciences, Gothenburg University, Sweden
Endo Plus
Mathys
Nederlandse Orthopaedische Vereniging (NOV)
Nederlandse Vereniging voor Biomaterialen en Tissue Engineering (NBTE)
Ortho Biotech
Smith and Nephew
Tornier
Zimmer

Outcome Assessment of Total Hip Arthroplasty in The Netherlands and Sweden

Outcome Assessment van Totale Heupprothese operaties
in Nederland en Zweden

(met een samenvatting in het Nederlands)

Proefschrift
Ter verkrijging van de graad van doctor aan de Universiteit Utrecht, op gezag van de Rector Magnificus, Prof. Dr. W.H. Gispen, in gevolge het besluit van het College voor Promoties in het openbaar te verdedigen op woensdag 8 december 2004 des ochtends om 10.30 uur

door

Marieke Ostendorf
Geboren 16 februari 1975 te Almelo

Promotores Prof. dr. A.J. Verbout
Faculteit Geneeskunde, Universiteit Utrecht

Prof. dr. A.J.P. Schrijvers.
Faculteit Geneeskunde, Universiteit Utrecht

Co-promotor Dr. W.J.A. Dhert
Faculteit Geneeskunde, Universiteit Utrecht

The research described in this thesis was carried out at the Department of Orthopaedics of the University Medical Centre in Utrecht, The Netherlands and at the Department of Orthopaedics, Institute of Surgical Sciences, Sahlgrenska Academy at Göteborg University, Sweden.

The project was financially supported by Dutch Healthcare Research (ZonMW).

This thesis is based upon the following publications

1 Ostendorf M, Johnell O, Malchau H, Dhert WJA, Schrijvers AJP, Verbout AJ. The epidemiology of total hip replacement in The Netherlands and Sweden: present status and future needs. Acta Orthopaedica Scandinavica 2002;73(3):282-6.

2 Ostendorf M, Matthijssen MAH, Dzaferagic A, Schrijvers AJ, Dhert WJA, Verbout AJ. Waiting for a hip: waiting lists and waiting times in orthopaedics [in Dutch]. Medisch Contact 2000; 55(12):416-8.

3 Ostendorf M, Buskens E, van Stel HF, Schrijvers AJP, Marting L, Dhert WJA, Verbout AJ. Waiting for total hip arthroplasty: Avoidable loss in quality time and preventable deterioration. J Arthroplasty 2004; 19(3):302-9.

4 Ostendorf M, van Stel HF, Buskens E, Schrijvers AJP, Marting LN, Verbout AJ, Dhert WJA. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. J Bone Joint Surg Br 2004; 86(6):801-8.

5 Ostendorf M, Malchau H, Schrijvers AJP, Verbout AJ, Dhert WJA. Indications for THA in Sweden and The Netherlands: Surveying Swedish and Dutch orthopaedic surgeons. Conditionally accepted Acta Orthopaedica Scandinavica 2004.

6 Ostendorf M, Eisler T, Hertberts P, Fleer A, van der Tweel I, Dhert WJA, Malchau H. Trends and risk factors in revision THA because of deep infection: a review of 960 first revisions from the Swedish National Hip Registry. In manuscript 2004.

Contents

Chapter 1	**Introduction and aims**	9
Chapter 2	**The epidemiology of total hip arthroplasty in The Netherlands and Sweden:** *Present status and future needs*	49
Chapter 3	**Indications for THA in The Netherlands and Sweden:** *Surveying Dutch and Swedish orthopaedic surgeons*	57
Chapter 4	**Trends and risk factors in revision THA because of deep infection:** *A review of 960 first revisions from the Swedish National Hip Registry*	75
Chapter 5	**Waiting for a hip:** *Waiting lists and waiting times in orthopaedics*	95
Chapter 6	**Waiting for total hip arthroplasty:** *Avoidable loss in quality time and preventable deterioration*	101
Chapter 7	**Patient-reported outcome in total hip in total hip replacement:** *A comparison of five instruments of health status*	115
Chapter 8	**Summary and conclusions**	133
	Samenvatting en conclusies	139
	Acknowledgements	145
	Curriculum Vitae	147

Chapter 1

Introduction to outcome assessment in total hip arthroplasty

Since the introduction of modern hip prosthetic surgery in 1961 by Charnley, total hip arthroplasty (THA) (Figure 1) has dramatically improved quality of life in many patients affected by disease of the hip joint.[1,2] Nowadays, more than one million patients worldwide are treated with a hip replacement annually, and the procedure has evolved as one of the most cost-effective interventions in surgery.[3-6] In the Netherlands, the number of primary THAs has increased with 270% from 6,751 procedures in 1980 until 18,186 procedures in 2000, because of widening of indications for THA due to the excellent results achieved with the procedure, as well as because of an aging population.[7] The most common indication for THA is primary osteoarthritis (OA), a chronic disease that increases in prevalence with age.[8] Apart from primary OA (74%), other hip diseases leading to THA are secondary OA after hip fracture (11.4%), rheumatoid arthritis (4.8%), avascular necrosis of the femoral head (2.9%) and sequelae after childhood diseases (1.5%).[9] In The Netherlands, the prevalence of radiological hip OA is about 15% in females and 11% in males, in the age group over 70 years of age.[10] The socioeconomic burden attributable to OA is tremendous. Patients report significant impairment in their ability to perform activities of daily living and in their quality of life.[11] The psychosocial dimension is also altered, with symptoms of depression having been reported.[12] As a consequence, OA results in high healthcare costs[13] as well as non-medical costs such as those related to work

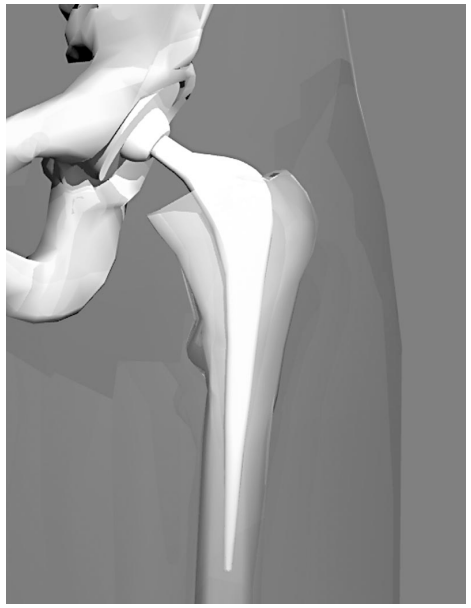


Figure 1
Total hip replacement using bone cement. (© JJ Verlaan, 2004)

disability.[14] During the past 15 years, the total direct and indirect costs of musculoskeletal diseases have risen, accounting for up to 1-3.5% of the gross national product of countries like Australia, Canada, the United States, the United Kingdom or the Netherlands.[15,16] However, the burden of arthritis on society remains underappreciated.[17] In the United States, a study investigating the relationship between funding by the National Institute of Health (NIH) and the burden of diseases included neither arthritis nor any other musculoskeletal condition.[18] In line with demographic changes and in view of the poor awareness of the burden of arthritis, the United Nations, the World Health Organisation and National governments have declared the years 2000-2010 as the 'Bone and Joint Decade' to draw attention to the increasing impact musculoskeletal conditions will have on world health as life expectancy increases.[15,19]

History

During the late 1800s and early 1900s, several attempts were made to replace the hip joint with interpositions of organic or inorganic material. The first artificial hip implant can be attributed to Thermistokles Gluck, who implanted in 1891 in Berlin a femoral head of ivory, fixed with plaster of Paris and glue.[20,21] The results of these experiments were poor, because of severe infection problems and foreign body tissue reactions. As of 1923, Smith-Petersen designed several cups made of respectively glass, Pyrex (a special glass) and Bakelite. The glass cups broke easily while the Bakelite produced too much foreign body reaction. In 1937 he introduced a cobalt-chrome (Vitallium) arthroplasty, resurfacing the femoral head.[22,23] In 1950, the Judet brothers designed a stemmed femoral acryl prosthesis, but the prosthesis often fractured and caused foreign body reactions.[24] Moore developed a cobalt-chrome, stemmed, intramedullar, femoral component in 1957.[25] All these prostheses were hemi-arthroplasties. With hemi-arthroplasties becoming popular for the treatment of intracapsular hip fractures, it was logical to expand the concept by including an acetabular component. One of the first total hip arthroplasties was introduced by McKee and Watson-Farrar, in 1966, who used an all-metal acetabular component combined with a metal stem.[26] Long term results of the metal-on-metal McKee-Farrar prosthesis showed insignificant wear of the articulating surfaces.[27] Metal-on-metal articulations are receiving renewed attention during the last decade.[28] When McKee used acrylate cement to fix the Moore stem into the femur, he achieved good durability in a majority of patients at 5-10 years.[26]
Sir John Charnley (1911-1982) recognized the importance of low friction in the artificial joint and reduced the diameter of the femoral head in his stainless steel prosthesis to reduce frictional torque at the implant-host interface.[29] His attention was called to the possibilities of using polymethylmethacrylate (PMMA) cement by Leon Wiltsie and Charnley quickly adopted it.[30] Charnley also introduced the use of

high molecular weight polyethylene (HMWPE) cups, after experiments with polytetrafluorethylene (PTFE) which showed many problems with wear and foreign body reactions.[31] This THA combination was inserted for the first time in November 1962, and although the design was modified since then, it has ever since provided long-term pain relief and improved function for vast numbers of patients world-wide.[32] Since Charnley, many THA designs have been introduced with different shapes, materials and fixation techniques. Most have undergone several modifications over time in order to solve problems with early loosening and hardware failure. The effect of these modifications is difficult to evaluate because several alterations were usually made simultaneously. As a rule, implants did well in the hands of their inventors but yielded worse results when used in common practice. This proves that apart from implant design, surgical technique is an important parameter in the success rate of THA.[32] For instance, revision percentages of the Charnley prosthesis in specialist centres are about 5% at 10 years[33,34] and 10% at 20 years, while the revision percentage at 5 years of a Charnley prosthesis from a regional register in England was only 8%.[35,36] Apart from the Charnley prosthesis, similar and sometimes even better results have been achieved with several other prosthesis designs like the Lubinus SP-II, the collared Spectron EF and the tapered, highly polished Exeter prosthesis.[9,37]

Indications for THA

The Dutch consensus statement on the indication for THA states that candidates for THA are patients, who - in spite of adequate conservative treatment - are that much disabled by pain and stiffness of the hip joint that impairment in function and dependency on others result.[38] Further, the consensus statement emphasizes the importance of considerable joint damage on radiographs. However, it is unclear whether and to what extent this guideline is applied in clinical practice.
Several studies found that patients with greater pain and dysfunction preoperatively, i.e., at a later point in the natural history of functional decline of the disease process, did not improve to the level achieved by those with higher preoperative function.[39,40] Therefore, traditional orthopaedic practice, to delay surgery as long as possible, might need to be re-evaluated in view of these findings, especially in elderly patients.[41]

Results after THA

The complication rate after THA, when operated with contemporary techniques is low. Because of the total number of operations, however, complications associated with THA will still affect a relative large number of patients. During the first three decades of THA surgery, the infection rate has been reduced from 9% to less than 1%,[42] and due to improved materials and design, implant fractures are rare. Today,

aseptic loosening is the main cause for late implant failure, requiring revision of the prosthesis.[3] Revision surgery is more difficult, more costly, has more complications and generates worse results than primary surgery does.[43,44] The main reason is structural bone loss, suboptimal soft tissues and increased infection risk.[32] The fraction of revisions in relation to primary plus revision operations is a crude measure by which the quality of primary THA can be compared internationally.[45] In the Netherlands, revision burden in the period 1986-2000 was around 11%.[7] The revision burden in other countries varies from 8% (Sweden), 15% (Norway, Denmark, England) to 18% in the USA and 24% in 8% Finland.[45]

Aseptic loosening

In the majority (76%) of revision operations, the reason for revision is aseptic loosening of the prosthesis.[46] Aseptic loosening is often preceded by periprosthetic osteolysis, which at present is believed mainly to be caused by particulate wear debris from the prosthesis. At first, earlier studies suggested that osteolysis was a 'cement disease', because PMMA was (and still is) thought to cause bone necrosis repaired by fibrosis.[47,48] However, the newer uncemented implants eventually developed even more severe osteolysis with similar histological patterns in the osteolytic lesions as compared to cemented implants.[49] Regular follow-up of THA is important to diagnose osteolysis in an early stage, before bone loss has increased to an extent at which revision becomes more difficult. It has been shown that the use of antibiotic bone cement prevents the number of revisions performed for aseptic loosening.[50,51] This at least suggests that the current incidence of infections is underestimated by conventional diagnostic methods and that a certain number of revisions is probably caused by low-virulent infection that is mistaken for aseptic loosening.[52] However, because of low sensitivity and specificity of diagnostic tests, the differentiation between aseptic loosening and low-grade infection remains a difficult problem.[42]

Septic loosening

In about 7% of the cases, revision of THA is performed because of deep infection of the prosthesis.[46] Deep infection after THA is a devastating complication, and remains a challenging diagnostic and therapeutic dilemma associated with decreased patient satisfaction, significant morbidity and mortality.[53] Improved surgical procedures and antibiotic prophylaxis have contributed to the current low infection rate in primary THA (0.6-1.3%).[54] However, the incidence of infection in THA will probably increase due to an increasing use of THA in elderly patients, who have an increased infection risk.[32]
Pathogenesis. There are three routes by which microorganisms can reach the periprosthetic space: contamination at the time of surgery, migration to the joint from a superficial wound infection, and haematogenous inoculation by spread from a distant focus.[42] The main source of bacteria is the skin in both patients and operating

personnel. The type of infection is usually mono-bacterial, and the most common bacteria (>50-70%) are Gram-positive aerobic bacteria (coagulase negative staphylococci, *Staphylococcus aureus* and streptococci). Gram-negative species represent about 20% and anaerobes about 10%.[42,55,56] Bacterial adhesion to the prosthetic surface is a crucial event in implant infection. This mechanism has been called 'the race for the surface', indicating the competition over the implant surface between host cells and bacteria.[57] Certain types of coagulase negative staphylococci produce a protective polysaccharide coating (glycocalyx), which increases their ability to persist on the prosthetic surface.

Risk factors for infection. Several patient-related factors affect the susceptibility to contract wound complications, such as obesity, rheumatoid arthritis and diabetes mellitus.[42] Also, circumstances in the operating theatre are important in preventing infection. Many efforts have been undertaken to minimize intraoperative inoculation by clean-air operating theatres with laminar air-flow, and exhaust ventilation suits, and by measures such as the frequent changing of gloves and suction tips.[58] Bacteraemia caused by dental surgery or urinary tract infections may cause bacterial seeding to the THA.[59]

Antibiotic prophylaxis. To prevent infection after THA, parenteral antibiotics are used intraoperatively or in bone cement. A study by Hill et al. demonstrated that intravenous cefazolin reduced the infection rate from 3.3 to 1.9% in conventional operating theatres.[60] Nowadays, mostly cephalosporins and flucloxacillins are used in antibiotic prophylaxis. A study from the Norwegian Hip registry found the lowest infection rate for the combination of systemic antibiotics and antibiotic containing bone cement, irrespective of ventilation of the operating theatre.[50]

Antibiotics in bone cement partly bypass biological inactivation and reduce toxic side effects. The cement eludes high amounts of antibiotics during the first days, after which the amount rapidly decreases during 2 weeks to 2 months.[61] However, small concentrations are still measured at 5 years, which could increase the risk of bacterial resistance.[62] Because of the high initial local concentration of antibiotics even bacteria defined as resistant may be eradicated.[63] The antibiotic most commonly used in bone cement is gentamycin, because of its good release properties, its thermal stability during the polymerization of the bone cement, and because of its coverage of the most important pathogens.[64]

The problem of antibiotic resistance. The diminished sensitivity of common staphylococcal pathogens to routine antibiotics has caused alarm among surgeons. In a nation-wide American study of resistant organisms in nosocomial infections in patients in Intensive Care Units, as many as 46.7% of Staphylococcus aureus strains and 85.7% of coagulase-negative staphylococci strains were methicillin resistant.[65] Coagulase-negative staphylococci are frequently found in infections of invasive and indwelling medical devices.[66] Bacteria within a biofilm on a plain polymer are several hundred times more resistant to antibiotics.[67] There are concerns that the use of

systemic antibiotic prophylaxis and antibiotic containing bone cement can lead to the development of antibiotic resistance.[62,68] Sanzén and Walder found an increased resistance to gentamicin in patients after joint replacement in most of whom gentamicin containing bone cement was used.[69] An experimental study in rats showed an increased resistance of Staphylococcus epidermidis to gentamicin after subcutaneous implantation of gentamicin containing bone cement.[70] However, gentamicin resistance has not shown to be more common in countries using bone cement containing antibiotics.[71] The problem of increasing multiresistant CNS has also been found in other medical fields, especially in infections of vascular prostheses and intravenous catheters.[72,66,73] There is also evidence that resistant skin flora can develop as a consequence of antimicrobial prophylaxis given.[74] In a large European study, a correlation was found between antimicrobial resistance of S. pneumoniae to penicillin and the use of beta-lactam antibiotics and macrolides.[75] Furthermore, in the development of resistant bacteria, it should also be remembered that more than half of the total production of antimicrobial agents is used in animal husbandry.[76] There is also evidence that multiresistant strains of CNS are often transferred from hospital staff to patients during the hospital stay.[74,77] In several European countries, there is increasing attention for the problem of antibiotic resistance in human and veterinary medicine, leading to international surveillance programs and research efforts.[78,79]

Diagnosis. Classic symptoms of infection are not always present, but most patients present with constant pain in the hip. Physical examination includes the evaluation of the hip area for any signs of inflammation, swelling, warmth, erythema, localized tenderness and wound inspection for any evidence of drainage or persistent sinuses.[42] Other investigations include the comparison of serial good quality radiographs, white blood cell count, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) and occasionally, nuclear imaging and intraoperative (frozen) sections. The most effective test to promptly identify the infecting organism is a hip aspiration, preferably under fluoroscopy to document the intra-articular location of the needle. Varying specificity and sensitivity have been reported for laboratory testing in infected THA.[80] Radionuclide imaging scans have high sensitivity but low specificity, as other processes such as bone remodeling, inflammation and fractures give uptake as well.[81] Intraoperative biopsies, either for frozen sections or conventional processing, have shown rather high specificity but lower sensitivity.[80,82]

Intraoperative tissue biopsies for culturing should be taken after discontinuation of any systemic antibiotics. However, in cemented THA, antibiotic containing cement may leak antibiotics from the cracked cement, even up to 10 years after primary surgery, and thus cause negative cultures.[83] Bacterial cultures can be problematic for several reasons. Sensitivity may be poor in bacteria demanding enriched or prolonged cultures. Low-virulent bacteria may have an altered metabolism and may require different culturing techniques.[84] To improve diagnostic accuracy, molecular biological techniques are currently being developed. The most established method for this

purpose uses polymerase chain reaction (PCR) to amplify specific bacterial DNA-sequences.[64,85] The method has several advantages: It requires only bacterial DNA, not living bacteria, it is relatively fast and it captures bacteria irrespective of their growing characteristics. However, high sensitivity is still the major drawback of PCR, because contaminants will be amplified as well.

Treatment. In most cases, in chronic infections, removal of the components is required to eradicate infection. Current options for treatment include lifelong antibiotic suppression, debridement with prosthesis retention and subsequent antibiotic treatment, insertion of a new prosthesis directly after debridement (one-stage revision) or after an interval (mostly months) of intensive antibiotic treatment (two-stage revision). Other options include permanent resection of the prosthesis, and (rarely) arthrodesis and amputation.[42] Selection of the antibiotic type, and of dose and length of treatment must be carefully made, based on the results of intra-operative cultures. A minimum post-peak serum bactericidal titer of 1:8 has been suggested.[86] Most studies employ a regimen of 4 to 6 weeks of intravenous treatment, and in two-stage revisions systemic therapy less than 4 weeks has been associated with a poor prognosis.[87] Local antibiotics, in the form of antibiotic beads or cement spacers, are often used in two-stage revisions. Spacers also help to mobilize the patient, and to prevent contractures of hip musculature between stages.[88-90] Debridement and retention of components can afford successful results in patients with early infections or sudden onset of symptoms in a previously well-functioning hip.[56] However, in patients with symptoms for a longer period (> 2 weeks) results were generally poor.[91] Successful outcome after one-stage revision depends on careful patient selection, a pre-operatively diagnosed mono-bacterial low virulent infection, and little bone loss.[32] Surgery should be performed with meticulous debridement and with antibiotic containing bone cement.[55,92] Because of current problems with resistant bacteria and the inability to identify the bacterial identity before insertion of a new prosthesis, one-stage revision is less applicable than it once was.[93] The advantages of the method lie in decreased morbidity and lower costs. A recent review of the literature showed that 83% of the patients treated with one stage revision were infection free after a follow-up of 4.8 years.[94] Two-stage revision is a safe and effective method to eradicate infection in THA, with reported success percentages around 91-93%.[55,95] The advantage of a staged procedure is that a more accurate bacterial diagnosis can be obtained, and that extended treatment with local antibiotics and debridement is possible before reimplantation. The most appropriate time interval between resection and reimplantation remains unclear,[87] but most studies employ an interval of about 3 months. The use of antibiotic containing bone cement is recommended.[55] Several studies have designed staging systems to optimize treatment for infected THA.[53,96,97]

Quality improvement: the importance of stepwise introduction of new techniques

Since the introduction of Charnley's low friction hip arthroplasty, limited progress has been booked in the development of new, better prostheses. Therefore, the importance of the stepwise introduction of new techniques in total hip arthroplasty has been emphasized.[98] A new prosthesis should only be introduced after thorough preclinical testing (materials, wear, mechanics), randomized and prospective trials, multicenter studies and, preferably, regional register studies.[3] Several reports from the English National Audit Office (NAO) stressed the importance to use well-established prosthetic designs with modern surgical techniques.[99,100] Since 1993, new prostheses need to be CE-marked before they are allowed on the market, but in practice this means that new implants only have to fulfil basic criteria regarding safety and design. No documented proof of long-term success of the prosthesis is required. Because of the low failure rates currently achieved in conventional THA, it is effectively impossible to show that a new design is significantly better than conventional ones. With current failure rates of about 1% a year, a new design that shows 30% improvement over an older one (a truly radical improvement) would require a trial of many thousands of patients followed up for at least a decade before a significant difference could be shown.[101] Therefore, Bulstrode et al. pleaded for the use of an implant registry to allow identification of prostheses, outcome measures more sensitive to failure than revision and better ways of assessing early how well an implant is functioning. The orthopaedic medical device industry regularly introduces new prosthetic designs, in an attempt to improve the performance of a prosthesis, which as a consequence will improve the market position of the manufacturer. New implants are often more expensive because of development, advertising and marketing costs. Several studies assessed whether the higher price of a new prosthesis can be justified by lower revision rates, but this seems seldom the case because of the low revision rates already achieved with the old systems, especially in elderly patients.[102] Therefore, Healy pleaded for an 'implant matching system' which matches patients with certain characteristics to a certain prosthesis type.[103]

Measuring outcome in THA

At the end of the 19th century, E.A. Codman implemented the 'End Result Idea' into his surgical practice.[104] This system was designed to routinely follow-up and document the outcome of various treatments for the stated purpose of improving the effectiveness of patient care. He believed that this type of evaluation should be performed systematically and without bias to provide as accurate a picture of the patient's clinical status as was possible. The importance of the idea was missed by most of Codman's colleagues, primarily because there was little incentive driving them toward a more sophisticated methodological approach to evaluating their clinical

outcomes and to using the results to drive the improvement of quality and efficiency.[105] However, during the past decades, outcome measurement has become increasingly important. The issue of cost containment has proven to be the alarm that has reawakened Codman's ideas. Since World War II, medicine has experienced an incredible growth.[106] Increased affluence, new technologies, and an aging population have led to this unprecedented growth. Transplantation surgery, coronary artery bypass surgery, arthroscopic surgery and total joint replacement are examples of developments over the past 2 to 3 decades. These advances have increased the demand for healthcare and have produced a cost crisis in most countries of the developed world.[107] As such, healthcare interventions have come under increasing scrutiny. Healthcare providers are interested in how good an intervention is and whether it is cost-effective.[108] Surgeons want to know which treatment has most benefit for their patients. The outcomes movement has evolved to answer many of these questions in a scientifically valid manner.[107]

Three factors were pivotal in stimulating the outcomes movement. The first issue, as already mentioned, is the increasing cost of medical care. As a percentage of the gross national product, the cost of care continues to rise.[109] The second issue concerns the appropriateness of care. Several studies have suggested that up to 30% of some surgical procedures are inappropriate.[110] The third factor is the phenomenon of geographic area variation, whereby the rates of surgical procedures vary according to where people live.[111,112] The Darthmouth Atlas of Musculoskeletal Health Care shows substantial regional variations in the use of many orthopaedic procedures performed in the United States.[113] Widely varying rates of surgical procedures raise the question: 'What is the correct rate?'[111] It has been suggested that patients living in areas with a higher rate receive too much, or inappropriate surgery, which is of course of interest to third-party payers and governments.[107]

The outcomes movement

The outcomes movement was defined in several ways by Paul Ellwood in 1988.[114] First, outcomes management emphasizes patient-based outcomes.[108,115] Many evaluations of medical therapy have used objective end-points with little relevance to how patients were affected by treatment. This particular aspect of the outcomes movement, called outcomes assessment, resulted in the development of numerous patient-based outcome measures. The second element of the outcomes movement involves outcome studies at the provider and population level.[116] These studies mostly involve large and more inclusive databases.[114] The perception was that the majority of clinical studies, performed primarily in tertiary care centres, were of limited relevance to practice in the community.[117] This is one of the rationales behind the desire to create national joint arthroplasty registries.[107] Examples of such population based outcome studies are studies on a large sample of Medicare patients, showing that low-volume surgeons tended to have higher mortality rates, more infections, higher rates of revision operations, and more serious complications during hospitalization for THA.[109,118] The third element of the outcomes movement is the development of practice guidelines. Practice guidelines, however, are often inadequate to deal with the complexity of the patient population and yet are not simple enough to allow them to be used and adopted by practitioners.[107] The fourth element of the outcomes movement is a cycle of continuous improvement, whereby practice guidelines based on research with use of patient based outcomes of community-based samples would lead to improved outcomes and a continuous cycle of improved care (Figure 2).[114] The development of clinical pathways in total hip arthroplasty is an example of optimizing patient care and outcomes while reducing costs.[119]

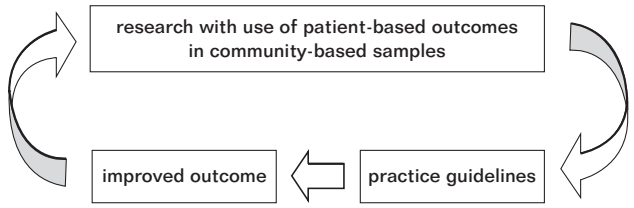


Figure 2
Cycle of continuous improvement using outcome assessment.

'Outcomes is a study of what works'.[120] However, the meaning of outcome assessment is highly dependent on who the observer is and what the incentives and expected benefits of outcome assessment are. For instance, outcomes may mean the fulfilment of long-awaited expectations of pain relief and restoration of function to the hopeful orthopaedic patient. In addition to these desirable outcomes, range of motion or radiographic findings may be of particular interest to the operating surgeon. A governmental approach to outcomes might be concerned with the volume and intensity of services for a condition, and their relationship with morbidity, mortality and rehospitalization. Likewise, private insurers may be primarily interested in outcomes as they determine economic risk.[105] Research, which involves societal aspects of medical technologies, is called medical technology assessment (MTA) or health technology assessment (HTA). HTA studies often compare the relative effectiveness (performance of a technology under routine clinical circumstances) or efficacy (performance of a technology under optimal circumstances) of a medical technology using an adequate control group. Sensitivity analysis should be used to test the robustness of the assumptions tested.

Individual outcome studies, whose methodologies will be discussed in the next paragraphs, are rarely sufficient for clinical decision making.[116] Even the most carefully performed studies, such as randomized controlled trials, leave many questions unanswered about clinical effectiveness.[121] For example, they may be too small to

provide information on effectiveness for important subgroups of patients, or fail to account for all consequences of treatment alternatives such as patient attitudes towards risk, quality of life, and longevity. Finally, they may fail to consider costs. To address these problems and to add value to information provided by individual outcome studies, several data synthesis studies have become increasingly popular.

Meta-analysis. Meta-analysis is a biostatistical tool for combining the results of multiple clinical trials. A large proportion of clinical trials lack sufficient sample size to detect clinical meaningful effects of treatment.[122] With its systematic approach to evaluating clinical studies and combining their results, meta-analysis explores reasons for contradictory findings across clinical studies.[116] A well-conducted quantitative review may resolve discrepancies between studies with conflicting results. However, a systematic review of the methodologies of meta-analyses in orthopaedic surgery showed methodological flaws in 88% of the studies.[123]

Decision analysis. Decision analysis is a 4-step process for evaluating treatment strategies. First, alternative therapeutic strategies are specified, and all important potential outcomes are represented in a decision tree model. Second, probabilities are assigned to each clinical outcome in the model. Third, each possible outcome is assigned a value, or utility such as quality adjusted life expectancy. In the final step, the expected value of each clinical strategy is calculated by multiplying outcome utilities by their probabilities of occurrence, and the stability of the results is tested using sensitivity analysis. Several studies used this type of evaluation for comparative economic appraisal of different total hip prostheses.[102,124-127]

Cost-effectiveness studies. In an era of resource constraints, clinical practice must not only be effective, it must provide benefit at reasonable costs. In cost-effectiveness analysis, the relative value of an intervention as a health investment is defined by its cost-effectiveness ratio, obtained by dividing the net cost of the intervention by its net benefit.[128] Benefits are most often assessed in terms of QALYs, a measure that accounts for the effect of an intervention on both longevity and quality of life. To improve the comparability and quality of studies, consensus-based recommendations guiding the conduct of cost-effectiveness analysis (CEA) have been developed.[129]

Methods to study total hip replacement

Randomized clinical trials (RCTs) remain the gold standard for the evaluation of different interventions. However, randomized trials are not very practical to evaluate the multitude of joint replacement options that are currently available. There are more than twelve manufacturers of hip implants, with more than fifty femoral components. The performance of RCTs to examine the relative effectiveness of all of these systems would be costly and of limited value because of the ongoing redesign of hip implants.[107,130] While RCTs are valuable, they are not optimal for market surveillance and do not serve as an early warning system for implant technologies that fail

prematurely. Meta-analysis, as mentioned before, is another method to study an intervention such as THA. However, when it comes to evaluating the failure rate of a given implant system, this technique is of limited value. Most implant systems do not have a reasonable body of literature on which to perform meta-analysis and there are few studies that compare different implant systems. A systematic review of primary hip replacement systems showed generally low methodological quality in evaluated studies.[126] Retrospective case series are a commonly used method for dissemination of information on the performance of hip arthroplasty components. These reports are frequently from a single surgeon or single (specialist) centre and are not generally applicable to the general orthopaedic community.[3] Additionally, the investigator may be the inventor of the device or may have other conflicts of interest, which introduces the potential for bias. Finally, timely reporting is not typical in this type of study because they depend on the inclusion of a critical number of cases with a substantial period of follow-up.

The Swedish Total Hip Replacement Registry

The Swedish Total Hip Replacement Registry was initiated in 1979 by Herberts and Ahnfelt after a pilot study, which showed that data on reoperations gave valuable information about serious complications after THA in Sweden.[131] The mission of the registry is to improve the outcome of total hip replacement.[132] In the registry, information is collected on all primary total hip replacements in Sweden. Further, information on reoperations, prophylactic measures and surgical technique is collected. All orthopaedic departments in Sweden participate on a voluntary basis. More than 95% of the total hip operations in Sweden are reported to the registry.[133] The aims of the registry are to perform epidemiological analyses of hip replacement surgery in Sweden, identify risk factors for failure of primary and revision surgery, perform bench marking by comparison between regions and hospitals and perform quality assurance of all hip replacements performed in Sweden.[134] The past five years, almost all information exchange between the reporting units and the register has been provided via the Internet (http://www.jru.orthop.gu.se).[45] From 1979 until 2003, 229,031 primary and 21,367 revision THAs have been collected.[46] In the registry, several implant and patient-related factors have been analyzed with estimation of the survival of the implants, depending on age, gender, diagnosis, type of implant and fixation technique. Multivariate Poisson models from the registry showed that male gender and young age increase the risk for revision because of aseptic loosening. As a result of the register, six implant types constitute 70% of the THA market in Sweden, all these implants have long-term documentation.[45] Two key features have contributed to the success of the registry. The first is the Swedish personal identification number, by which patients can be easily followed even if they are operated in different hospitals, and which can be used to add information from other databases, such as the Death Registry. The second feature is the willingness of Swedish orthopaedic sur-

geons to cooperate.[45] Joint replacement registries have several advantages in studying the results of THA. They provide good statistical power, because of the large number of patients involved. Further, they give a more realistic estimate of the results of THA in large patient populations, while the results from highly specialized centres might be misleading and overly optimistic.[107] The Ontario Joint Replacement Registry has assisted in the management of surgical waiting lists: information is available on the number of people waiting, the length of their wait and the severity of disease.[135] Registries also provide feedback of outcome data to surgeons. They allow post-market surveillance of new medical devices and technologies. Together with accurate preclinical testing and limited randomized pilot trials, a register could prevent disasters with new technologies in THA, as experienced with the Christiansen implant and Boneloc cement.[136] The registry's impact has been credited with helping to decrease the revision burden.[107] Estimating the cost of the Swedish THA Registry at $400,000 and the direct cost of a THA revision at $12,000, the registry already becomes cost-effective if the annual number of revisions is reduced by 33 or more.[45]
Several other countries started national joint registries in the past fifteen years, such as Finland and Norway.[137,138] Denmark, New Zealand, Hungary, Australia, Canada and the United Kingdom started more recently.[100,107,139] Also in the United States there have been incentives to start a national joint replacement registry, but liability, compliance, cost and confidentiality are important issues impeding introduction of a registry.[107,140,141] In the Netherlands, no joint replacement registry exists apart from a non-compulsory implant registration system with limited compliance of reporting hospitals.[38]

Measuring outcome of THA: subjective and objective evaluation

Health status can be assessed by a number of methods, which are classified as either subjective or objective.[115] Objective outcome measures in THA include survival analysis of the prosthesis (revision rate), the measurement of radiological changes (plain radiographs as well as radiostereometric analysis (RSA)),[142] and of pain, function and range of movement of the hip joint, which involve a clinician making a judgment about the patient. A subjective outcome measure is essentially an assessment of change, which judges how the patient is now, by his or her own standards, as compared with a previous occasion, such as before surgery.[115]

Objective evaluation of THA

Survival analysis. In 1980, Dobbs introduced the statistical technique of survivorship analysis, originally a technique mainly used in cancer research, in the setting of total hip arthroplasty.[143] Until then, many studies just reported the percentage of prostheses that had to be revised because of failure. Nowadays, survivorship analysis or survival analysis is a commonly used tool to evaluate outcome of total hip replacement. Basically, survival analysis calculates the probability for an implant to survive a

particular length of time. The most common end-point in survival analysis in THA is revision of the prosthesis. Using survival analysis, it is possible to analyze data from patients with different length of follow-up, allowing cases to enter and be withdrawn from a trial at any stage and for whatever reason. The two methods used in survival analysis are the life-table method and the product-limit method or Kaplan-Meier method.[144] The life-table method calculates the cumulative estimate of success at the end of each chosen time interval, usually after each year. With the product-limit method, the cumulative estimate of success is calculated on a daily basis and changes with each failure.
Revision as an endpoint for failure of the prosthesis is a very clear endpoint as no one can deny that such an implant has failed from a scientific perspective.[32] However, indication of revision is relative and depends on the wish of the patient to be operated, on the opinion of the surgeon, the knowledge how to detect loosening and skill to perform a pending revision. Further, revisions also depend on local resources for the operations and the way patients are followed after their primary THA. Several studies advocated the use of other, less crude endpoints for survival analysis, such as pain levels, radiographic failure, outcome instruments and satisfaction of the patient.[117,145] There are some pitfalls using survival analysis. First, a surgical method (a prosthesis but also the cementing technique and skill of the surgeon) must be fairly similar over the observed period, at least not dramatically changed. Further, the number of the patients (prostheses) at risk will be smaller at longer follow-up because of revision, death, loss to follow-up and because of patients operated more recently. This results in an unreliable survival curve, where the failure of only one case can dramatically reduce the survival rate. Therefore Dorey et al. recommended that survival curves are interrupted when less than 20 implants remain and that confidence limits are always included.[146] Another problem is the assumption that patients who are deceased or lost to follow-up were at the same risk of failure as those under continuous surveillance. However, it is shown that patients lost to follow-up have a worse outcome than those who continue to be assessed. A way to deal with this problem is to minimize loss to follow-up and to include a worst-case curve, which is calculated by assuming that patients lost to follow-up had failed implants.[147] Recently, Dorey and Amstutz also pleaded to account for patient activity levels when using survival analysis in the evaluation of THA.[148] Survival analysis cannot be used to assess several covariates influencing survival, such as age, gender and surgical technique. For this reason, the multivariate Cox proportional hazard regression method is commonly used. When assessing time-dependent risk factors or when assessing time as a risk factor on its own, multivariate Poisson regression models should be used.[32]
Radiographs. Conventional radiography is a common method to evaluate THA. The aim is to evaluate stability and wear of the components and to study responses of the surrounding bone over time. Radiographs are important to assess early changes predictive of failure or adverse tissue reactions. A study by Lieberman et al. found a sen-

sitivity and specificity around 92-100% for acetabular and femoral loosening.[149]
To improve the accuracy of radiographic analyses, standardized positioning of the patient, source of radiation and film are very important. For diagnostic standardization, prosthetic regions are divided into three acetabular regions (according to Charnley and DeLee)[150] and into usually seven femoral zones (according to Gruen).[151] These zones are used to describe the cement mantle, fractures, radiolucencies and bone remodelling. Radiolucencies between the cement and bone, occurring postoperatively, have a negative prognostic value for aseptic loosening of the prosthesis.[152] However, radiolucencies are difficult to detect and are subjected to a large inter-observer variability.[153]

Radiostereometric analysis (RSA). Implant micromotions 1-2 years after surgery have a proved predictive ability for later aseptic loosening and revision.[142,154] Radiostereometric analysis (RSA) is currently the only non-invasive method that enables reliable measurement of component migration. The method involves insertion of tantalum markers in the bone, cement and prosthesis followed by sequential radiographic examinations made simultaneously from two directions, where the three-dimensional coordinates of the tantalum markers are measured in relation to fixed calibration markers. Three-dimensional movements can thus be estimated with a precision lying between 0.15-0.6 mm, depending on the technique used.[155] RSA has been used in a research setting, but not in a routine clinical setting.

Other objective outcome instruments. In the early days of THA, outcome of total hip arthroplasty was, apart from evaluation methods like revision rates, range of motion, walking distance and radiographs, mainly measured by retrospective evaluation according the surgeons standards. Examples of objective outcome measures designed in those days are the Harris Hip Score[156] and the Merle d'Aubigne Score (MdA).[157] Other examples of objective outcome measures in THA include the Charnley score[158] and the Lequesne index (L-ISH).[159] It has become increasingly clear that clinical assessment of key aspects of outcome such as pain, physical function and range of joint movement are often inaccurate and not reproducible.[160] They may also overly represent the concerns of the clinician, rather than those of the patient.[115] Several studies have shown that patients and the orthopaedic surgeon do not share the same definition of success.[161-163] The physician's evaluation was regularly more favorable than the patient's, especially in patients with more unfavorable results.[164] Further, from a practical point of view, persuading patients to attend for review, particularly if they are working, live some distance away or are no longer inconvenienced by pain and disability, may prove to be difficult and requires much organization.[115]

Subjective outcome instruments

The subjective measures rely on obtaining responses directly from patients about their perceptions of health and illness. Questions relating to patient satisfaction are another consideration and do not necessarily accord with other measures. These findings, when discordant, can often be explained by a failure to achieve the prior expectations of the patient.[165] The administration of questionnaires is more straightforward to organize than clinical assessments, is less costly and response rates are usually much higher.[105] Most investigators accept that a combination of objective and subjective measures is desirable in order to provide a complete assessment of health-related quality of life and overall outcome.[115]
There are two main types of patient-based outcome measures, namely, disease-specific instruments, which focus on patients' perception in relation to a single condition (such as hip disease) and generic instruments, which are intended to address a wide range of health problems. Disease-specific patient-based outcome instruments used in the evaluation of THA include the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC),[166] the Oxford Hip Score,[167] the McMaster Toronto Arthritis (MACTAR) scale ,[168] and the hip disability and osteoarthritis outcome score (HOOS). Disease-specific scores appear to generate a higher sensitivity to change in outcome of THA as compared to generic instruments.[167] Generic measurements offer the possibility to assess health states over different disease categories, and in some cases to an age- and gender related population sample. A well-known, extensively used and tested generic measure in outcome research is the Short-Form 36 (SF-36).[170] Other examples are the Nottingham Health Profile (NHP) and the Sickness Impact Profile (SIP).[171,172] Some generic questionnaires, so-called multiattribute utility measures, such as the EuroQol (EQ-5D), the Quality of Well-Being index (QWB), SF-6D and the Health Utilities Index (HUI) give the possibility to compare cost-utility between different interventions, which is important in this era of limited health care budgets.[173-176] With this cost-utility approach, the cost of an intervention can be related to the number of quality-adjusted life years (QALYs), that is, a ratio between the cost and the effect of the treatment times the duration of the improvement. However, there has been some debate on the validity of the QALY principle, because the QALY utility values are based on opinions from the general population on a certain health state and might not represent patient preferences.[177,178] Another study showed extensive variation in the construction of QALYs in cost-effectiveness analyses and revealed that most studies did not adhere to practices now recommended by leaders in the field.[179] Adequate and valid methodology should be used if the results of such studies are to be used for purposes of allocating resources. It is recommended to use both a disease-specific and generic measure in the outcome studies of joint replacement surgery.[180] The WOMAC and SF-36 are especially recommended to assess outcome in THA.[105,120]
An outcome instrument, disease-specific as well as generic, should be valid or accu-

rate in its description of a real clinical state. It must be reliable or precise when used on a stable patient, and responsive in demonstrating a quantitative change that is appropriate to the observed clinical change. However, only a limited number of clinical hip osteoarthritis scores has been tested for reliability and validity.[181]
Classically, the response to total hip replacement has been measured by the extent to which the hip score changes. It must be remembered, however, that the noise level, or variation, caused by less than perfect reliability needs to be accounted for in defining the clinically important difference.[105] For instance, when measuring function after joint arthroplasty, other changes in general health (such as comorbidity) will potentially act as interference (or 'noise'), which may have the effect of obscuring the particular outcome of interest.[182-184] Further, the scale with the greatest responsiveness may not always be measuring change which is important to patients.[185] Brinker et al. found several demographic biases in four hip scoring systems that were used in healthy adult subjects without hip disease, such as advanced age, low income and the presence of two or more medical conditions.[186] Ritter et al. found a significant decline in the Harris Hip Score after a follow-up of 10 years after THA, in the absence of implant-related problems. This decline, caused by deterioration in the functional capacity of ageing patients, is an important factor in longitudinal studies using outcome instruments.[187]

Epidemiology of THA

Variations in THA rates

In the 1980s, orthopaedic surgeons became under increasing pressure to defend the use of expensive new technology, including total hip replacement.[188] Quantitative data on the actual need for THA are useful to public and private policy makers in deciding where to deliver this service, in planning the size and direction of orthopaedic training programs, and in estimating the potential economic impact of the procedure. Geographic variations in incidence of THA have been the subject of intense study by third-party payers.[189] High incidences are often attributed to the styles of practice of the surgeons or to a greater number of surgeons in relation to the population in a given area.[112] Interest in geographic variations has increased because of the high total cost of hip arthroplasties and because the number of total hip artroplasties performed has risen in many Western countries during the past decades.[139,190-193]
THA is now also used in patient groups that were formerly considered as too old or too impaired to benefit from surgery.[194] Differences in hip replacement incidence rates, within districts (small-area variations) or between countries, may be due to various causes. These include different coding systems, country-specific differences in the healthcare system, in total expenditure on health per capita and in population structure.[15] Also differences in incidence of underlying disease (osteoarthritis, rheumatoid arthritis), in indication criteria for THA and in patient's willingness to undergo surgery have been reported as explaining variables for the variation in THA incidences.[195,196] One study found that after controlling for population characteristics and access to care, orthopaedic surgeons' opinions or enthusiasm was the dominant modifiable determinant for area variation in the utilization of knee replacement surgery.[112] Incidence rates of THA vary between different age groups, races[197] and gender.[198] Generally, incidences rise with age, are much higher for the white population compared to Asians, Hispanics, and blacks and are higher for women than for men.[199,200] A study by Hawker et al. suggested however an underuse in THA for women compared to men.[198]

Practice guidelines for THA

The stated purpose of clinical guidelines is to 'help healthcare professionals and patients make the right decisions about health care in specific clinical circumstances.' [201] Practice guidelines have been welcomed by those who wish to contain costs by reducing variations in practice. Further, they play a role in improving clinical practice by describing an array of appropriate and inappropriate treatments for well-defined conditions.[105] Initially, most guidelines arise from a consensus of experts together with information from the literature. These guidelines, however, must be modified utilizing outcomes data gathered from many different practice environments involving a variety of populations to ensure generalizability, and to allow the necessary flexibility for patient preference.[114] However, there is an imperfect evidence base to support decisions about which guideline dissemination and implementation strategies are likely to be efficient under different circumstances.[202]
In several countries, consensus statements with respect to total hip replacement have been developed.[38,203,204] Among other topics they discuss indications for THR, pre- and postoperative considerations, surgical technique and expected outcomes. In The Netherlands and Sweden, only consensus statements for the indication for THA exist, based on expert opinion. These are not real guidelines in the sense that they are based on evidence from extensive review of the literature or on information from prospective outcome studies. Nowadays, it is recommended to use evidence-based guidelines instead of consensus-based guidelines. Apart from scientific evidence, guidelines should also deal with cost-effectiveness, patient preferences and availability issues. [201,205,206] Practice guidelines play an important role in the 'best practice cycle' earlier mentioned, whereby practice guidelines based on research with use of patient based outcomes of community-based samples should lead to improved outcomes and a continuous cycle of improved care.[114]

Demand, need and supply for THA: the waiting list problem

The gap between demand for health care services and supply is increasing and needs assessment has become increasingly important in allocating finite resources.[207] The prevalence of disease in a population is indicative of the actual need for health care, but not all these needs will be translated into a demand for care. An English study investigated the population requirement for primary THA in England by self-report screening questionnaires and subsequent clinical investigation in a population sample.[208] They found that satisfaction of demand for THA, given agreed criteria for surgery, is a realistic objective. Several other studies also found an unmet need for THA in the population.[111,196,209]

There is a wide gap between potential need for THA as defined from the point of view of surgical indication and actual need as defined from the point of view of the patient. Two studies showed that, apart from the difference in prevalence of severe arthritis, the willingness of patients to undergo arthroplasty plays an important role in estimating the demand for the procedure.[111,196] Among patients with severe arthritis, no more than 15% were definitely willing to undergo arthroplasty.[196] On the other hand, as patients are getting more informed about treatment possibilities and less inclined to accept disability, demand for THA is expected to increase.[101]

However, the demand for THA still exceeds the supply judging by the considerable waiting lists that exist in countries with a publicly funded health care system.[135,210-212] It has been suggested that shorter waits might be associated with improved patient outcomes and health care savings.[213,214] Others suggest that delays may cause losses in quality of life, characterized by progressive pain and immobility, and may lead to poorer outcomes because of more advanced hip disease when surgery is finally done.[215] It is not always clear whether patients are dealing with an acceptable waiting time or an unacceptable delay in receiving appropriate care.[216] A certain waiting time gives the possibility of reflective second thought or opinion, adaptation to a new condition or trial management with more conservative measures.[217] Canadian data have shown that for knee replacement, patients are satisfied with waiting times averaging four weeks (but not 8 weeks) for specialist consultation and about 8 weeks (but not 32 weeks) for surgery itself.[218] Being on a long waiting list is not necessarily a problem, although being on a long list for a long time may be.[217] Validity of waiting lists is a problem: Independent chart reviews and clinical assessments, as well as patient surveys, have revealed proportions of patients inappropriately placed on lists ranging from 15 to 70%, clustering in the 20-40% range.[219-223] Further it is not always clear how waiting time is defined: from the moment the patient is referred to an orthopaedic surgeon or from the moment the patient is put on a waiting list for the operation, because it may be that reduction in waiting time on true inpatient waiting lists is gained at the expense of longer periods to be placed on the lists.[224] Also the fact whether waiting time is measured prospectively, retrospectively or in a cross-sectional way is a source of variability in perceptions about waiting times.[219] Initiatives to publish waiting times for surgery on the Internet have been found inaccurate and misleading by some physicians, but useful by government organizations because the figures might help patients choose their surgeon.[225,226]

Several studies did not find a relation between post-operative outcomes of THA and length in waiting times.[210,227] However, a Canadian study reported clinically important losses in quality of life and mobility in patients waiting longer than 6 months.[215] This finding, and the fact that osteoarthritis of the hip is common, imply a large burden from prolonged waits for surgery and support a case for limiting waits to less than 6 months. In a large multicenter study, Hajat et al. found a significant association between waiting for THA surgery and worse health status pre-operatively and 12 months post-operatively, even after adjustment for confounding variables.[40] Further, there have been pleas for prioritized waiting lists based on symptom severity and burden of disease.[228-231]

Aging populations: demographic changes and future requirements for THA

The world is witnessing an unprecedented, irremediable, and long-lasting aging process.[232] Western Europe will see the numbers of its inhabitants aged below 50 decrease, while those above 60 will sharply increase. The number of Europeans older than 60 is projected to rise from 84 million (21.8% of Europeans) at present to 107 million in 2050 (32.8%).[17] In The Netherlands, the number of people aged above 60 was 2,9 million (18.1%) in 2000 and will be over 5 million (28.1%) in 2030 (data from Dutch Census Bureau, 2002). Figure 3 shows the 'baby boomers' wave that is going to continuously extend the numbers of people over 60 years of age of in The Netherlands. Apart from changes in medical technology and individual factors, future healthcare management and costs will result from an increase in the prevalence of age related chronic and disabling conditions such as osteoarthritis, which belongs to

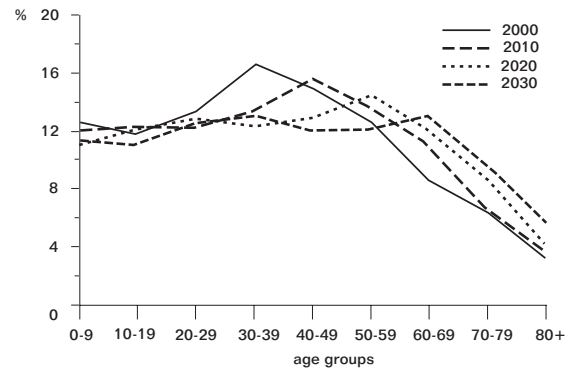


Figure 3

Population distribution by 10-year age group in The Netherlands from 2000 to 2030 (Central Bureau of Statistics, The Netherlands).

the most prevalent age related chronic conditions and which has been an important cause of disability through wide population health survey.[233] The large projected growth of the population above 60 suggests a sharp increase in the prevalence of arthritis for at least the four forthcoming decades.[17]
Very few studies have investigated the future demand for total hip replacement. These include scenarios for the United Kingdom, Sweden and Denmark.[234-236] In all these studies, age-specific incidence rates for THA rates are extrapolated to future population forecasts for different age groups. However, these are conservative estimations, because they do not take into account the yearly increase in incidence rates of the past decades as described in most studies. If outcome of THA continues to improve, indications for the procedure might be extended with consequences for the future incidence rate.[194] Further, as patients become better informed and more aware of present treatment possibilities, they may be less inclined to accept their disability and prefer treatment in an earlier stage of their disease.[101] Over a 20-year period, a two-to-threefold increase in the number of hip fractures is to be expected.[237] Studies of femoral neck fractures have shown that THA is a safe procedure if osteosynthesis fails, and that THA can be considered as a primary procedure and even provides a better outcome than internal fixation for elderly, relatively healthy patients with a displaced fracture of the femoral neck.[238-240] Together with the number of primary THA, the number of revision THA (which are more expensive per operation) will also increase. In conclusion, the predicted increase in number of THA because of demographic changes and extending indications will have large consequences for future healthcare budgets.

Generally, little is known about the epidemiology of THA in The Netherlands. This is due to limited registration systems with low compliance and stringent privacy regulations. Until 2002, there was no information in the literature on the effect of population aging on the future demand for THA in The Netherlands and Sweden, except for some predictions for the year 2000 in Sweden.[235] In Sweden, much more data is available on the epidemiology of THA due to the existence of the Swedish National Hip Registry. We had the opportunity to investigate the population from the Registry who underwent a revision for deep infection. Because of the large number of patients in the registry and the excellent follow-up, we could investigate some unique material such as risk factors for re-revision and trends in bacteriological patterns over the years. Studies on infected THA are usually too small to study these phenomena.
In spite of the existence of recommendations for the indication for THA, it was never investigated whether orthopaedic surgeons in Sweden or The Netherlands complied with these recommendations or whether there were any systematic differences in indication between the two countries. As mentioned earlier, recommendations or guidelines are important features in improving patient outcome. In spite of evidence from abroad about variability in perceptions about waiting lists and times,[219,241] little research is performed in The Netherlands regarding this topic.[226,242] Further, few

prospective studies are performed on the effect of waiting times on patient outcome.[210,215] Most studies are retrospective or cross-sectional in design.[135,218,243] Because of the large number of different questionnaires used in our patient group, we were for the first time able to study the relative performance of the Oxford Hip Score, WOMAC, SF-12, SF-36 and EQ-5D with respect to sensitivity to change and baseline characteristics.

Questions to be addressed in this thesis

Until now, limited data were available on the differences in incidence rates for THA in Sweden and The Netherlands, countries with a similar population and health care system. Also the effect of aging populations on the future demand for THA is unknown, which is important in planning future healthcare budgets for the operation. This prompted us to pose the following question in this thesis:

- What is the demographic profile of patients receiving THA in The Netherlands and Sweden and what is the expected demand for THA caused by the aging population in both countries?

For more than a decade, indication guidelines for primary THA exist both in The Netherlands and Sweden. However, it is not clear whether and to what extent these guidelines are applied in clinical practice, especially in this era with problems of limited health care budgets and, on the other hand, a more informed patient population wanting treatment for their complaints. Further, differences in incidence rates of THA in both countries might be explained by differences in indications for the operation. Therefore, the next question in this thesis was:

- To what extent are indication guidelines for THA applied in clinical practice, and what are the differences in indications for THA between The Netherlands and Sweden?

As discussed earlier, large databases can supply a lot of information with respect to treatment modalities in THA. Especially in infection of THA, registry information is important because infection is a relatively rare complication and reported sample sizes in studies on infected THA are usually small.[55] Often, reports are based on the results of one tertiary clinic, which limits the validity of the conclusions. Therefore, we investigated a large number of revisions for deep infection from the Swedish National Hip Registry and addressed the following question in this thesis:

- What are the demographic and bacteriological profiles and surgical practice in revision THA surgery for deep infection, and what is the risk for repeated revision for both aseptic loosening and recurrent infection?

Because of recurrent problems with long waiting lists for THA in The Netherlands and limited data on the possibly harmful effects of long waits for the procedure, the next question was:

- Does waiting for THA affect pre-operative health status and post-operative outcome of the operation?

Contents of this thesis

We investigated several aspects of the outcome and epidemiology of THA in The Netherlands and Sweden.
To quantify variations in practice and the effect of the aging process in Western countries on the incidence of THA, we calculated age-specific incidence rates of THA, studied the demographic profiles of the populations receiving THA and predicted demands for THA in The Netherlands and Sweden. This is described in **Chapter 2**. Because variations in incidence rates of THA might be explained by differences in indications for THA, we performed a survey among orthopaedic surgeons in Sweden and The Netherlands regarding their indications and modifying factors for primary THA (**Chapter 3**). In **Chapter 4**, we report on patient demographics, bacteriology and surgical practice in a national, Swedish cohort of 960 first revisions for deep infection, and investigated the risk and risk factors for repeated revision for both aseptic loosening and recurrent infection.
The current discrepancies in demand and supply of THA result in waiting lists for THA. This prompted us in **Chapter 5** to study the validity of a waiting list for THA in our university hospital, assess discrepancies with predicted waiting times and investigate the effects of extra resources on the length of the waiting list and the average waiting time for THA.
Chapter 6 reports on the effect of waiting times for THA in terms of loss of quality adjusted life years and additional burden perceived, and on the effect of waiting times and pre-operative outcome scores on postoperative outcome scores.
In **Chapter 7**, the minimum set of patient reported outcome measures required to assess health status after total hip replacement is determined.
Finally, in **Chapter 8**, we discuss results and conclusions from previous chapters and put forward some considerations on future perspectives in the field of outcome assessment of THA.

References.

1. Charnley J. Arthroplasty of the hip. A new operation. Lancet 1961:1129-1132.
2. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. J Bone Joint Surg Am 2004; 86(5):963-974.
3. Garellick G. On Outcome Assessment of Total Hip Replacement. Dept. Orthopaedics, Institute of Surgical Sciences. Göteborg: Göteborg University, 1998.
4. Chang RW, Pellisier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. Jama 1996; 275(11):858-65.
5. Maynard A. Developing the health care market. Economic Journal 1991; 101:1277-1286.
6. Williams A. Seeking priorities in health care: an economist's view. Journal of Bone and Joint Surgery Br 1991; 73(3):365-367.
7. National Medical Registration. Prismant, Institute for Healthcare Management 2000.
8. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998; 41(8):1343-55.
9. Malchau H, Herberts P, Garellick G. Annual Report of the Swedish Total Hip Arthroplasty Registry. Dept. Orthopaedics, Joint Replacement Unit, University of Gothenburg 2003.
10. van Saase JL, van Romunde LK, Cats A, Vandenbroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis 1989; 48(4):271-80.
11. Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, Kington RS, Lane NE, Nevitt MC, Zhang Y, Sowers M, McAlindon T, Spector TD, Poole AR, Yanovski SZ, Ateshian G, Sharma L, Buckwalter JA, Brandt KD, Fries JF. Osteoarthritis: New Insights. Part 1: The Disease and Its Risk Factors. Ann Intern Med 2000; 133(8):635-646.
12. Dexter P, Brandt K. Distribution and predictors of depressive symptoms in osteoarthritis. J Rheumatol 1994; 21(2):279-86.
13. Dunlop DD, Manheim LM, Yelin EH, Song J, Chang RW. The costs of arthritis. Arthritis Rheum 2003; 49(1):101-13.
14. Badley EM, Crotty M. An international comparison of the estimated effect of the aging of the population on the major cause of disablement, musculoskeletal disorders. J Rheumatol 1995; 22(10):1934-40.
15. Merx H, Dreinhofer K, Schrader P, Sturmer T, Puhl W, Gunther KP, Brenner H. International variation in hip replacement rates. Ann Rheum Dis 2003; 62(3):222-6.
16. R.I.V.M. Volksgezondheid toekomst verkenning. De gezondheidstoestand van de Nederlandse bevolking in de periode 1950-2010. Den Haag: SDU uitgeverij, 1993.
17. Ethgen O, Reginster JY. Degenerative musculoskeletal disease. Ann Rheum Dis 2004; 63(1):1-3.

18. Gross CP, Anderson GF, Powe NR. The relation between funding by the National Institutes of Health and the burden of disease. N Engl J Med 1999; 340(24):1881-7.
19. Lidgren L. The Bone and Joint Decade 2000-2010: an update. Acta Orthop Scand 2000; 71(1):3-6.
20. Glück T. Die Invaginationsmethode der Osteo- und Arthroplastik. Berl Klin Wochenschr Circulation 1890; 33:752.
21. Glück T. Referat über die Durch das Moderne Chirurgische Experiment Gewonnenen Positiven Resultaten, Betreffende die Naht und den Ersatz von Defekten Höherer Gewebe, Sowie über die Verwerthung Resorbirbarer und Lebendiger Tampons in der Chirurgie. Arch Klin Chir 1891; 41:186.
22. Smith-Petersen MN. Evolution of mould arthroplasty of hip joint. J Bone Joint Surg Br 1948; 30:59-75.
23. Aufranc OE. Constructive hip surgery with the Vitallium mold. A report of 1000 cases of arthroplasty of the hip over a fifteen-year period. J Bone Joint Surg Am 1957; 39:237-248.
24. Judet J, Judet R. The use of an artificial femoral head femoral head for arthroplasty of the hip joint. J Bone Joint Surg Br 1950; 32:166-173.
25. Moore A. The self-locking metal hip prosthesis. J Bone Joint Surg Am 1957; 39:811.
26. McKee GK, Watson-Farrar J. Replacement of arthritic hips by the McKee-Farrar prosthesis. J Bone Joint Surg Br 1966; 48:245-259.
27. Schmalzried TP, Szuszczewicz ES, Akizuki KH, Petersen TD, Amstutz HC. Factors correlating with long term survival of McKee-Farrar total hip prostheses. Clin Orthop 1996(329 Suppl):S48-59.
28. Heisel C, Silva M, Schmalzried TP. Bearing surface options for total hip replacement in young patients. Instr Course Lect 2004; 53:49-65.
29. Charnley J. Low Friction Arthroplasty of the Hip. Berlin Heidelberg New York: Springer-Verlag, 1979.
30. Charnley J. Anchorage of the femoral head prosthesis to the shaft of the femur. J Bone Joint Surg Br 1960; 42:28-30.
31. Peltier LF. A history of hip surgery. In Callaghan J, Rosenberg A, Rubash HE, eds. The Adult Hip, Vol. I. Philadelphia: Lippincott-Raven Publishers, 1998.
32. Eisler T. On loosening and revision in total hip arthroplasty. Dept. of Orthopaedics, Stockholm Söder Hospital, Karolinska Institutet. Stockholm, 2003.
33. Eftekhar NS. Long-term results of cemented total hip arthroplasty. Clin Orthop 1987(225):207-17.
34. Older J. Low-friction arthroplasty of the hip. A 10-12-year follow-up study. Clinical orthopaedics and related research 1986; 211:36-42.
35. Fender D, Harper WM, Gregg PJ. Outcome of Charnley total hip replacement across a single health region in England: the results at five years from a regional hip register. J Bone Joint Surg Br 1999; 81(4):577-81.

36. Fender D, van der Meulen JH, Gregg PJ. Relationship between outcome and annual surgical experience for the charnley total hip replacement. Results from a regional hip register. J Bone Joint Surg Br 2003; 85(2):187-90.
37. Garellick G, Malchau H, Herberts P. The Charnley versus the Spectron hip prosthesis: clinical evaluation of a randomized, prospective study of 2 different hip implants. J Arthroplasty 1999; 14(4):407-13.
38. Herziening consensus totale heupprothese. Quality Institute for Health Care CBO 1994.
39. Fortin PR, Penrod JR, Clarke AE, St-Pierre Y, Joseph L, Belisle P, Liang MH, Ferland D, Phillips CB, Mahomed N, Tanzer M, Sledge C, Fossel AH, Katz JN. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002; 46(12):3327-30.
40. Hajat S, Fitzpatrick R, Morris R, Reeves B, Rigge M, Williams O, Murray D, Gregg P. Does waiting for total hip replacement matter? Prospective cohort study. J Health Serv Res Policy 2002; 7(1):19-25.
41. Garellick G, Malchau H, Herberts P, Hansson E, Axelsson H, Hansson T. Life expectancy and cost utility after total hip replacement. Clin Orthop 1998(346):141-51.
42. Salvati E, Della Valle AG, Masri BA, Duncan CP. The infected total hip arthroplasty. Instr Course Lect 2003; 52:223-245.
43. Lavernia CJ, Drakeford MK, Tsao AK, Gittelsohn A, Krackow KA, Hungerford DS. Revision and primary hip and knee arthroplasty. A cost analysis. Clinical orthopaedics and related research 1995; 311:136-141.
44. Robinson AH, Palmer CR, Villar RN. Is revision as good as primary hip replacement? A comparison of quality of life. J Bone Joint Surg Br 1999; 81(1):42-5.
45. Malchau H, Herberts P, Eisler T, Garellick G, Soderman P. The Swedish Total Hip Replacement Register. J Bone Joint Surg Am 2002; 84-A Suppl 2:2-20.
46. Malchau H, Herberts P, Garellick G. Annual Report of the Swedish Total Hip Arthroplasty Registry. Dept. Orthopaedics, Joint Replacement Unit, University of Gothenburg 2004.
47. Willert HG, Ludwig J, Semlitsch M. Reaction of bone to methacrylate after hip arthroplasty: a long-term gross, light microscopic, and scanning electron microscopic study. J Bone Joint Surg Am 1974; 56(7):1368-82.
48. Horowitz SM, Doty SB, Lane JM, Burstein AH. Studies of the mechanism by which the mechanical failure of polymethylmethacrylate leads to bone resorption. J Bone Joint Surg Am 1993; 75(6):802-13.
49. Kim KJ, Chiba J, Rubash HE. In vivo and in vitro analysis of membranes from hip prostheses inserted without cement. J Bone Joint Surg Am 1994; 76(2):172-80.
50. Espehaug B, Engesaeter LB, Vollset SE, Havelin LI, Langeland N. Antibiotic prophylaxis in total hip arthroplasty. Review of 10,905 primary cemented total hip replacements reported to the Norwegian arthroplasty register, 1987 to 1995. J Bone Joint Surg Br 1997; 79(4):590-5.

51. Persson U, Persson M, Malchau H. The economics of preventing revisions in total hip replacement. Acta Orthop Scand 1999; 70(2):163-9.
52. Tunney MM, Patrick S, Gorman SP, Nixon JR, Anderson N, Davis RI, Hanna D, Ramage G. Improved detection of infection in hip replacements. A currently underestimated problem. J Bone Joint Surg Br 1998; 80(4):568-72.
53. Cierny G, 3rd, DiPasquale D. Periprosthetic total joint infections: staging, treatment, and outcomes. Clin Orthop 2002; 403:23-8.
54. Hanssen AD, Rand JA. Evaluation and treatment of infection at the site of a total hip or knee arthroplasty. Instructional Course Lectures, American Academy of Orthopaedic Surgeons 1999; 48.
55. Garvin KL, Hanssen AD. Infection after total hip arthroplasty. Past, present and future. J Bone Joint Surg Am 1995; 77 (10):1576-1588.
56. Tsukayama DT, Estrada R, Gustilo RB. Infection after total hip arthroplasty. A study of the treatment of one hundred and six infections. J Bone Joint Surg Am 1996; 78(4):512-23.
57. Gristina AG, Costerton JW. Bacterial adherence to biomaterials and tissue. The significance of its role in clinical sepsis. J Bone Joint Surg Am. 1985; 67(2):264-73.
58. Ritter MA. Operating room environment. Clin Orthop 1999(369):103-9.
59. Drangsholt MT. Current concepts review. Prophylactic use of antibiotics for procedures after total joint replacement. J Bone Joint Surg Am 1998; 80(9):1394-5.
60. Hill C, Flamant R, Mazas F, Evrard J. Prophylactic cefazolin versus placebo in total hip replacement. Report of a multicentre double-blind randomised trial. Lancet 1981; 1(8224):795-796.
61. Lodenkamper H, Lodenkamper U, Trompa K. Elimination of antibiotics from Palacos bone cement (personal experience from a bacteriological viewpoint after 10-year application in joint replacement surgery). Z Orthop 1982; 120(6):801-5.
62. van de Belt H, Neut D, Schenk W, van Horn JR, van der Mei HC, Busscher HJ. Infection of orthopedic implants and the use of antibiotic-loaded bone cements. A review. Acta Orthop Scand 2001; 72(6):557-71.
63. Scott CP, Higham PA, Dumbleton JH. Effectiveness of bone cement containing tobramycin. J Bone Joint Surg Br. 1999; 81(3):440-443.
64. Nijhof MW. Musculoskeletal infections. Developments in diagnosis and treatment. Dept. Orthopaedics. Utrecht: Utrecht University, 2001. pp. 1-221.
65. Garvin KL, Hinrichs SH, Urban JA. Emerging antibiotic-resistant bacteria. Their treatment in total joint arthroplasty. Clin Orthop 1999(369):110-23.
66. Rupp ME, Archer GL. Coagulase-negative staphylococci: pathogens associated with medical progress. Clin Infect Dis 1994; 19(2):231-43; quiz 244-5.
67. Ramage G, Tunney MM, Patrick S, Gorman SP, Nixon JR. Formation of Propionibacterium acnes biofilms on orthopaedic biomaterials and their susceptibility to antimicrobials. Biomaterials 2003; 24(19):3221-7.

68. Hope PG, Kristinsson KG, Norman P, Elson RA. Deep infection of cemented total hip arthroplasties caused by coagulase- negative staphylococci. J Bone Joint Surg Br 1989; 71(5):851-5.
69. Sanzen L, Walder M. Antibiotic resistance of coagulase-negative staphylococci in an orthopaedic department. J Hosp Infect 1988; 12(2):103-8.
70. Thornes B, Murray P, Bouchier-Hayes D. Development of resistant strains of Staphylococcus epidermidis on gentamicin-loaded bone cement in vivo. J Bone Joint Surg Br 2002; 84(5):758-60.
71. Lidgren L. Joint prosthetic infections: a success story. Acta Orthop Scand 2001; 72(6):553-6.
72. Huebner J, Goldmann DA. Coagulase-negative staphylococci: role as pathogens. Annu Rev Med 1999; 50:223-36.
73. Laverdiere M, Weiss K, Rivest R, Delorme J. Trends in antibiotic resistance of staphylococci over an eight-year period: differences in the emergence of resistance between coagulase positive and coagulase-negative staphylococci. Microb Drug Resist 1998; 4(2):119-22.
74. Archer GL. Alteration of cutaneous staphylococcal flora as a consequence of antimicrobial prophylaxis. Rev Infect Dis 1991; 13 Suppl 10:S805-9.
75. Bronzwaer SL, Cars O, Buchholz U, Molstad S, Goettsch W, Veldhuijzen IK, Kool JL, Sprenger MJ, Degener JE. A European study on the relationship between antimicrobial use and antimicrobial resistance. Emerg Infect Dis 2002; 8(3):278-82.
76. Wilcox MH. Antibiotic use and abuse. Lancet 1998; 352:1152.
77. Thore M, Kuhn I, Lofdahl S, Burman LG. Drug-resistant coagulase-negative skin staphylococci. Evaluation of four marker systems and epidemiology in an orthopaedic ward. Epidemiol Infect 1990; 105(1):95-105.
78. Swedres 2002. A report on Swedish Antibiotic Utilisation and Resistance in Human Medicine. The Swedish Strategic Programme for the Rational Use of Antimicrobial Agents (STRAMA) and the Swedish Institute for Infectious Disease Control 2003.
79. Bronzwaer S, Lonnroth A, Haigh R. The European Community strategy against antimicrobial resistance. Euro Surveill 2004; 9(1):1-3.
80. Spangehl MJ, Masri BA, O'Connell JX, Duncan CP. Prospective analysis of preoperative and intraoperative investigations for the diagnosis of infection at the sites of two hundred and two revision total hip arthroplasties. J Bone Joint Surg Am 1999; 81(5):672-83.
81. Hakki S, Harwood SJ, Morrissey MA, Camblin JG, Laven DL, Webster WB, Jr. Comparative study of monoclonal antibody scan in diagnosing orthopaedic infection. Clin Orthop 1997; 335(335):275-85.
82. Della Valle CJ, Bogner E, Desai P, Lonner JH, Adler E, Zuckerman JD, Di Cesare PE. Analysis of frozen sections of intraoperative specimens obtained at the time of reoperation after hip or knee resection arthroplasty for the treatment of infection. J Bone Joint Surg Am 1999; 81(5):684-9.

83. Powles JW, Spencer RF, Lovering AM. Gentamicin release from old cement during revision hip arthroplasty. J Bone Joint Surg Br 1998; 80(4):607-10.
84. von Eiff C, Bettin D, Proctor RA, Rolauffs B, Lindner N, Winkelmann W, Peters G. Recovery of small colony variants of Staphylococcus aureus following gentamicin bead placement for osteomyelitis. Clin Infect Dis 1997; 25(5):1250-1251.
85. Tunney MM, Patrick S, Curran MD, Ramage G, Hanna D, Nixon JR, Gorman SP, Davis RI, Anderson N. Detection of prosthetic hip infection at revision arthroplasty by immunofluorescence microscopy and PCR amplification of the bacterial 16S rRNA gene. J Clin Microbiol 1999; 37(10):3281-90.
86. Lieberman JR, Callaway GH, Salvati EA, Pellicci PM, Brause BD. Treatment of the infected total hip arthroplasty with a two-stage reimplantation protocol. Clin Orthop 1994(301):205-12.
87. McDonald DJ, Fitzgerald RH, Jr., Ilstrup DM. Two-stage reconstruction of a total hip arthroplasty because of infection. J Bone Joint Surg Am 1989; 71(6):828-34.
88. Magnan B, Regis D, Biscaglia R, Bartolozzi P. Preformed acrylic bone cement spacer loaded with antibiotics: use of two-stage procedure in 10 patients because of infected hips after total replacement. Acta Orthop Scand 2001; 72(6):591-4.
89. Yamamoto K, Miyagawa N, Masaoka T, Katori Y, Shishido T, Imakiire A. Clinical effectiveness of antibiotic-impregnated cement spacers for the treatment of infected implants of the hip joint. J Orthop Sci 2003; 8(6):823-8.
90. Younger AS, Duncan CP, Masri BA, McGraw RW. The outcome of two-stage arthroplasty using a custom-made interval spacer to treat the infected hip. J Arthroplasty 1997; 12(6):615-23.
91. Crockarell JR, Hanssen AD, Osmon DR, Morrey BF. Treatment of infection with debridement and retention of the components following hip arthroplasty. J Bone Joint Surg Am 1998; 80(9):1306-13.
92. Hanssen AD, Osmon DR. Assessment of patient selection criteria for treatment of the infected hip arthroplasty. Clin Orthop 2000(381):91-100.
93. Buchholz HW, Elson RA, Engelbrecht E, Lodenkamper H, Rottger J, Siegel A. Management of deep infection of total hip replacement. J Bone Joint Surg Br 1981; 63-B(3):342-53.
94. Jackson WO, Schmalzried TP. Limited role of direct exchange arthroplasty in the treatment of infected total hip replacements. Clin Orthop 2000(381):101-5.
95. Pagnano MW, Trousdale RT, Hanssen AD. Outcome after reinfection following reimplantation hip arthroplasty. Clin Orthop 1997(338):192-204.
96. McPherson EJ, Woodson C, Holtom P, Roidis N, Shufelt C, Patzakis M. Periprosthetic total hip infection: outcomes using a staging system. Clin Orthop 2002; 403:8-15.
97. Hanssen AD, Osmon DR. Evaluation of a staging system for infected hip arthroplasty. Clin Orthop 2002; 403:16-22.

98. Malchau H. On the importance of stepwise introduction of new hip implant technology. Assessment of total hip replacement using clinical evaluation, radiostereometry, digitised radiography and a national hip registry. Goteborg University, Sweden, 1995.
99. National Audit Office (NAO). Hip replacements: getting it right first time. 2000.
100. National Audit Office (NAO). Hip replacements: an update. 2003.
101. Bulstrode CJ, Murray DW, Carr AJ, Pynsent PB, Carter SR. Designer hips. Bmj 1993; 306(6880):732-3.
102. Gillespie WJ, Pekarsky B, O'Connell DL. Evaluation of new technologies for total hip replacement. Economic modelling and clinical trials. Journal of Bone and Joint Surgery Br 1995; 77(4):528-533.
103. Healy WL, Kirven FM, Iorio R, Patch DA, Pfeifer BA. Implant standardization for total hip arthroplasty. An implant selection and a cost reduction program. Journal of Arthroplasty 1995; 10(2):177-183.
104. Codman EA. The shoulder. Boston, 1934.
105. Johanson NA. Outcomes assessment. In Callaghan J, Rosenberg A, Rubash HE, eds. The Adult Hip, Vol. II. Philadelphia: Lippincott-Raven Publishers, 1998.
106. Bourne RB. Con: some clinical research is not cost-effective--weaknesses of outcome studies. Spine 1995; 20(3):386-7.
107. Bourne RB, Maloney WJ, Wright JG. An AOA critical issue. The outcome of the outcomes movement. J Bone Joint Surg Am 2004; 86 (3):633-40.
108. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. Bmj 1996; 312(7023):71-2.
109. Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States medicare population. J Bone Joint Surg Am 2001; 83 (11):1622-9.
110. Herndon JH, Davidson SM, Apazidis A. Recent socioeconomic trends in orthopaedic practice. J Bone Joint Surg Am 2001; 83 (7):1097-105.
111. Hudak PL, Clark JP, Hawker GA, Coyte PC, Mahomed NN, Kreder HJ, Wright JG. "You're perfect for the procedure! Why don't you want it?" Elderly arthritis patients' unwillingness to consider total joint arthroplasty surgery: a qualitative study. Med Decis Making 2002; 22(3):272-8.
112. Wright JG, Hawker GA, Bombardier C, Croxford R, Dittus RS, Freund DA, Coyte PC. Physician enthusiasm as an explanation for area variation in the utilization of knee replacement surgery. Med Care 1999; 37(9):946-56.
113. The Center of the Evaluative Clinical Sciences DMS. The Dartmouth atlas of musculoskeletal health care. Chicago: AHA Press, 2000.
114. Ellwood PM. Shattuck lecture--outcomes management. A technology of patient experience. N Engl J Med 1988; 318(23):1549-56.
115. Dawson J, Carr A. Outcomes evaluation in orthopaedics. J Bone Joint Surg Br 2001; 83(3):313-5.

116. Birkmeyer JD. Outcomes research and surgeons. Surgery 1998; 124(3):477-83.
117. Garellick G, Malchau H, Herberts P. Survival of hip replacements. A comparison of a randomized trial and a registry. Clin Orthop 2000(375):157-67.
118. Kreder HJ, Deyo RA, Koepsell T, Swiontkowski MF, Kreuter W. Relationship between the volume of total hip replacements performed by providers and the rates of postoperative complications in the state of Washington. J Bone Joint Surg Am 1997; 79(4):485-94.
119. Howe JG, Lambert B. Critical Pathways in Total Hip Arthroplasty. In Callaghan J, Rosenberg A, Rubash HE, eds. The Adult Hip, Vol. II. Philadelphia: Lippincott-Raven Publishers, 1998.
120. Ritter MA, Albohm MJ. Overview: maintaining outcomes for total hip arthroplasty. The past, present, and future. Clin Orthop 1997(344):81-7.
121. Rosendal H. Comparative Cohort Studies in Transmural Care: Three cases of structural embedded practice in The Netherlands. Julius Center for Health Sciences and Primary Care. Utrecht: Utrecht University, 2002.
122. Freedman KB, Back S, Bernstein J. Sample size and statistical power of randomised, controlled trials in orthopaedics. J Bone Joint Surg Br 2001; 83(3):397-402.
123. Bhandari M, Morrow F, Kulkarni AV, Tornetta P, 3rd. Meta-analyses in orthopaedic surgery. A systematic review of their methodologies. J Bone Joint Surg Am 2001; 83 (1):15-24.
124. Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, Dawson J, Carr A, Britton A, Briggs A. Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses. Health Technol Assess 1998; 2(20):1-64.
125. Daellenbach HG, Gillespie WJ, Crosbie P, Daellenbach US. Economic appraisal of new technology in the absence of survival data-- the case of total hip replacement. Soc Sci Med 1990; 31(12):1287-93.
126. Faulkner A, Kennedy LG, 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. Health Technol Assess 1998; 2(6):1-133.
127. Pynsent PB, Carter SR, Bulstrode CJ. The total cost of hip-joint replacement; a model for purchasers. J Public Health Med 1996; 18(2):157-68.
128. Gold MR, Russell RB, Siegel JE, Weinsyein MC. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press, 1996.
129. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. Jama 1996; 276(14):1172-7.
130. Rorabeck CH, Bourne RB, Laupacis A, Feeny D, Wong C, Tugwell P, Leslie K, Bullas R. A double-blind study of 250 cases comparing cemented with cementless total hip arthroplasty. Cost-effectiveness and its impact on health-related quality of life. Clin Orthop 1994(298):156-164.

131. Ahnfelt L, Andersson G, Herberts P. Reoperation av totala höftledsplastiker. Läkartidningen 1980; 77(30-31):2604-7.
132. Herberts P, Ahnfelt L, Malchau H, Stromberg C, Andersson GB. Multicenter clinical trials and their value in assessing total joint arthroplasty. Clin Orthop 1989(249):48-55.
133. Söderman P, Malchau H, Herberts P, Johnell O. Are the findings in the Swedish National Total Hip Arthroplasty Register valid? J Arthroplasty 2000; 15(7):884-9.
134. Söderman P. On the validity of the results from the Swedish National Total Hip Arthroplasty Register. Dept. Of Orthopaedics, Institute of Surgical Sciences. Göteborg: Göteborg University, 2000.
135. Williams JI, Llewellyn Thomas H, Arshinoff R, Young N, Naylor CD. The burden of waiting for hip and knee replacements in Ontario. Ontario Hip and Knee Replacement Project Team. J Eval Clin Pract 1997; 3(1):59-68.
136. Linder L. Boneloc--The Christiansen experience revisited. Acta Orthop Scand 1995; 66(3):205-6.
137. Paavolainen P, Hamalainen M, Mustonen H, Slatis P. Registration of arthroplasties in Finland. A nationwide prospective project. Acta Orthopedica Scandinavica Suppl 1991; 241(1):27-30.
138. Havelin LI, Espehaug B, Vollset SE, Engesaeter LB, Langeland N. The Norwegian arthroplasty register. A survey of 17,444 hip replacements 1987-1990. Acta Orthopedica Scandinavica 1993; 64(3):245-251.
139. Lucht U. The Danish Hip Arthroplasty Register. Acta Orthop Scand 2000; 71(5):433-9.
140. Maloney WJ. National Joint Replacement Registries: has the time come? J Bone Joint Surg Am 2001; 83 (10):1582-5.
141. Swiontkowski MF, Maloney WJ. The Feasibility of a National Joint Replacement Registry in the United States. J Bone Joint Surg Am 2002; 84(5):871-872.
142. Walker PS, Mai SF, Cobb AG, Bentley G, Hua J. Prediction of clinical outcome of THR from migration measurements on standard radiographs. A study of cemented Charnley and Stanmore femoral stems. J Bone Joint Surg Br 1995; 77(5):705-14.
143. Dobbs HS. Survivorship of total hip replacements. J Bone Joint Surg Br 1980; 62 (2):168-73.
144. Murray DW, Carr AJ, Bulstrode C. Survival analysis of joint replacements. J Bone Joint Surg Br 1993; 75(5):697-704.
145. Britton AR, Murray DW, Bulstrode CJ, McPherson K, Denham RA. Pain levels after total hip replacement: their use as endpoints for survival analysis. J Bone Joint Surg Br 1997; 79(1):93-8.
146. Dorey F, Amstutz HC. Survivorship analysis in the evaluation of joint replacement. J Arthroplasty 1986; 1(1):63-9.
147. Murray DW, Britton AR, Bulstrode CJ. Loss to follow-up matters. J Bone Joint Surg Br 1997; 79(2):254-7.

148. Dorey FJ, Amstutz HC. The Need to Account for Patient Activity When Evaluating the Results of Total Hip Arthroplasty with Survivorship Analysis. J Bone Joint Surg Am 2002; 84(5):709-710.
149. Lieberman JR, Huo MH, Schneider R, Salvati EA, Rodi S. Evaluation of painful hip arthroplasties. Are technetium bone scans necessary? J Bone Joint Surg Br 1993; 75(3):475-8.
150. DeLee JG, Charnley J. Radiological demarcation of cemented sockets in total hip replacement. Clin Orthop 1976(121):20-32.
151. Gruen TA, McNeice GM, Amstutz HC. "Modes of failure" of cemented stem-type femoral components: a radiographic analysis of loosening. Clin Orthop 1979(141):17-27.
152. Ritter MA, Zhou H, Keating CM, Keating EM, Faris PM, Meding JB, Berend ME. Radiological factors influencing femoral and acetabular failure in cemented Charnley total hip arthroplasties. J Bone Joint Surg Br 1999; 81(6):982-6.
153. McCaskie AW, Brown AR, Thompson JR, Gregg PJ. Radiological evaluation of the interfaces after cemented total hip replacement. Interobserver and intraobserver agreement. J Bone Joint Surg Br 1996; 78(2):191-4.
154. Freeman MA, Plante-Bordeneuve P. Early migration and late aseptic failure of proximal femoral prostheses. J Bone Joint Surg Br 1994; 76(3):432-8.
155. Karrholm J, Herberts P, Hultmark P, Malchau H, Nivbrant B, Thanner J. Radiostereometry of hip prostheses. Review of methodology and clinical results. Clin Orthop 1997(344):94-110.
156. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. J Bone Joint Surg Am 1969; 51(4):737-55.
157. Merle d'Aubigné R, Postel M. Functional results of hip arthroplasty with acrylic prosthesis. J Bone Joint Surg Am 1954; 36:451-75.
158. Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. J Bone Joint Surg Br 1972; 54(1):61-76.
159. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation--value in comparison with other assessment tests. Scand J Rheumatol Suppl 1987; 65:85-9.
160. Drake BG, Callahan CM, Dittus RS, Wright JG. Global rating systems used in assessing knee arthroplasty outcomes. J Arthroplasty 1994; 9(4):409-17.
161. Brokelman RB, van Loon CJ, Rijnberg WJ. Patient versus surgeon satisfaction after total hip arthroplasty. J Bone Joint Surg Br 2003; 85(4):495-8.
162. Bullens PH, van Loon CJ, de Waal Malefijt MC, Laan RF, Veth RP. Patient satisfaction after total knee arthroplasty: a comparison between subjective and objective outcome assessments. J Arthroplasty 2001; 16(6):740-7.
163. McGee MA, Howie DW, Ryan P, Moss JR, Holubowycz OT. Comparison of patient and doctor responses to a total hip arthroplasty clinical evaluation questionnaire. J Bone Joint Surg Am 2002; 84 (10):1745-52.

164. Lieberman JR, Dorey F, Shekelle P, Schumacher L, Thomas BJ, Kilgus DJ, Finerman GA. Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. J Bone Joint Surg Am 1996; 78(6):835-8.
165. Wright JG, Rudicel S, Feinstein AR. Ask patients what they want. Evaluation of individual complains before total hip replacement. J Bone Joint Sur Br 1994; 76(2):229-234.
166. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. Journal of Orthopaedic Rheumatology 1988; 1:95-108.
167. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78(2):185-90.
168. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire--an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987; 14(3):446-51.
169. Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003; 4(1):10.
170. Ware JE, Jr., Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. Med Care 1995; 33(4 Suppl):AS264-79.
171. Hunt SM, McKenna SP, McEwen J, Backett EM, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. J Epidemiol Community Health 1980; 34(4):281-6.
172. de Bruin AF, de Witte LP, Stevens F, Diederiks JP. Sickness Impact Profile: the state of the art of a generic functional status measure. Soc Sci Med 1992; 35(8):1003-14.
173. Brooks R. EuroQol: the current state of play. Health Policy 1996; 37(1):53-72.
174. Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. Arch Phys Med Rehabil 1999; 80(5):572-9.
175. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002; 21(2):271-92.
176. Blanchard C, Feeny D, Mahon JL, Bourne R, Rorabeck C, Stitt L, Webster-Bogaert S. Is the Health Utilities Index responsive in total hip arthroplasty patients? J Clin Epidemiol 2003; 56(11):1046-54.
177. Neumann PJ, Zinner DE, Wright JC. Are methods for estimating QALYs in cost-effectiveness analyses improving? Med Decis Making 1997; 17(4):402-8.
178. Johannesson M. On aggregating QALYs: a comment on Dolan. J Health Econ 1999; 18(3):381-6.

179. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. Med Care 2003; 41(7):791-801.
180. Lingard E, Hashimoto H, Sledge C. Development of outcome research for total joint arthroplasty. J Orthop Sci 2000; 5(2):175-7.
181. Sun Y, Sturmer T, Gunther KP, Brenner H. Reliability and validity of clinical outcome measurements of osteoarthritis of the hip and knee--a review of the literature. Clin Rheumatol 1997; 16(2):185-98.
182. Greenfield S, Apolone G, McNeil BJ, Cleary PD. The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. Med Care 1993; 31(2):141-54.
183. Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology (Oxford) 1999; 38(4):355-61.
184. Dawson J, Fitzpatrick R, Murray D, Carr A. The problem of 'noise' in monitoring patient-based outcomes: generic, disease-specific and site-specific instruments for total hip replacement. J Health Serv Res Policy 1996; 1(4):224-31.
185. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997; 50(3):239-46.
186. Brinker MR, Lund PJ, Barrack RL. Demographic biases of scoring instruments for the results of total knee arthroplasty. J Bone Joint Surg Am 1997; 79(6):858-65.
187. Ritter MA, Thong AE, Davis KE, Berend ME, Meding JB, Faris PM. Long-term deterioration of joint evaluation scores. J Bone Joint Surg Br 2004; 86(3):438-42.
188. Melton LJ, Stauffer RN, Chao EY, Ilstrup DM. Rates of total hip arthroplasty; a population-based study. N Engl J Med 1982; 307(20):1242-5.
189. Peterson MG, Hollenberg JP, Szatrowski TP, Johanson NA, Mancuso CA, Charlson ME. Geographic variations in the rates of elective total hip and knee arthroplasties among Medicare beneficiaries in the United States. J Bone Joint Surg Am 1992; 74(10):1530-9.
190. Ingvarsson T, Hagglund G, Jonsson H, Jr., Lohmander LS. Incidence of total hip replacement for primary osteoarthrosis in Iceland 1982-1996. Acta Orthop Scand 1999; 70(3):229-33.
191. Herberts P, Malchau H. How outcome studies have changed total hip arthroplasty practices in Sweden. Clin Orthop 1997(344):44-60.
192. Dunsmuir RA, Allan DB, Davidson LA. Increased incidence of primary total hip replacement in rural communities. Bmj 1996; 313(7069):1370.
193. Wells VM, Hearn TC, McCaul KA, Anderton SM, Wigg AE, Graves SE. Changing incidence of primary total hip arthroplasty and total knee arthroplasty for primary osteoarthritis. J Arthroplasty 2002; 17(3):267-73.
194. Friedman B, Elixhauser A. Increased use of an expensive, elective procedure: total hip replacements in the 1980s. Med Care 1993; 31(7):581-99.

195. Mancuso CA, Ranawat CS, Esdaile JM, Johanson NA, Charlson ME. Indications for total hip and total knee arthroplasties. Results of orthopaedic surveys. J Arthroplasty 1996; 11(1):34-46.
196. Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R, Wilkins A, Badley EM. Determining the need for hip and knee arthroplasty: the role of clinical severity and patients' preferences. Med Care 2001; 39(3):206-16.
197. Oishi CS, Hoaglund FT, Gordon L, Ross PD. Total hip replacement rates are higher among Caucasians than Asians in Hawaii. Clin Orthop 1998(353):166-74.
198. Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R, Badley EM. Differences between men and women in the rate of use of hip and knee arthroplasty. N Engl J Med 2000; 342(14):1016-22.
199. Hoaglund FT, Oishi CS, Gialamas GG. Extreme variations in racial rates of total hip arthroplasty for primary coxarthrosis: a population-based study in San Francisco. Ann Rheum Dis 1995; 54(2):107-10.
200. Mahomed NN, Barrett JA, Katz JN, Phillips CB, Losina E, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Rates and outcomes of primary and revision total hip replacement in the United States medicare population. J Bone Joint Surg Am 2003; 85 (1):27-32.
201. Guideline development methods: Information for National Collaborating Centres and Guideline Developers. National Institute for Clinical Excellence 2004.
202. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L, Wensing M, Dijkstra R, Donaldson C. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6):iii-iv, 1-72.
203. National Institute of Health (NIH). Total Hip Replacement. NIH Consensus Statement 12-14 1994; 12(5):1-31.
204. Total Hip-Joint Replacement in Sweden. JAMA 1982; 248(15):1822-1824.
205. Wailoo A, Roberts J, Brazier J, McCabe C. Efficiency, equity, and NICE clinical guidelines. Bmj 2004; 328(7439):536-7.
206. Guideline development within the Quality Institute for Health Care CBO, manual for workgroup members. Quality Institute for Health Care CBO 2000.
207. Wright J, Williams R, Wilkinson JR. Development and importance of health needs assessment. Bmj 1998; 316(7140):1310-3.
208. Frankel S, Eachus J, Pearson N, Greenwood R, Chan P, Peters TJ, Donovan J, Smith GD, Dieppe P. Population requirement for primary hip-replacement surgery: a cross- sectional study. Lancet 1999; 353(9161):1304-9.
209. Sanders C, Donovan JL, Dieppe PA. Unmet need for joint replacement: a qualitative investigation of barriers to treatment among individuals with severe pain and disability of the hip and knee. Rheumatology (Oxford) 2004; 43(3):353-7.

210. Kelly KD, Voaklander DC, Johnston DW, Newman SC, Suarez-Almazor ME. Change in pain and function while waiting for major joint arthroplasty. J Arthroplasty 2001; 16(3):351-9.
211. Quan H, Lafreniere R, Johnson D. Health service costs for patients on the waiting list. Can J Surg 2002; 45(1):34-42.
212. Sheldon T. Dutch patients travel to Spain for orthopaedic surgery. Bmj 2001; 322(7302):1565C.
213. Derrett S, Paul C, Morris JM. Waiting for elective surgery: effects on health-related quality of life. Int J Qual Health Care 1999; 11(1):47-57.
214. Saleh KJ, Wood KC, Gafni A, Gross AE. Immediate surgery versus waiting list policy in revision total hip arthroplasty. An economic evaluation. J Arthroplasty 1997; 12(1):1-10.
215. Mahon JL, Bourne RB, Rorabeck CH, Feeny DH, Stitt L, Webster-Bogaert S. Health-related quality of life and mobility of patients awaiting elective total hip arthroplasty: a prospective study. Cmaj 2002; 167(10):1115-21.
216. Barkun JS. Patients waiting for elective general surgery: a culture of waiting? Can J Surg 2002; 45(1):6-7.
217. Lewis S, Barer ML, Sanmartin C, Sheps S, Shortt SE, McDonald PW. Ending waiting-list mismanagement: principles and practice. Cmaj 2000; 162(9):1297-300.
218. Ho E, Coyte PC, Bombardier C, Hawker G, Wright JG. Ontario patients' acceptance of waiting times for knee replacements. J Rheumatol 1994; 21(11):2101-5.
219. Sanmartin C, Shortt SE, Barer ML, Sheps S, Lewis S, McDonald PW. Waiting for medical services in Canada: lots of heat, but little light. Cmaj 2000; 162(9):1305-10.
220. Rigge M. Quality of life of long wait orthopaedic patients before and after admission: a consumer audit. Qual Health Care 1994; 3(3):159-63.
221. Hochuli VK. Orthopaedic waiting list reduction through a review of service provision: the problems encountered. J R Soc Med 1988; 81(8):445-7.
222. Donaldson LJ, Pugh EJ, Smyth JE. A census of patients waiting for treatment or consultation in surgical specialties in an English health region. Community Med 1989; 11(1):13-20.
223. Elwyn GJ, Williams LA, Barry S, Kinnersley P. Waiting list management in general practice: a review of orthopaedic patients. Bmj 1996; 312(7035):887-8.
224. Smith T. Waiting times: monitoring the total postreferral wait. Bmj 1994; 309(6954): 593-6.
225. Kent H. Waiting-list Web site "inaccurate" and "misleading," BC doctors complain. Cmaj 1999; 161(2):181-2.
226. Stoop AP, Berg M. De Juiste Tijd: Totstandkoming, Betekenis en Gebruik van wachttijdgegevens. ZonMW 2002.
227. Nilsdotter AK, Lohmander LS. Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis. Rheumatology (Oxford) 2002; 41(11):1261-7.
228. Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 1: Overview. Bmj 1997; 314(7074):131-4.

229. Conner-Spady BL, Arnett G, McGurran JJ, Noseworthy TW. Prioritization of patients on scheduled waiting lists: validation of a scoring system for hip and knee arthroplasty. Can J Surg 2004; 47(1):39-46.
230. Naylor CD, Williams JI. Primary hip and knee replacement surgery: Ontario criteria for case selection and surgical priority. Quality in Health Care 1996; 5:2 0-30.
231. Gauld R, Derrett S. Solving the surgical waiting list problem? New Zealand's 'booking system'. Int J Health Plann Manage 2000; 15(4):259-72.
232. Population Division of the Department of Economic and Social Affairs of the United Nations. World population prospects: the 2000 revision and world urbanization prospects: the 2001 revision. http://esa.un.org/unpp. 2001.
233. Alonso J, Ferrer M, Gandek B, Ware JE, Jr., Aaronson NK, Mosconi P, Rasmussen NK, Bullinger M, Fukuhara S, Kaasa S, Leplege A. Health-related quality of life associated with chronic conditions in eight countries: results from the International Quality of Life Assessment (IQOLA) Project. Qual Life Res 2004; 13(2):283-98.
234. Birrell F, Johnell O, Silman A. Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. Ann Rheum Dis 1999; 58(9):569-572.
235. Goldie I. Total hip and knee replacements. Acta Orthop Scand Suppl 1991; 241:23-6.
236. Overgaard S, Hansen LN, Knudsen HM, Mossing N. Epidemiological studies on total hip arthroplasty in the country of south Jutland, Denmark, in the 1980s. Acta Orthopaedica Scandinavica 1991; 62(suppl 246):42.
237. Johnell O. The socioeconomic burden of fractures: today and in the 21st century. Am J Med 1997; 103(2A):20S-25S; discussion 25S-26S.
238. Rogmark C, Carlsson A, Johnell O, Sernbo I. A prospective randomised trial of internal fixation versus arthroplasty for displaced fractures of the neck of the femur. Functional outcome for 450 patients at two years. J Bone Joint Surg Br 2002; 84(2):183-8.
239. Squires B, Bannister G. Displaced intracapsular neck of femur fractures in mobile independent patients: total hip replacement or hemiarthroplasty? Injury 1999; 30(5):345-8.
240. Tidermark J, Ponzer S, Svensson O, Soderqvist A, Tornkvist H. Internal fixation compared with total hip replacement for displaced femoral neck fractures in the elderly. A randomised, controlled trial. J Bone Joint Surg Br 2003; 85(3):380-8.
241. Newton JN, Henderson J, Goldacre MJ. Waiting list dynamics and the impact of earmarked funding. Bmj 1995; 311(7008):783-5.
242. Stoop AP, Brouwer WBF. De patiënt achter de cijfers. Medisch Contact 2001; 56(44):1610-1612.
243. Coyte PC, Wright JG, Hawker GA, Bombardier C, Dittus RS, Paul JE, Freund DA, Ho E. Waiting times for knee-replacement surgery in the United States and Ontario. N Engl J Med 1994; 331(16):1068-71.

Chapter 2

The epidemiology of total hip replacement in The Netherlands and Sweden:

Present status and future needs.

Ostendorf M
Johnell O
Malchau H
Dhert WJA
Schrijvers AJP
Verbout AJ

Acta Orthopaedica Scandinavica 2002;73(3):282-86.

Introduction

The improvements in surgical technique, implant material, and implant design have led to excellent long-term results of total hip arthroplasty (THA) with implant survival rates of 93% after 10 years.[1] In consequence, the demand for total hip replacement has increased. Another reason for the growing demand is the increasing number of elderly people. Because of capacity and budget restraints, many countries have waiting times for the procedure.[2] To plan for future demands and costs, it is important to estimate the number of total hip replacements to be expected. We calculated age-specific incidences, studied the demographic profile of the population undergoing THA and predicted the number of THAs to be expected in The Netherlands and Sweden.

Patients and methods

We obtained information on all patients admitted to Dutch hospitals for total hip arthroplasty in the period 1986-1997 from the National Medical Registration system as provided by Prismant, Institute for Healthcare Management. Swedish data were obtained from the Swedish Discharge registry for the period 1987-1997. In The Netherlands, we obtained data on THA for every other year in the period 1986-1996 and 1997. We did not use data from the Swedish Hip Arthroplasty Register because no such register exists in The Netherlands and the Discharge registers can be more easily compared. The Swedish Hip Arthroplasty Register was validated by comparing it with the Discharge register.[3] Information was obtained about the number of primary total hip arthroplasties. For every procedure, the gender and age group of the patient were registered. We also obtained information on the (projected) demographic profile of the Dutch and Swedish populations from the Census Bureaus in both countries (Centraal Bureau voor de Statistiek (CBS), The Netherlands, and Statistiska Centralbyrån (SCB), Sweden). We calculated annual and age-specific incidences in both countries and estimated the need for primary total hip arthroplasty by using the projected changes of the Dutch and Swedish populations until 2020.

Results

We found an increase of 68% in the number of primary total hip replacements in the Netherlands from 10,359 operations in 1986 to 17,401 in 1997. This is an increase from 71 operations to 112 operations per 100,000 inhabitants (Figure 1). Only 15% of the increase in number of hip replacements could be explained by changes in the size and age-profile of the population. On the basis of the incidence of THA in 1997, we predicted the annual number of primary total hip replacements. Assuming no further change in the age- and sex-specific arthroplasty rates, the annual number of total hip replacements in The Netherlands by the year 2020 will increase by 44% to 25,090 operations.

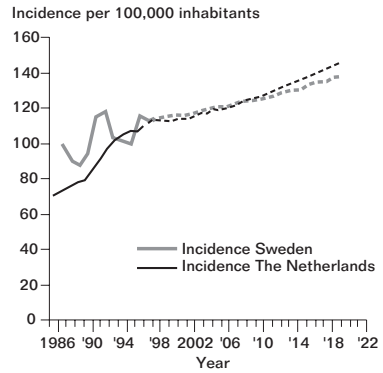


Figure 1

Incidence of total hip replacement per 10^5 inhabitants in the period 1986-2020 in The Netherlands and Sweden. (Predictions from 1997).

In Sweden, the number of arthroplasties increased by 20% from 8,336 in 1987 to 10,015 in 1997, an increase from 99 operations to 113 operations per 100,000 inhabitants (Figure 1). Only 3% of the increase in number of hip replacements could be explained by demographic changes in the population. The number of THA varied considerably each year. On the basis of the incidence of THA in 1997, the predicted annual number of THA in Sweden by the year 2020 will rise to 12,773 operations, an increase of 28% compared with 1997.

When comparing the age of patients undergoing THA in the period 1986-1997, it turned out that in The Netherlands the distribution of the age-groups had changed relatively little, but in Sweden, more elderly people were operated on in 1997 (39% over 75 years of age) compared to 1987 (28% over 75 years of age) (Figure 2).

Although the overall incidences of THA in 1997 were similar in both countries, it appears that the age-specific incidences for THA in women were higher in The Netherlands while the age-specific incidences for THA in men were higher in Sweden (Table). To correct for differences in the population profile (Figure 3), we calculated the number of total hip replacements in The Netherlands based on the Swedish incidence of THA in different age-classes.[4] (Figure 4). After correction for differences in population structure the rate of THA in The Netherlands is 20% higher than in Sweden.

Table 1

Age-standardized incidence of total hip replacement per 10^5 inhabitants in the Netherlands and Sweden in 1997.

Age	Both genders		Males		Females	
	Netherlands	Sweden	Netherlands	Sweden	Netherlands	Sweden
< 44	4	4	4	4	4	5
45-54	51	52	45	46	57	58
55-64	229	206	151	190	306	223
65-74	570	431	328	391	768	465
75-84	713	560	383	445	905	640
> 85	408	381	281	289	452	422
Total	112	113	61	88	161	138

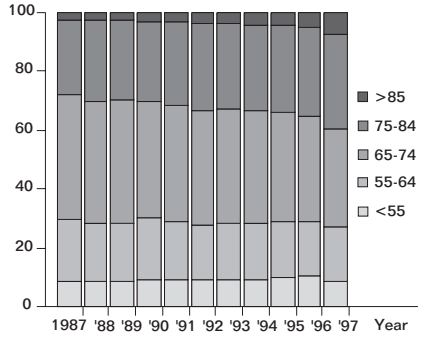


Figure 2a

Percentage of patients in Sweden receiving total hip replacement in different age-classes in different years.

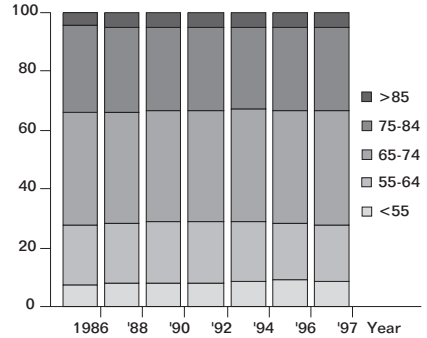


Figure 2b

Percentage of patients in The Netherlands receiving total hip replacement in different age-classes in different years.

Discussion

From 1986 the number of THAs has been increasing in Sweden and The Netherlands. Only a small percentage of the increase could be attributed to demographic changes. We found that the procedure has been offered to an increasing number of patients, who were previously regarded as too old or too sick.[5,6] The shift in age-classes of patients undergoing THA in Sweden over the period 1987-1997 may be due to Swedish orthopaedic surgeons becoming more reluctant to operate on younger patients because of the disappointing results of THA in younger patients reported in the Swedish National Hip Register.

We found a difference in the gender ratio of the population undergoing THA in both countries (Table). This could be explained by different rates of osteoarthritis (OA), the main indication for THA[1] in both genders in the Swedish and Dutch populations. Dutch studies have reported more radiological hip OA in women than in men, but in a Swedish study, the incidence was similar.[7-9] However, these studies used different criteria to define OA radiologically and not all patients with radiological OA have symptoms. Hawker et al. reported that the degree of underuse of total hip and knee replacements is more than three times higher in women than in men, which could also be a factor in the lower incidence of THA in Swedish women. Another explanation could be that heavy physical work, which has been described as a risk factor for hip OA,[10] was commoner in Swedish men. A limitation of the current study is the difficulty of comparing index diagnoses for THA in both countries to explain differences in incidences, because no valid data on diagnoses were available from the Dutch Discharge register. The Dutch Discharge register has not been validated like the Swedish Discharge registry.[3] However, because more than 99% of hospital stays are recorded in the Dutch register and the operation codes are clearly defined, we assume that it has the same validity as the Swedish one.

The variation in the number of THA in Sweden is due to economic restrictions in the budgets for THA and does not reflect the need for this operation in the population. [3] As a result of a special program started in 1992 to guarantee treatment within three months, the incidence of THA increased greatly and the waiting times decreased. The program was discontinued in 1993, with a resultant reduction in the number of operations and an increase in waiting times. [11] In The Netherlands, the effect of extra funds to reduce waiting times in the Netherlands was also limited. More operations could be performed, but the waiting list did not decrease because of a growing need.[12]

Our results also showed that overall incidences of THA in various studies are difficult to compare due to differences in demographics, methods and reporting.[4] The reason for the similarity in overall incidences of THA in each country in 1997 is that the population is generally older in Sweden than in The Netherlands (Figure 3). The reason for the difference in normalized incidence rates in The Netherlands and Sweden (Figure 4) is less clear. The population characteristics and the health care systems are

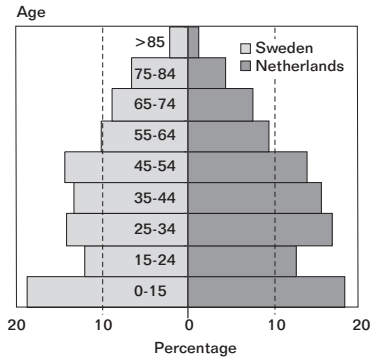


Figure 3

Population distribution in The Netherlands and Sweden in 1997. The Swedish population is older than the Dutch population.

quite similar. The higher incidence of THA in The Netherlands can be explained by a larger demand in the population, but also by differences in budgets for THA between the two countries.

Several variables are known to affect the need for arthroplasties.[13] No consistent trend of a change in the prevalence of osteoarthritis was found in Swedish or Dutch research.[8,14] Ten years ago comparable consensus statements for THA were introduced in both The Netherlands and Sweden.[15,16] However, it is not known whether these guidelines reduce the variation in surgery rates described in the literature.[17] As patients become better informed and more aware of present treatments, they may be less inclined to accept their disability and prefer treatment at an earlier stage of their disease. A British study showed, however, that the satisfaction of demand for total hip replacement in England, given the agreed criteria for surgery, is a realistic objective and requires a relatively small increase in the number of operations (7%).[18] Therefore, the influence of changing demography will be the main variable in predicting arthroplasty rates.

The predicted annual increase of THAs on the basis of demographic changes until 2020 is greater in The Netherlands (44%) than in Sweden (28%). The effect of ageing of the population is greater in The Netherlands, due to a relatively larger and more prolonged baby-boom phenomenon after the Second World War. Birell et al. reported an expected rise of 40% in the number of THAs from 1996 until 2026 in the United Kingdom.[13] All developed countries with an ageing population will witness a similar process and the THA rate must be increased to satisfy the future needs for THA. The predictions of THA rates in our study should be considered as conservative because the demand for THA in Sweden and The Netherlands is not satisfied yet, as can be concluded from the current waiting times for the procedure. The indications for THA will probably be wider if the outcome continues to improve.

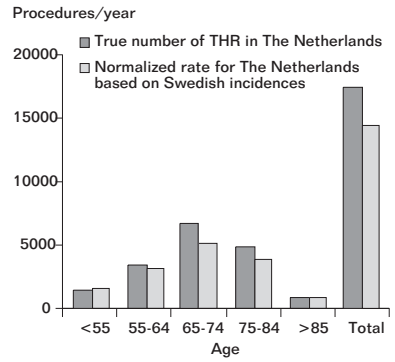


Figure 4

Normalized numbers of total hip replacement in The Netherlands. The dark grey bars indicate age-class rates for THR in 1997. The light grey bars indicate the number of procedures per year if the Swedish age-class data were applied to the Dutch population structure. Thus, incidence rates for The Netherlands and Sweden could be compared taking into account differences in population structure between the two countries.

References

1. Malchau H, Herberts P. Prognosis of total hip replacement. Update and validation of Results from the Swedish National Hip Arthroplasty Registry 1979-1998. Scientific Exhibition presented at the 67th Annual Meeting of the American Academy of Orthopaedic Surgeons, March 15-19, Orlando, USA: University of Gothenburg, Department of Orthopedics 2000.
2. Llewellyn-Thomas HA, Arshinoff R, Bell M, Williams JI, Naylor CD. In the queue for total joint replacement: patients' perspectives on waiting times. Ontario Hip and Knee Replacement Project Team. J Eval Clin Pract 1998; 4(1):63-74.
3. Söderman P, Malchau H, Herberts P, Johnell O. Are the findings in the Swedish National Total Hip Arthroplasty Register valid? J Arthroplasty 2000; 15(7):884-9.
4. Ingvarsson T, Hagglund G, Jonsson H, Jr., Lohmander LS. Incidence of total hip replacement for primary osteoarthrosis in Iceland 1982-1996. Acta Orthop Scand 1999; 70(3):229-33.
5. Friedman B, Elixhauser A. Increased use of an expensive, elective procedure: total hip replacements in the 1980s. Med Care 1993; 31(7):581-99.
6. Robertsson O, Dunbar MJ, Knutson K, Lidgren L. Past incidence and future demand for knee arthroplasty in Sweden: a report from the Swedish Knee Arthroplasty Register regarding the effect of past and future population changes on the number of arthroplasties performed. Acta Orthop Scand 2000; 71(4):376-80.

7. van Saase JL, van Romunde LK, Cats A, Vandenbroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis 1989; 48(4):271-80.
8. Danielsson L, Lindberg H. Prevalence of coxarthrosis in an urban population during four decades. Clin Orthop 1997(342):106-10.
9. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998; 57(4):203-8.
10. Vingård E, Alfredsson L, Goldie I, Hogstedt C. Occupation and osteoarthrosis of the hip and knee: a register-based cohort study. Int J Epidemiol 1991; 20(4):1025-31.
11. Socialstyrelsen. 1992 års vårdgaranti - Trender och volymer för köer och väntelistor 1990-1996. 1997.
12. Laeven AM, van Hulst BL, Mathijssen SW, van Eijk W. Het wachtlijstfonds 1999 - de derde inhaalslag ? Prismant 2000.
13. Birrell F, Johnell O, Silman A. Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. Ann Rheum Dis 1999; 58(9):569-572.
14. R.I.V.M. Volksgezondheid toekomst verkenning. De gezondheidstoestand van de Nederlandse bevolking in de periode 1950-2010. Den Haag: SDU uitgeverij, 1993.
15. Socialstyrelsen. God vård i rätt tid-1992 års vårdgaranti. 1991; 1991:11.
16. Herziening consensus totale heupprothese. Quality Institute for Health Care CBO 1994.
17. Peterson MG, Hollenberg JP, Szatrowski TP, Johanson NA, Mancuso CA, Charlson ME. Geographic variations in the rates of elective total hip and knee arthroplasties among Medicare beneficiaries in the United States. J Bone Joint Surg Am 1992; 74(10):1530-9.
18. Frankel S, Eachus J, Pearson N, Greenwood R, Chan P, Peters TJ, Donovan J, Smith GD, Dieppe P. Population requirement for primary hip-replacement surgery: a cross- sectional study. Lancet 1999; 353(9161):1304-9.

Chapter 3

Indications for THA in The Netherlands and Sweden:

surveying Swedish and Dutch orthopaedic surgeons.

Ostendorf M
Malchau H
Schrijvers AJP
Verbout AJ
Dhert WJA

Conditionally accepted Acta Orthopaedica Scandinavica 2004.

Introduction

Osteoarthritis (OA) is one of the leading causes of disability and pain in the Dutch and Swedish population[1,2] and has a high impact on quality of life also in comparison to many other chronic conditions.[3] Osteoarthritis of the hip is a common chronic disease that increases in prevalence with age.[4] Primary osteoarthritis is the most common diagnosis in patients undergoing total hip arthroplasty (THA).[5] The procedure results in marked pain relief and functional improvement.[6] Aging of the population causes an increase in the prevalence of arthritic disease and, consequently, in the prevalence of joint failure. Hence, the demand for total hip arthroplasty is expected to increase.[7,8] The need for THA in the general population is estimated to be even higher than the current provision of THA.[9] Considerable variations in the rate of THA have been described in the United States and Sweden.[10,11] These differences have been attributed to various factors, including a lack of consensus regarding the indications for these procedures. It is generally acknowledged that the indications for primary THA are joint pain, functional limitation and some evidence of intra-articular disease on the radiograph. In The Netherlands and Sweden consensus statements for the indication for THA were introduced in the previous decade.[12,13] Uniformity of indications is important to create similar chances of access to a procedure for all patients. In an earlier study, we found significant differences in incidence rates of THA in The Netherlands and Sweden, even after correction for age, gender and population structure.[8] Because one of the explanations for the difference found could be a difference in indication for the procedure, we wanted to compare indications for THA between orthopaedic surgeons in both countries.

The goals of this study were (1) to survey orthopaedic surgeons in Sweden and The Netherlands regarding the uniformity of their indications and modifying factors for primary THA, and to have these surgeons rate the likelihood of an excellent outcome based on certain critical patient factors and (2) to compare the indications and modifying factors for THA between surgeons in Sweden and The Netherlands.

Materials and Methods

Study population

The directory of the Dutch Orthopaedic Association (NOV) was used to identify all actively practicing, board-certified orthopaedic surgeons in The Netherlands. In Sweden, we used the directory of the Swedish Orthopaedic Association (SOF). This directory could not distinguish between residents, consultants and orthopaedic surgeons who had been retired. Therefore we sent questionnaires to all members and asked them to specify their precise occupation. Early 2002, an explanatory letter with the questionnaires was sent to the individuals listed in The Netherlands and Sweden.

Questionnaires

We used questionnaires developed by Mancuso et al. that were especially designed to investigate indications and modifying factors for THA and TKA, with permission from the authors.[14] Questionnaires were translated in Swedish and Dutch and pilot-tested in a group of 10 orthopaedic surgeons in The Netherlands. We report on the results of the hip questionnaire here.

The questionnaire was originally developed from review of the literature and consultation with hip and knee arthroplasty surgeons. The hip questionnaire was divided into three sections: (1) indications for surgery, (2) modifying factors for surgery and (3) likelihood of excellent outcome based on patient characteristics. Survey responses were in a multiple-choice ordinal format.

The indications for surgery used in this study are listed in Table 1. In answering these questions, orthopaedic surgeons were instructed to select the least severe level of symptom/sign that would be consistent with performing THA. For example, for the THA indication of severe pain, possible responses were one day/month, one day/week, several days/week, daily, and constant. If a surgeon chose several

Table 1

Indications for THA, with percentage of Dutch and Swedish surgeons complying to guideline criteria

	Criterium derived from Swedish and Dutch consensus statement	Percent of surgeons complying to consensus when selecting least severe level consistent with performing THA (%)	
		Sweden	Netherlands
Pain			
Severe pain	≥ several days/week	84	88
At rest	≥ several days/week	77	87
With transfer	≥ several days/week	89	76
Function			
Walking distance	<1000 meter	95	94
Need for cane or crutch	≥ one day/week	58	64
Difficulty climbing stairs	Can climb a few steps or less	79	92
Difficulty putting on shoes	At least some difficulty	89	73
Physical examination			
Range of motion	Flexion < 90 degrees	68	55
Radiographs			
Amount of joint space preserved on radiograph	< 50 % preserved	74	69

days/week, then that surgeon considered severe pain at least several days per week an indication for THA. From the existing, similar Dutch and Swedish guidelines we estimated criteria levels for the indication for THA (Table 1).
The factors that affect the decision to perform THA surgery used in this study are listed in Table 2. In answering these questions, surgeons were instructed as follows: 'Assume that all other factors are optimal, in what way would these factors modify you decision to perform THA?' Possible responses ranged from 1 to 7 with anchors of 'against' towards 1, 'neutral' at 4, and 'for' towards 7. In the data analysis, these variables were categorized as follows: 1-3, against surgery; 4, neutral; 5-7, for surgery. Thus, if a surgeon chose 2 for patient's age over 80, then that surgeon considered this factor to weigh against surgery. Because of their limited applicability in the Dutch and Swedish society, we left out 3 questions from the original questionnaire: on the

Table 2
Factors affecting the decision to perform THA

Age and comorbid factors
Age > 80 years
Age < 50 years
Comorbidity
Alcohol abuse
Weight > 90 kg.
Physical factors
History of deep venous thrombosis
Leg varicosities
Arterial insufficiency of legs
Neurologic disease of hip
Severe bone loss
Poor soft tissue coverage
Poor hip musculature
Psychologic factors
Chronic depression
Dementia
Poor motivation
Limited cooperation
Hostile personality
Function and legal factors
Return to work
Be independent
Return to sports
Unrealistic expectations

availability of home care, psychological benefit for the patient and pending court cases.
The patient characteristics that were assessed for likelihood of an excellent outcome are listed in Table 3. In answering these questions, surgeons were instructed to 'rank the likelihood of an excellent outcome in terms of the following pre-operative conditions.' Possible responses ranged from 1 to 10 with anchors of 'slight' toward 1, 'moderate' around 5-6 and 'high' toward 10. In the data analysis, these responses were categorized as follows: 1-3, slight; 4-7, moderate; and 8-10; high. Thus, if a surgeon selected 9 for osteoarthritis, then that surgeon considered there was a high likelihood of an excellent outcome if the patient had osteoarthritis.

Table 3
Patient characteristics and likelihood of excellent outcome from THA

Severe pain
Osteoarthritis
Rheumatoid arthritis
Obesity
Poor bone quality
Significant comorbidity

Statistical analysis

We calculated means and frequencies for all responses. The continuous variables, number of years in practice and number of total hip replacements performed per year were dichotomized according to the median values of these variables: In Sweden, categories were 1-16 and 17-42 years in practice and 1-30 and 31-100 THAs performed per year. In The Netherlands categories were 1-12 and 13-30 years in practice and 1-60 and 61-200 THAs performed per year.
We used Chi-square tests and t-tests to compare responses based on location (Sweden or The Netherlands), number of years in practice and number of procedures performed per year. The statistical significance level was established at $p < 0.05$. All statistical analyses were performed using SPSS 11.0.

Results

Response rates in Swedish and Dutch Survey

In 2002, a total of 1206 and 419 questionnaires were sent to the individuals in Sweden and The Netherlands respectively. In Table 4, response rates for both countries are summarized. In Sweden, 643 individuals returned questionnaires after two mailings of which 525 were consultant orthopaedic surgeons. Of all responding consultant surgeons, 111 (21%) did not perform THA. In 2002, there were 900 orthopaedic consultants practicing in Sweden according to the Swedish Orthopaedic Association, which means that 58.3% of the orthopaedic surgeons in Sweden completed the questionnaires.

Table 4
THA Questionnaire response rates in The Netherlands and Sweden

	Sweden	Netherlands
Number of questionnaires sent	1205	419
Respondents	643	263
Consultant orthopaedic surgeons	525	263
Performed THA	387	253
Did not perform THA	113	10
Did not answer question about performing THA	25	0
Nr. of surgeons in each country	±900*	419
Response rates		
Questionnaires returned	53%	63%
Consultant surgeons who completed questionnaire	58%*	63%

* Number of Swedish orthopaedic surgeons as estimated by the Swedish Orthopaedic Association (SOF).

In The Netherlands, 263 surgeons returned questionnaires after two mailings (63%). Of all responding surgeons, 10 (4%) did not perform THA. Dutch orthopaedic surgeons performed on average 64 THA per year (SD 29.3, median 60, range 5-200 prostheses). They were in practice for on average 12.1 years (SD 7.5, median 12, range 1-30 years). Swedish orthopaedic surgeons performed on average 33 prostheses per year (SD 20.4, median 30, range 1-100 prostheses) and were in practice for on average 16.8 years (SD 8.1, median 16 and range 1-42 years).

Surgeons performing THA versus surgeons not performing THA

In The Netherlands, 4% of the responding surgeons did not perform THA compared to 21% of the responding surgeons in Sweden. We compared responses from the surgeons who did (performers) and did not perform THA (non-performers) in Sweden. In The Netherlands only ten surgeons who did not perform THA filled out the questionnaire, therefore we did not analyze this group. In Sweden, we found significant differences between the two groups in ratings of the indications for THA: non-performers rated more severe levels of disability in 3 of 9 indication questions (rest pain ($p < 0.03$), difficulty in climbing stairs ($p = 0.03$), putting on shoes and socks ($p = 0.05$)). Further, non-performers found 5 of 24 modifying factors less favourably for THA than surgeons who performed THA: obesity, severe bone loss, poor soft tissue cover, chronic depression and hostile personality ($p < 0.03$). Besides, non-performers found 2 of 24 modifying factors more favourably for THA than surgeons who performed THA: desire to return to work and sports ($p < 0.01$). Lastly, surgeons who performed THA tended to rate the likelihood of an excellent outcome higher in 3 of 6 patient characteristics compared to surgeons who did not perform THA: obesity, poor bone quality or severe comorbidity ($p < 0.004$). Because of the differences found between the two groups and because the operating surgeon usually decides about the indication for THA even after referral from colleagues, we decided to perform a separate analysis for the responses of the Dutch (253) and Swedish surgeons (384) who performed THA to be able to make a valid comparison between groups.

Indications and modifying factors for THA

In Figure 1, surgeons' ratings of the indications for THA are shown for both countries. Indications are grouped in separate figures according to (a) pain, (b) and (c) function, and (d) physical examination/radiographic findings. The first graph of Figure 1 shows that most surgeons in Sweden as well as The Netherlands believe that patients should at least have severe pain daily, and rest pain several days a week. Most Swedish surgeons (88%) required transfer pain at least several days a week before performing a THA compared to 74% of the Dutch surgeons ($p < 0.001$). With regard to function, most Swedish surgeons (55%) required that patients be unable to walk more than 300 meter compared to 45% of the Dutch surgeons ($p < 0.001$). Surgeons did not uniformly require an assistive device for walking, with most surgeons indicating the use of a cane or crutch one day per week or less often to be consistent with performing THA. Further, most Swedish surgeons required that patients are able to climb a few steps while Dutch surgeons required more difficulty in climbing stairs to be consistent with THA ($p < 0.001$). Most Dutch surgeons also required use of assistance in putting on shoes and socks (59%) compared to 41% of the Swedish surgeons. ($p < 0.001$). Flexion greater than 45° was considered to be consistent with performing THA in both countries (85% of Swedish surgeons compared to 84% of Dutch surgeons). Finally, most surgeons, in Sweden (74%) as well as The

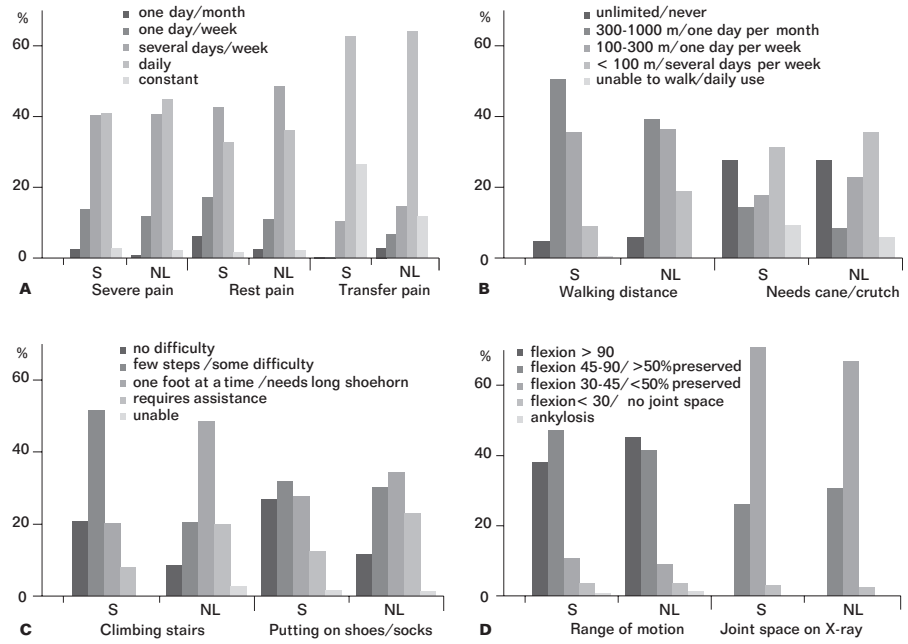


Figure 1

a. Pain as indication for THA in Sweden (S) and The Netherlands (NL).

b. Function (1) as indication for THA in Sweden (S) and The Netherlands (NL).

c. Function (2) as indication for THA in Sweden (S) and The Netherlands (NL).

d. Physical examination and radiographs as indication for THA in Sweden (S) and The Netherlands (NL).

Netherlands (69%), required that the majority of joint space be destroyed. With regard to pain symptoms and walking distance, most surgeons in Sweden and The Netherlands complied with the existing guidelines for the indication for THA while there was less agreement with regard to other functional impairments, range of motion and loss of joint space on radiographs (Table 1).

We did not find any systematic differences in Sweden or The Netherlands in rating indications for surgery when surgeons were stratified for number of THAs performed per year. When surgeons were stratified for number of years in practice, we found no systematic differences in rating indications for surgery either, apart from rest pain and use of assistive devices for walking that were more often required by Dutch surgeons who were longer in practice (35% vs 48% required assistive devices more than several days, ($p<0.02$)).

Figure 2 shows how Swedish and Dutch surgeons rated potentially important factors in considering THA. These factors have been grouped in several categories: (a) age and

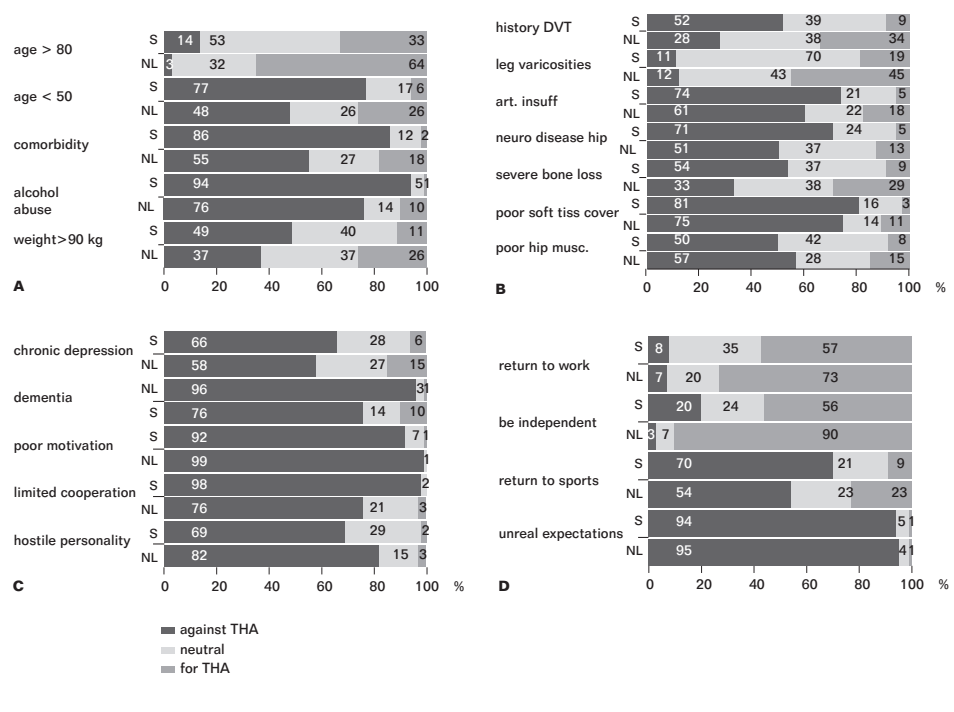


Figure 2

a. Factors affecting decision to perform THA: age and comorbid factors

b. Factors affecting decision to perform THA: physical factors

c. Factors affecting decision to perform THA: psychological factors

d. Factors affecting decision to perform THA: function and legal factors

comorbidity, (b) physical factors, (c) psychological factors and (d) function and legal factors. Surgeons indicated whether each factor would sway the decision for or against surgery or whether each factor would have no effect in the decision for THA (neutral). Most Dutch surgeons rated age over 80 as a positive factor in the decision for THR while most Swedish surgeons considered this as a neutral factor. Age less than 50, however, swayed the decision against surgery for most surgeons in Sweden as well as in The Netherlands, as would comorbidity, alcohol abuse and obesity. Swedish surgeons considered these factors significantly more unfavourable than Dutch surgeons ($p < 0.001$).

The presence of certain physical factors, such as arterial insufficiency of the legs, neurological disease of the hip, severe loss of bone stock, poor soft tissue cover and poor hip musculature would sway the decision against surgery. In Sweden, most surgeons considered a history of deep venous thrombosis (DVT) to weigh against surgery, whereas in The Netherlands an equal proportion was for, neutral and against THA if a patient had a history of DVT. Most surgeons considered the presence of leg

varicosities a neutral factor in the decision for THA. Again, Swedish surgeons considered these physical factors significantly more unfavourable in the decision to perform THA than Dutch surgeons ($p < 0.001$), except for the presence of poor hip musculature where we found no significant difference between the two countries.
The psychological factors shown in Figure 2c were found to be very important when considering a patient for THA. Chronic depression, dementia, poor motivation, limited cooperation and a hostile personality were all negative factors in the decision to perform THA for the vast majority of surgeons. When comparing Swedish and Dutch surgeons with respect to psychological factors, we found that chronic depression, dementia and limited cooperation weighed the decision more against surgery for Swedish surgeons, while poor motivation and hostile personality were rated more unfavourably by Dutch surgeons ($p < 0.01$).
On the other hand, certain functional factors such as the desire to return to work and be independent weighed very much in favour of THA for most Swedish and Dutch surgeons. In contrast, the majority of surgeons considered the desire to return to sports and unrealistic expectations of the patient as factors against THA. Swedish surgeons rated wanting to return to work or sports and the desire to be independent less favourably than Dutch surgeons ($p < 0.01$).
We did not find any systematic differences in The Netherlands in modifying factors for surgery when surgeons were stratified for number of THAs performed per year, apart from the fact that high-volume Dutch surgeons less often indicated dementia or poor soft tissues as factors swaying against surgery ($p = 0.04$ and $p = 0.03$, respectively). In Sweden, we found that high volume surgeons less often thought that obesity was a negative factor in the decision for THA ($p < 0.01$). Age over 80 or under 50 years, severe bone loss, poor soft tissue cover and poor hip musculature also less often swayed the decision against surgery in high-volume surgeons compared to low-volume surgeons ($p = 0.01$-0.05). When surgeons were stratified for number of years in practice, we found no systematic differences in modifying factors for surgery, apart from comorbidity being less often viewed as a factor against surgery by Dutch surgeons who were more than 12 years in practice ($p < 0.01$), while a hostile personality was an more unfavourable factor in the decision for THA for Swedish surgeons who were over 16 years in practice ($p < 0.001$).

Likelihood of excellent outcome from surgery

Figure 3 shows surgeons' ratings of the likelihood of an excellent outcome based on certain important patient characteristics. The vast majority of surgeons in Sweden as well as The Netherlands rated there to be a high likelihood of an excellent outcome if the patient had severe pain, osteoarthritis or rheumatoid arthritis. However, obesity, poor bone quality and significant comorbidity were considered to result only in a moderate likelihood of an excellent outcome for THA. We found no difference in rating the likelihood of an excellent outcome between Swedish and Dutch surgeons

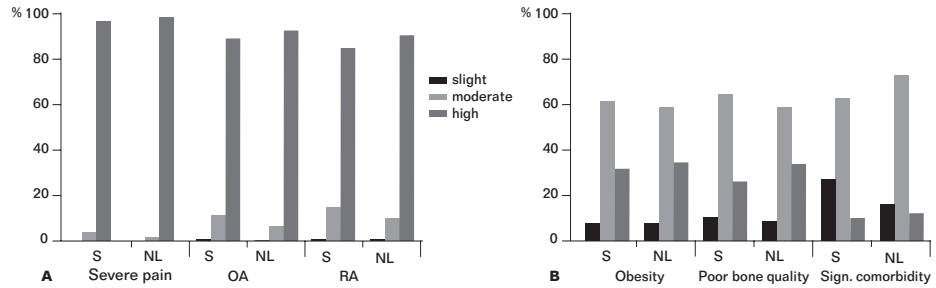


Figure 3

a. Likelihood of excellent outcome according to Dutch and Swedish surgeons: Pain, OA, RA.

b. Likelihood of excellent outcome according to Dutch and Swedish surgeons: Obesity, poor bone quality, significant comorbidity.

regarding severe pain and osteoarthritis, but Dutch surgeons tended to rate a higher likelihood of an excellent outcome than Swedish surgeons regarding obesity, poor bone quality and significant comorbidity ($p < 0.01$).
We did not find any systematic differences between surgeon groups in Sweden or The Netherlands in rating the likelihood of an excellent outcome when surgeons were stratified for number of years in practice. When surgeons were stratified for number of THAs performed per year, we found no systematic differences in rating the likelihood of an excellent outcome for Dutch surgeons, but Swedish surgeons who performed more procedures per year tended to rate the likelihood of an excellent outcome higher than surgeons who performed fewer procedures per year, with respect to preoperative rheumatoid arthritis, obesity and poor bone quality ($p < 0.02$).

Discussion

This is the first study to investigate and compare indications for THA among a large group of orthopaedic surgeons in two European countries.
Pain, functional limitation and evidence of intra-articular disease on radiographs are the primary indications for total hip arthroplasty.[15] However, there is no agreement between orthopaedic surgeons on the extent of disease that should exist before THA.[14] The Dutch guideline for the indication for THA states that candidates for THA are patients, who - in spite of adequate conservative treatment - are that much disabled by pain and stiffness of the hip joint that impairment in function and dependency on others results.[12] Swedish guidelines state that THA should be performed in patients with moderate to severe pain, serious functional impairment and decreased quality of life.[13] Both guidelines emphasize the importance of considerable joint damage on radiographs and the Swedish guideline requires that joint space be reduced with at least 50%. Despite the existing guidelines we found considerable

variations in indications for THA among surgeons in Sweden as well as The Netherlands. Apart from severity of transfer pain and joint space on radiograph, there was no agreement among surgeons over 60% per response category, which is the minimum percentage of agreement needed to designate opinions as practice guidelines.[16] While most surgeons in both countries agreed on the requirement of severe pain, rest pain and transfer pain several days per week or daily for the indication of THA, we found large variations in the required use of assistive devices, the ability in climbing stairs and the putting on of shoes and socks in the indication for THA. Apparently, activities of daily living are considered less important than the presence of pain in the decision process for THA. This trend is also found in the compliance of surgeons from both countries to existing guidelines in the indication for THA. Swedish surgeons required relatively more transfer pain and impairment in walking, while Dutch surgeons required more disability in climbing stairs and putting on shoes in the indication for THA compared to Swedish surgeons. While we did not see surgeons from Sweden or The Netherlands generally requiring more disability compared to surgeons from the other country, we did see clear differences in the rating of modifying factors in the decision for THA, where Swedish surgeons scored most factors (20 of 22) less favourably in the indication for THA compared to Dutch surgeons. This might be explained by a higher awareness among Swedish surgeons about potential risk factors such as young age because of the annual reports by the Swedish National Hip Registry, or relatively more attention for modifying factors because of budget problems for the procedure in many hospitals the past few years.[17] Also certain cultural differences between both countries, although difficult to quantify, might be important in clinical decision-making. A study by Wright et al. showed that orthopaedic surgeons' enthusiasm for total knee replacement was the dominant modifiable determinant of area variation in the use of the procedure.[18] In rating the likelihood of an excellent outcome, Dutch surgeons tended to rate a higher likelihood of an excellent outcome than Swedish surgeons regarding obesity, poor bone quality and significant comorbidity. An explanation for this could be that Dutch surgeons generally perform more THAs per surgeon, which is associated with higher ratings of the likelihood of a good outcome.[14] When we looked for differences in low- and high-volume surgeons in each country, we found that especially Swedish high-volume surgeons were more liberal concerning modifying factors in the decision for THA and rated the likelihood of an excellent outcome higher than surgeons who performed fewer procedures per year. We also found that surgeons who did not perform THA in Sweden required more morbidity in the indication for THA and were stricter concerning modifying factors in the decision for THA. A possible reason for the fact that high volume surgeons rate the modifying factors less unfavourable and tend to rate a higher likelihood of excellent outcome compared to low-volume surgeons and, even more, to non-performers could be that surgeons who perform more procedures have better outcome of surgery in the sense of better surgical technique, lower complication rates and mortality.[19,20] Further, higher volume surgeons usually also have the support of a sophisticated and experienced rehabilitation staff.[14] Lastly, surgeons who perform more procedures may also be more experienced in weighing indications and modifying factors and, therefore, select better candidates for surgery.[21] We did not find systematic differences in rating indications, modifying factors or likelihood of excellent outcome when surgeons were stratified for number of years in practice.
The original study by Mancuso et al. among 122 orthopaedic surgeons from New York City reported moderate agreement (>50%) for thresholds for pain, walking distance, range of motion and degree of joint space loss on radiographs (5 of 9 variables). Although they found consensus for some of the modifying factors, there was no consensus for all factors.[14] Concerning age and comorbid factors, American surgeons rated age > 80 years and weight > 90 kg more unfavourably and rated the other factors in-between the levels rated by Dutch and Swedish surgeons. Further, they rated physical factors more unfavourably than Dutch and Swedish surgeons. Concerning psychological factors, American surgeons rated depression, dementia and hostile personality more unfavourably while limited cooperation and poor motivation were in-between the ratings of the Dutch and Swedish surgeons. We found no structural differences between American, Dutch and Swedish surgeons concerning function and legal factors. Overall, American surgeons tended to rate modifying factors more unfavourably compared to European surgeons (12 of 22 factors), but it should be noted that the US survey was performed in 1992. Like the Swedish high-volume surgeons, American surgeons who performed more procedures per year (31-250) tended to rate the likelihood of an excellent outcome higher than surgeons who performed fewer procedures per year (5 of 6 pre-operative characteristics).
There are several possible reasons for the lack of consensus among orthopaedic surgeons found in our study. First, isolated indications are not as important as integrating and weighing several indications and modifying factors in the indication of THA.[14] Further, there is constant movement in technical developments in orthopaedic surgery and perioperative medicine, such as new prosthesis materials, improved anaesthesia, surgical technique and rehabilitation methods. In addition, there is a wide variation in other important factors that influence the postoperative course, such as hospital characteristics and available home care and rehabilitation facilities, which affect the way surgeons rate indications.[21]
The response rates were 58% and 63% for Swedish and Dutch orthopaedic surgeons, respectively. These moderate response rates may raise the possibility of response bias. However, for a community survey among medical professionals with participation on a voluntary basis these response rates are not unusual. Further, a certain amount of bias could have occurred because of cultural, translation or interpretation differences between surgeons from the two countries. Another limitation of the study could be that surgeons were asked to consider the minimum level of symptom severity, not the most common level, for which they would consider surgery. Using this approach may

have made it appear that surgeons are operating on patients with less severe symptoms than is actually the case. In addition, this study measured surgeons' responses to a hypothetical situation and may not reflect what is done in actual practice.[14]
A lack of consensus regarding the indications for THA has been cited as one reason for the variations in the rates of THA.[10,22] More unfavourable rating of modifying factors for THR in Sweden might explain partly the lower incidence (20%) of primary THR in Sweden compared to The Netherlands, even after correction for population structure.[8] It has been suggested that an increased number of surgeons per capita could increase the number of operations.[23] However, this seems not to be the case in Sweden (6.9 surgeons per 100,000 inhabitants) as compared to The Netherlands (2.5 surgeons per 100,000 inhabitants).[24] Another important reason for variations in the rates of THA is the willingness of the patient to undergo surgery for his complaints. In a study by Hawker et al., only 15% of patients with severe arthritis were willing to undergo total hip or knee arthroplasty.[25] Further, a considerable amount of the variation in the rates of THA in Sweden as well as The Netherlands the past few years can be attributed to limited financial resources for THA in both countries with considerable waiting list problems.[26-28] Also the incidence of osteoarthritis, the access to care and the percentage of people doing heavy physical work can play a role in the rate of THA in a certain area.[10,29]
Several studies showed that patients with worse pre-operative functional status had poor outcomes compared to patients with better baseline functioning. Performing surgery earlier in the course of functional decline may be associated with better outcome.[28,30,31] This could mean that the indication for THA would have to be acknowledged earlier in the disease process and that indications in existing guidelines would have to be expanded. Further, as patients become better informed and more aware of present treatment possibilities, they may be less inclined to accept their disability and prefer treatment in an earlier stage of their disease. This development, in combination with an aging population, longer life expectancy will put further pressure on health care resources. It has been argued that waiting lists for THA should be need-based, e.g. based on severity of symptoms and burden of disease instead of on order of patients on the waiting list.[32,33]
In conclusion, this study showed good compliance of Dutch and Swedish surgeons to existing guidelines for THA with respect to pain and walking ability and moderate compliance with respect to function, physical examination and radiographs. Swedish surgeons rated the majority of modifying factors more unfavourably compared to Dutch surgeons. Comparability of indications for THA is important to grant similar access to the procedure for all patients. Therefore, the results of this study are important, especially in this era when appropriateness of surgery and geographic variations in rates of procedures are under the scrutiny of administrators, health insurance companies and a more informed patient population.[14] Further, evaluation of existing guidelines is important when estimating the best time for the procedure in the disease process, especially with regard to the expected increase in need for THA.

References

1. R.I.V.M. Volksgezondheid toekomst verkenning. De gezondheidstoestand van de Nederlandse bevolking in de periode 1950-2010. Den Haag: SDU uitgeverij, 1993.
2. Åkesson K. Bone and joint diseases around the world. Sweden: a brief update on burden and priority. J Rheumatol Suppl 2003; 67:38-40.
3. Alonso J, Ferrer M, Gandek B, Ware JE, Jr., Aaronson NK, Mosconi P, Rasmussen NK, Bullinger M, Fukuhara S, Kaasa S, Leplege A. Health-related quality of life associated with chronic conditions in eight countries: results from the International Quality of Life Assessment (IQOLA) Project. Qual Life Res 2004; 13(2):283-98.
4. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998; 41(8):1343-55.
5. Malchau H, Herberts P, Eisler T, Garellick G, Soderman P. The Swedish Total Hip Replacement Register. J Bone Joint Surg Am 2002; 84 Suppl 2:2-20.
6. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. J Bone Joint Surg Am 2004; 86(5):963-974.
7. Birrell F, Johnell O, Silman A. Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. Ann Rheum Dis 1999; 58(9):569-572.
8. Ostendorf M, Johnell O, Malchau H, Dhert WJ, Schrijvers AJ, Verbout AJ. The epidemiology of total hip replacement in The Netherlands and Sweden: present status and future needs. Acta Orthop Scand 2002; 73(3):282-6.
9. Frankel S, Eachus J, Pearson N, Greenwood R, Chan P, Peters TJ, Donovan J, Smith GD, Dieppe P. Population requirement for primary hip-replacement surgery: a cross- sectional study. Lancet 1999; 353(9161):1304-9.
10. Peterson MG, Hollenberg JP, Szatrowski TP, Johanson NA, Mancuso CA, Charlson ME. Geographic variations in the rates of elective total hip and knee arthroplasties among Medicare beneficiaries in the United States. J Bone Joint Surg Am 1992; 74(10):1530-9.
11. Söderman P, Malchau H, Herberts P, Johnell O. Are the findings in the Swedish National Total Hip Arthroplasty Register valid? J Arthroplasty 2000; 15(7):884-9.
12. Herziening consensus totale heupprothese. Quality Institute for Health Care CBO 1994.
13. Socialstyrelsen. God vård i rätt tid-1992 års vårdgaranti. 1991; 1991:11.
14. Mancuso CA, Ranawat CS, Esdaile JM, Johanson NA, Charlson ME. Indications for total hip and total knee arthroplasties. Results of orthopaedic surveys. J Arthroplasty 1996; 11(1):34-46.
15. National Institute of Health (NIH). Total Hip Replacement. NIH Consensus Statement 12-14 1994; 12(5):1-31.
16. Eddy DM. Clinical decision making: from theory to practice. Designing a practice policy. Standards, guidelines, and options. Jama 1990; 263(22):3077, 3081, 3084.
17. Socialstyrelsen. Hälso- och sjukvårdsrapport 2001: sjukdomar och skador i rörelseorganer. 2001.

18. Wright JG, Hawker GA, Bombardier C, Croxford R, Dittus RS, Freund DA, Coyte PC. Physician enthusiasm as an explanation for area variation in the utilization of knee replacement surgery. Med Care 1999; 37(9):946-56.
19. Kreder HJ, Deyo RA, Koepsell T, Swiontkowski MF, Kreuter W. Relationship between the volume of total hip replacements performed by providers and the rates of postoperative complications in the state of Washington. J Bone Joint Surg Am 1997; 79(4):485-94.
20. Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States medicare population. J Bone Joint Surg Am 2001; 83 (11):1622-9.
21. Tierney WM, Fitzgerald JF, Heck DA, Kennedy JM, Katz BP, Melfi CA, Dittus RS, Allen DI, Freund DA. Tricompartmental knee replacement. A comparison of orthopaedic surgeons' self reported performance rates with surgical indications, contraindications, and expected outcomes. Knee Replacement Patient Outcomes Research Team. Clin Orthop 1994(305):209-17.
22. Malchau H, Herberts P, Garellick G. Annual Report of the Swedish Total Hip Arthroplasty Registry. Dept. Orthopaedics, Joint Replacement Unit, University of Gothenburg 2003.
23. Welch WP, Miller ME, Welch HG, Fisher ES, Wennberg JE. Geographic variation in expenditures for physicians' services in the United States. N Engl J Med 1993; 328(9):621-7.
24. Manpower data 1999, Orthopaedic Section. European Union of Medical Specialists (UEMS) 1999.
25. Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R, Wilkins A, Badley EM. Determining the need for hip and knee arthroplasty: the role of clinical severity and patients' preferences. Med Care 2001; 39(3):206-16.
26. Laeven AM, van Hulst BL, Mathijssen SW, van Eijk W. Het wachtlijstfonds 1999 - de derde inhaalslag ? Prismant 2000.
27. Socialstyrelsen. 1992 års vårdgaranti - Trender och volymer för köer och väntelistor 1990-1996. 1997.
28. Ostendorf M, Buskens E, Van Stel H, Schrijvers A, Marting L, Dhert W, Verbout A. Waiting for total hip arthroplasty: Avoidable loss in quality time and preventable deterioration. J Arthroplasty 2004; 19(3):302-9.
29. Vingård E, Alfredsson L, Goldie I, Hogstedt C. Occupation and osteoarthrosis of the hip and knee: a register-based cohort study. Int J Epidemiol 1991; 20(4):1025-31.
30. Fortin PR, Penrod JR, Clarke AE, St-Pierre Y, Joseph L, Belisle P, Liang MH, Ferland D, Phillips CB, Mahomed N, Tanzer M, Sledge C, Fossel AH, Katz JN. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002; 46(12):3327-30.
31. Holtzman J, Saleh K, Kane R. Effect of baseline functional status and pain on outcomes of total hip arthroplasty. J Bone Joint Surg Am 2002; 84 (11):1942-8.
32. Kingston R, Carey M, Masterson E. Need-based waiting lists for hip and knee arthroplasty. Ir J Med Sci 2000; 169(2):125-6.
33. Conner-Spady BL, Arnett G, McGurran JJ, Noseworthy TW. Prioritization of patients on scheduled waiting lists: validation of a scoring system for hip and knee arthroplasty. Can J Surg 2004; 47(1):39-46.

Chapter 3

Chapter 4

Results after revision THA because of deep infection

Trends and risk factors in 960 first revisions in the Swedish National Hip Registry

Ostendorf M
Eisler T
Herberts P
Fleer A
van der Tweel I
Dhert WJA
Malchau H

Introduction

Since the introduction of total hip arthroplasty, the incidence of deep infection has decreased rapidly due to the introduction of prophylactic antibiotics, improved aseptic techniques and identification of risk patients.[1] Currently, infection rates of 1 to 2 percent are being reported.[2,3] Even though the incidence of deep infection is low, the prevalence number is large because of the increasing number of THAs performed annually.[4] In 1994, over 4,000 patients were treated annually for deep infection in the United States only, which sums up to 200,000 million dollars in treatment costs.[5] Treatment of deep infection after THR is often prolonged, frequently requiring re-operations, and is expensive, which imposes significant economical burden on health care[6,7] and can entail poor clinical results.[8]

The diagnosis of infection is difficult to make in total joint arthroplasty, and it has been reported that in 20% of infected cases, no bacteria could be isolated in intra-operative cultures.[9] Especially low virulent infections can mimic the clinical symptoms of aseptic loosening, which is why it has been argued that a certain number of revisions for aseptic loosening are actually performed because of low virulent septic loosening. The protective effect of antibiotic bone cement against aseptic loosening indicates such a relation, but the clinical relevance of low-virulent infections still remains unclear.[10,11]

It has been suggested that virulent organisms such as Gram-negative bacteria and haemolytic streptococci are more difficult to eradicate.[3] Although coagulase negative staphylococci (CNS) are generally considered to be of low virulence, they are known to entrench themselves in a polysaccharide layer (glycocalyx), making them less susceptible to eradication. Also problems with multi-resistant forms of CNS have been reported.[12,13] In the treatment of infected THA, two-stage revisions have in some studies been associated with a lower recurrence risk.[14] It has been recommended that one-stage revisions should be reserved for infections caused by sensitive organisms in the elderly, infirm patient who is poorly suited for the rigor of two operations.[15] Revision surgery without the use of antibiotic containing bone cement has been associated with higher recurrence rates, in one-stage as well as two stage revisions.[14,16] The results of treatment of early postoperative and acute haematogenous infections are better than those for chronic infections.[17] The use of bone allograft at reimplantation has not been associated with higher recurrent infection rates (7.5-11%), but only small series from specialized centers have been reported.[18-20] Several patient related factors are associated with worse outcome of treatment for infection, such as poor medical condition and severe local infection with poor soft-tissue surroundings.[21,22] Although the use of systemic antibiotics and antibiotic-loaded bone cement leads to lower recurrence rates, meticulous debridement is critical and failure to adequately remove infected, necrotic, devitalized tissue and all foreign material can compromise the outcome.[23]

The purpose of the present study was to assess the demographics, bacteriology and surgical practice of revision surgery for deep infection in Sweden in the period 1979-2000. Moreover, to investigate the risk factors for repeat revision for both aseptic loosening and recurrent infection.

Patients and Methods

Since January 1979, information on any open surgical intervention after primary THA is recorded in the Swedish National Total Hip Arthroplasty Registry. All Swedish orthopaedic units participate on a voluntary basis. The current end-point is revision of either component; the validity and capture of the registry have recently been evaluated.[2] Information from medical records is stored in the revision database of the registry. Until the year 2000, 16,577 revisions have been entered.[24] Another database from the registry concerns the prophylactic measures against aseptic and septic loosening, which are reported annually per department.

We studied 960 first revisions in 952 patients performed because of deep infected THA between 1979 and 2000. Only revisions of both components were included. Infection was defined by the orthopaedic surgeon performing the revision based on erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), the number of positive intra-operative cultures and the clinical impression. Scintigraphic methods and pre-operative aspiration of joint fluid have not been widely used in Sweden.[25-28] We identified four types of infected THA after analysis of the medical records of the patients. First, patients in whom there was a strong clinical suspicion of infection of the prosthesis and with two or more positive intra-operative cultures. Second, patients in which there is a very strong clinical suspicion of infection of the prosthesis but with negative intra-operative cultures. Third, patients with strong clinical suspicion of infection but with only one positive intra-operative culture. The last group were patients with low clinical suspicion of infection but with two or more positive intra-operative cultures.[29]

The complete medical records of 868 patients (90%) were analyzed for clinical data thought to be of prognostic importance for the outcome, defined as repeat revision for deep infection, aseptic loosening or other causes.[30] The Central Bureau of Statistics, in Stockholm, Sweden, provided information on deaths (until November 1, 2002), which was also used as end-point in survival analysis.

Hospitals in Sweden are generally characterized by their catchment areas; university/regional hospitals provide care for about 300,000-600,000 inhabitants, central hospitals and rural hospitals for about 100,000-200,000 and <60,000 inhabitants, respectively.

Statistical analysis

We used ANOVA and Pearson's Chi-squared tests to test differences between groups. Prosthetic survival was estimated using Kaplan-Meier's product-limit method.[31] Survival curves (with 95% CI) were truncated when less than 50 prostheses remained at risk. Multivariable Cox proportional hazard regression models were used to study

risk factors for re-revision. Tests of the proportional hazards assumption were made for all risk factors included. In the Cox models, we used score tests to calculate p-values. Risk factors with more than two levels were represented as categorical variables. Statistical calculations were performed using SPSS software version 11.0 (SPSS Inc., Chicago, IL).

Results

Demographics

The demographic characteristics of the study group and bacterial strains cultured intra-operatively are summarized in Table 1. Age and index diagnoses did not differ in comparison to patients revised because of aseptic loosening,[24] but we found that relatively more male patients had a revision for deep infection (54.4%). Of the 890 (92%) cemented prostheses revised, 290 (33%) had been cemented with antibiotic containing bone cement at the primary procedure. Sixty-five percent of the revisions was performed in the central hospitals, 27% in university/regional hospitals and 8% in rural hospitals. Twenty-seven percent had a one-stage procedure and 56% were performed according to a two-staged protocol. Since the beginning of the 1990:s, one-stage revisions decreased in favor of staged revisions (Figure 1). Seventeen percent were permanent extractions as far as we know at present. Infected hybrid THRs ($p<0.01$) and septic loosenings caused by S. aureus ($p<0.003$) were more often managed with a staged procedure. There were no other systematic demographic or bacteriological differences between single-staged and two-staged procedures. Thirty-seven percent of the patients with a two-stage revision had a re-implantation interval between 1 and 3 months (Table 1) with a median of 2 (0.2—62) months. Patients undergoing permanent extraction of the infected prosthesis were on average 4.3 to 5.5 years older than patients who had one- or two-stage revision, respectively ($p<0.001$), had more often a previous diagnosis of hip fracture (29.6% vs. 11.6% and 10.7% in one- and two-stage revisions, $p<0.001$), and had more frequently intra-operative cultures with Gram-negative species infecting the primary THR (18.8% vs. 7.0% and 9.2% in one- and two-stage revisions, $p<0.02$).

In the 540 staged revisions, 79% were cemented revisions of which the majority (98%) was cemented with antibiotic containing bone cement. In 22% and 19% of the re-implantations, bone grafting (used as a dichotomous variable in the calculations) was performed on the acetabular and femoral side, respectively.

In 96% of the staged revisions, a local antibiotic carrier was used between resection and re-implantation. Gentamicin-loaded cement beads were most frequently used until 1985, after which several other systems were used, and since 2000 antibiotic-loaded spacers were used in 49% of the staged revisions (Figure 2).

Over the 20-year study period, flucloxacillin has been the most frequently used systemic prophylaxis in Sweden antibiotic at primary THR (Figure 4). However, there was a change in type of intravenous and oral antibiotics used to treat the peri-prosthetic infections (Figure 5), and since 1990 the types of antibiotics became more diverse. The previously more common mono-therapies decreased in favor of therapy with multiple antibiotics (37% in 1979-1984 vs. 27% in 1997-2000). Moreover, there was an increase in the use of glycopeptides (vancomycin, teicoplanin) and lincomycin (klindamycin) at the expense of cloxacillin and cephalosporines ($p<0.001$). Between 1979 and 1984, most patients were treated with cloxacillin or cephalosporins only (75%). Between 1997 and 2000 only 19% were treated with these antibiotics as a monotherapy, but still in combination regimes in about 41% of hips (Figure 5).

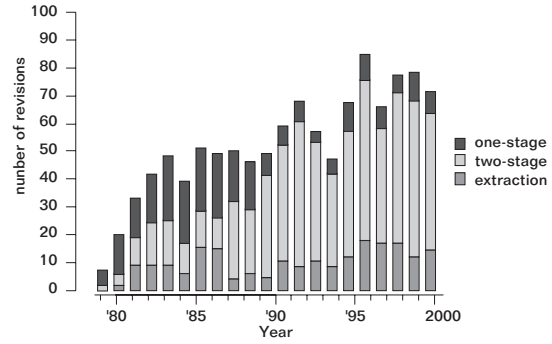


Figure 1
Type of revision for infection in Sweden 1979-2000

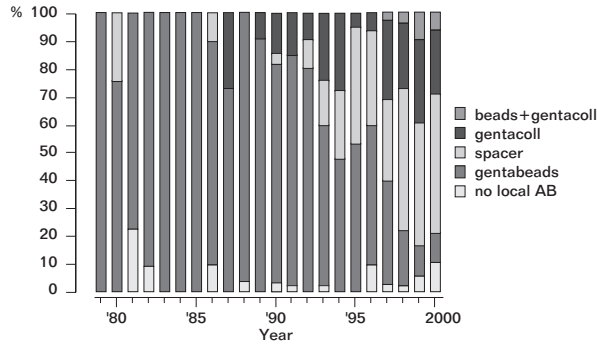


Figure 2
Local antibiotic therapy in patients undergoing two-stage revision total hip arthroplasty for infection in Sweden 1979-2000

Chapter 4

Table 1
Characteristics of revision groups

	Total group (N=960), N(%)	One-stage group (N=258)(%)	Two-stage group (N=540) (%)	Permanent extraction group (N=162) (%)
Male	523 (54.5)	58.5	55.9	43.2
Female	437 (45.5)	41.5	44.1	56.8
Clinics				
University/regional hospitals	262 (27.3)	32.9	25.9	22.8
Central hospitals	624 (65.0)	58.5	65.6	73.5
Rural hospitals	74 (7.7)	8.5	8.5	3.7
Prosthesis				
Cemented	1060 (91.5)	95.7	90.7	96.9
Uncemented	33 (2.8	2.7	3.5	0.6
Hybrid	65 (5.6)	1.6	5.7	2.5
Diagnosis				
Primary OA	699 (72.8)	75.2	75.9	58.6
Rheumatoid Arthritis	77 (8.0)	8.9	8.1	6.2
Post-traumatic	136 (14.2)	11.6	10.7	29.6
Other	48 (5.0)	4.3	5.2	5.6
Open procedure before revision, not for infection	44 (4.6)	4.7	4.8	3.7
Bacteria				
S.Aureus	107 (18.7)	10.7	22.3	23.4
CNS	207 (36.1)	39.6	36.0	29.8
Streptococci	49 (8.6)	9.6	9.2	4.3
Other Gram positive	48 (8.4)	11.2	6.8	7.4
Gram negative	57 (9.9)	7.0	9.2	18.1
Anaerobic bacteria	43 (7.7)	7.0	8.9	4.3
Mixed cultures	62 (10.8)	15.0	7.5	12.8
Mean time between 1st and 2nd stage in months (SD; median; range)			3.9 (5.4, 2.0, 0.2-61.7)	
0-1 month (N, %)			131 (24.3)	
1-3 months (N, %)			199 (36.9)	
3-6 months (N, %)			101 (18.7)	
> 6 months (N, %)			131 (24.3)	
Mean age at revision (SD; median, range)	69.3 (10.1; 71.0; 28-96)	69.1 (9.5; 71; 35-91)	68.0 (10.0; 70; 28-88)	73.9 (10.2; 76; 37-96)
Mean time between index THR and revision in years (SD; median, range)	2.9 (3.2; 1.6; 0.04-20.0)	2.5 (2.4; 1.6; 0.2-11.5)	3.0 (3.3; 1.6; 0.04-17.1)	3.5 (4.0; 2.1; 0.04-20.0)
Mean follow-up patients alive (N=434) (SD; median, range)	7.2 (5.2; 5.6; 0.0-21.5)	10.4 (6.4; 12.6; 1.1-21.5)	6.2 (4.3; 5.2; 0.0-20.6)	7.7 (5.9; 5.3; 1.4-21.1)
Mean follow-up patients deceased (N=393) (SD; median, range)	6.5 (4.7; 5.7; 0.0-21.2)	8.8 (5.1; 8.7; 0.1-21.2)	6.5 (4.2; 6.1; 0.0-19.0)	3.9 (3.4; 3.2; 0.0-17.0)
Mean follow-up patients rerevised (N=134) (SD; median, range)	3.7 (3.6; 2.6; 0.01-17.0)	4.3 (4.0; 2.6; 0.04-16.6)	3.3 (3.3; 2.4; 0.01-17.0)	

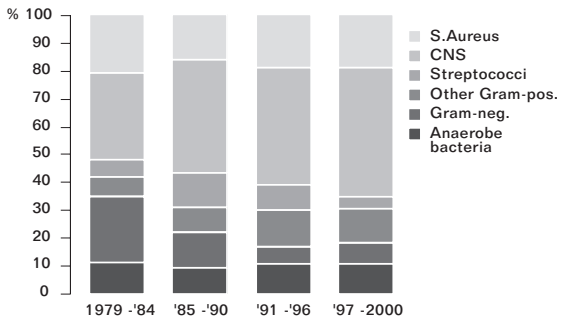


Figure 3

Type of bacteria cultured intraoperatively in revision THR for infection in Sweden

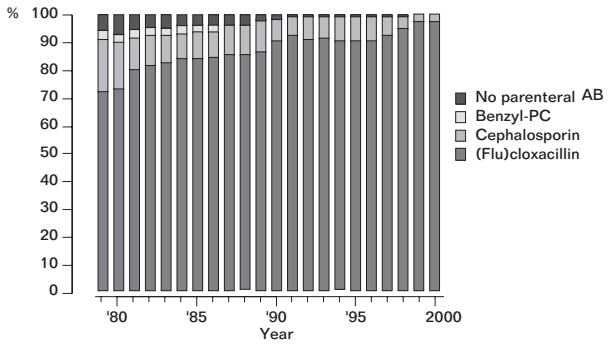


Figure 4

Systemic prophylaxis at primary THR in Sweden 1979-2000

Bacteriology

In 841 patients (73%) information on intra-operative cultures was available (Table 2). The median number of intra-operative cultures was 5 (1-20) per patient. One-hundred and fifty seven (19%) were negative. Of the positive cultures, 9% showed growth of more than one type of bacterial species. The type of bacterial species varied significantly over the study period. We found an increase in CNS (30% of positive cultures in the period 1979-1984 vs. 45% in the period 1997-2000 and a decrease of Gram-negative species (23% of positive cultures in the period 1979-1984 vs. 7% in the period 1997-2000; $p < 0.003$) (Figure 3).

Survival analysis and risk for re-revision

Fifty-eight of 798 (7.3%) patients revised (one- or two-stage revision protocols) relapsed into infection and were re-revised, implying a 90% (CI95% 88.2–93.0)

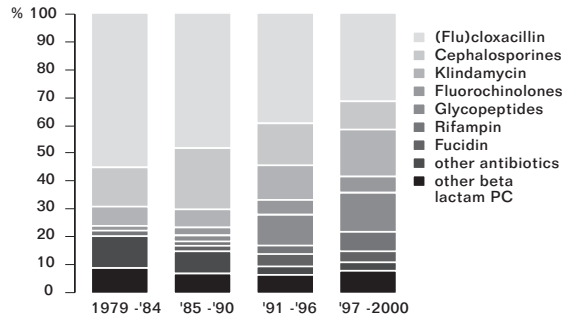


Figure 5

Systemic antibiotics (intravenously and orally) used by patients undergoing revision total hip arthroplasty for deep infection in Sweden

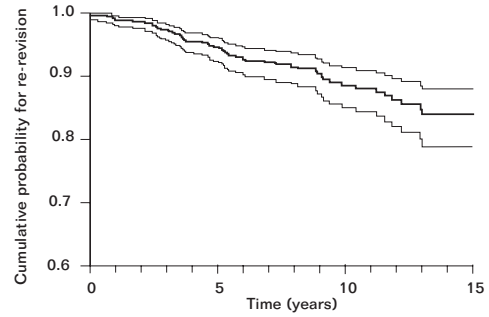


Figure 6

Survival curve (with 95% CI) of primary revision group, endpoint re-revision for infection

prosthesis survival at 10 years (Figure 6). Patients revised in rural hospitals were nearly two times (RR=1.9, CI95% 0.9–4.2) more likely to have a repeat revision for infection compared to patients revised in the central hospitals (RR=1.0) and the university hospitals (RR=1.0, CI95% 0.6–2.0) (Table 3). Patients who had minor open procedures (e.g. refixation of the major trochanter, removal of cerclage wires or heterotopic bone) after primary THR had no increased risk for septic re-revision. (RR=1.0, CI95% 0.3–3.3). The use of a one-stage revision protocol carried an increased risk (RR=1.5, CI95% 0.9–2.6) for re-revision for deep infection compared to the two-stage procedure. Patients with wound healing problems within six weeks after primary THR, revised with a single-staged procedure because of deep infection, were at particular risk (RR=1.9, CI95% 1.1–3.3) for a repeat revision. The presence of a draining sinus, the type of organism cultured intra-operatively, or the use of bone

Table 2
Bacterial species cultured intra-operatively in all revisions, including multiple species cultured in mixed infections.

	N	%
Coagulase negative staphylococci	294	39,9
Staphylococcus aureus	137	18,6
Streptococci	64	8,7
Propionibacterium acnes	52	7,1
Enterococci	44	6,0
Escherichia coli	40	5,4
Diphteroid	15	2,0
Peptostreptococci	15	2,0
Proteus mirabilis	13	1,8
Pseudomonad	11	1,5
Enterobacter	9	1,2
Klebsiella	6	0,8
Corynebacterium	6	0,8
Gram + cocci, not specified	5	0,7
Staphylococcus species, not specified	5	0,7
Tubercle bacteria	3	0,4
Anaerobic Gram - rods, not specified	2	0,3
Citrobacter	2	0,3
Acinetobacter	2	0,3
Candida	1	0,1
Salmonella	1	0,1
Bacillus subtilis	1	0,1
Haemophilus paraprophilus	1	0,1
Staphylococcus warneri	1	0,1
Bacteroides bivius	1	0,1
Aerococcus viridans	1	0,1
Aerococcus urinae	1	0,1
Actinomyces odontolyticus	1	0,1
Serratia	1	0,1
Micrococcus	1	0,1
Total	736	100

grafts did not predict re-revision for deep infection for either one- or two-stage revisions. However, the longer the interval between extraction and re-implantation the lower was the risk of a re-revision (Table 4). An interval longer than 6 months was associated with lowest risk for re-revision for deep infection (RR=0.1 (CI95% 0.02–0.94)).
Sixty out of 798 patients (7.5%) were re-revised for aseptic loosening, implying 89% (CI95% 85.7–92.0) - 84% (CI95% 79.8–88.8) prosthesis survival at 10–15 years (Figure 7). Moreover, patients younger than 65 years were 3 times more likely to undergo a re-revision for aseptic loosening than those older than 65 (RR 3.0, CI95% 1.8-5.2). Hospital type did not have a significant influence on the re-revision rate for aseptic loosening. Patients who had had minor open procedures in the hip joint after

Table 3
Cox regression model, endpoint rerevision for infection

	RR	95% CI for RR	P value
Gender	1.1	0.65 - 1.96	0.66
Age > 65	1		
< 65	1.2	0.80 - 2.47	0.24
Clinic			
Central	1		
University	1.0	0.51 - 1.88	0.95
Rural	1.9	0.89 - 4.24	0.09
Previous open procedure	1.0	0.31 - 3.26	0.98
One-stage/two-stage	1.5	0.85 - 2.58	0.16
Wound complications at index THR	1.9	1.08 - 3.34	0.03

Table 4
Two stage revisions: Cox regression model, endpoint re-revision for infection

	RR	95% CI for RR	P value
Gender	1.5	0.75 - 2.97	0.25
Age > 65 years	1		
< 65 years	1.6	0.78 - 3.12	0.21
Time between stages			
0-1 month (n=131)	0.8	0.36 - 1.74	0.56
1-3 months (n=199)	1		
3-6 months (n=109)	0.7	0.28 - 1.83	0.48
> 6 months (n=101)	0.1	0.02 - 0.94	0.04

primary THR were 2.8 times more likely to undergo re-revision for aseptic loosening compared to patients who did not (CI 1.2-6.6). Finally, the type of revision did not influence the re-revision rate for aseptic loosening, but the use of bone-graft at re-implantation did (RR=1.8, CI95% 0.9–3.4) (Table 5).
Using re-revision because of any cause as an endpoint, 134 patients out of 798 (16.8%) were re-revised implying a 79% (CI95% 75.0–82.3) - 74% (CI95% 69.6–78.8) prosthesis survival at 10 and 15 years follow-up (Figure 8).

Discussion

How should we treat patients with a deep peri-prosthetic infection, and who should take on this formidable task? This study applied an epidemiological approach to

Table 5
Cox regression model, endpoint re-revision for loosening

	RR	95% CI for RR	P value
Gender	1.1	0.65 - 1.89	0.70
Age > 65 years	1		
< 65 years	3.0	1.75 - 5.16	<0.001
Clinic			
Central	1		
University	0.9	0.51 - 1.72	0.84
Rural	1.4	0.58 - 3.33	0.46
Previous open procedure	2.8	1.17 - 6.56	0.02
One-stage/two-stage	1.2	0.66 - 2.07	0.60
Use of bonegraft at reimplantation	1.8	0.92 - 3.36	0.08

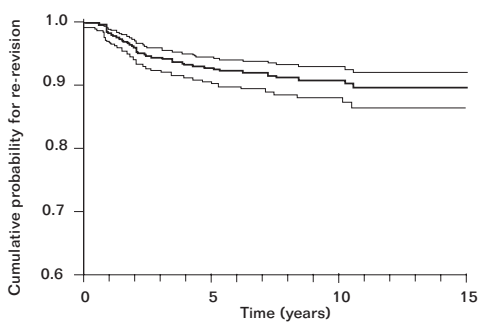


Figure 7
Survival curve (with 95% CI) for primary revision group, endpoint re-revision for aseptic loosening

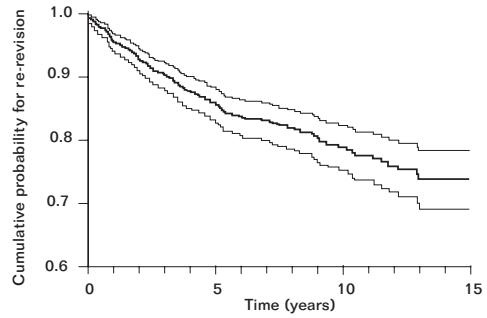


Figure 8
Survival curve (with 95% CI) for primary revision group, endpoint re-revision for any cause

investigate national results after revision THA performed for deep infection. The strength of this study lies in its sample size, and the fact that it depicts the results of the average surgeon rather than those from highly specialized centers.[32]
The weakness lies in the current end-point since the study lacks information on patients relapsing into either sepsis or aseptic loosening, but in whom a repeat revision for various reasons (e.g. co-morbidity, and unwillingness of the patient) has not been undertaken. Moreover, the registry has no information on patients treated by suppressive antibiotic therapy. Nevertheless, we are currently investigating the clinical outcome of first revision patients using patient-centered questionnaires to increase sensitivity. A further limitation to the study is that the definition of septic and aseptic loosening depended on the surgeons' clinical interpretation, and is thus highly subjective and also not uniform. On the other hand, there is no generally established definition of a deep infection, most diagnostic tools are hampered by varying accuracy, and the current low prevalence of deep infection make new diagnostic tools (such as PCR) difficult to evaluate. In addition, aseptic and septic loosening may supersede each other at different times.
Intra-operative cultures are considered the gold standard for the diagnosis.[1] Still, several errors in culture technique occur,[33] and anaerobes and other low-virulent bacteria may be lost in the culturing process.[34] Also the clinical significance of coagulase negative staphylococci can be difficult to interpret.[35] This is why the shift in bacterial species observed in this study not necessarily implies that septic loosening was caused by them. However, low-virulent infections seem to be a growing problem also in other fields,[3,36-39] which shows why there is an obvious need for improved culturing techniques.[35,40,41] Although difficult to explain, the predicting effect of minor open procedures on increased risk for re-revision due to aseptic loosening may just be an example of inferior detection techniques.
At present, there is concern about increasing microbial resistance in total joint arth-

roplasty infections, particularly regarding resistant forms of S. epidermidis.[1,9,13] Over the years, we found an increase in the percentage of CNS in intra-operative cultures and a decrease in Gram-negative species. There are concerns that the use of systemic antibiotic prophylaxis and antibiotic containing bone cement can lead to the development of antibiotic resistance.[12,42,43] Sanzén and Walder found an increased resistance to gentamicin in patients after joint replacement in most of whom gentamicin containing bone cement was used.[44] This has also been observed in an animal study.[45] Although the shift in antibiotics indicate such a relation (Figure 3), gentamicin resistance has not been shown to be more common in countries using bone cement containing antibiotics.[1]

We found relatively more males in our group that had a revision for deep infection compared to the gender ratio in the population undergoing primary THR.[24] There might be several reasons for this observation. One reason could be the relative underuse of revision surgery in female patients, because of their higher age at primary THA with associated co-morbidity or because of their decreased willingness to undergo surgery.[46] It is probably fair to assume that a more aggressive surgical treatment of implant infections is offered to patients with a more robust health status (e.g. young patients, patients without other illnesses),[47] who, in addition, have often per se improved conditions for tissue repair.

Generally, prosthesis survival after a first revision for deep infection appeared reasonable in comparison to prosthesis survival after revision for aseptic loosening.[48] Among factors studied, we believe this to be attributable to increased use of a staged revision protocol. Another explanation may be that low-virulent infections can be successfully treated in a one-stage procedure since these dormant bacteria are integrated in the surface of the implant rather than infiltrating the peri-prosthetic tissues as in more high-virulent cases. Still, the main problems seem to be recurrence of infection and aseptic loosening of the prosthesis.[3] Peri-implant infections associated with chronic inflammation, impaired circulation and osteonecrosis might in fact prove to be a causal factor for aseptic loosening, making the peri-prosthetic bone and soft tissue more susceptible to effects of micromotion, increased hydrostatic pressure and wear debris-triggered pro-inflammatory cytokines.
Interestingly, also the use of bone graft at the acetabular or femoral side at re-implantation increased the risk for re-revision for aseptic loosening. Bone grafting is technically demanding, and especially when using fresh-frozen (dead) impacted allografts for re-implantation after deep infection. Undiagnosed infections may flame up; perhaps seen as early migration but clinically misinterpreted as inferior primary stability and not infection. Results from the experienced group in Exeter suggest, however, that impaction bone grafting can be performed successfully in a staged procedure.[19] Nevertheless, differences shown in results for this technique [49] may at least partly be explained by missed low-virulent infections. Several studies have emphasized the dangers of using bone grafts in association with septic loosening.

We found that patients having a revision for deep infection in the rural clinics were more likely to have a re-revision for recurrent infection compared to patients operated in the central and the university/regional hospitals. This might be explained by the fact that high-volume surgeons have better outcome of surgery in the sense of better surgical technique, lower complication rates and mortality.[50,51] Further, larger hospitals also have increased availability of more sophisticated techniques, revision tools and intensive cooperation with other disciplines such as infection specialists. On the other hand it cannot be ruled out that indications and the severity of each case was systematically different over the hospitals and different regions in the country. Nevertheless, we believe it is fair to assume that with the small numbers of revisions for deep infection performed each year, it would be a benefit to the patients to centralize these difficult treatment procedures in specialized clinics. Changes in bacterial resistance could then perhaps also be detected earlier.[1]
The question to use a one-stage or two-stage revision protocol is controversial. While some authors feel that they are equally successful and results depend on a devoted surgeon, others believe that there are in fact very few indications today for a single-stage procedure.[52] Thus, what could perhaps be agreed upon is that the indication for each protocol may differ. We confirmed the results of other studies in that there was an increased risk for recurrent infection in patients treated with one-stage revision and experiencing wound healing problems after primary surgery.[52] Wound complications after primary THR may indicate an infection with virulent strains, which is better treated with a two-stage protocol. However, the registry has no indicator of the surgical quality of each of the protocols, and there is of course an important selection of patients to the different treatments - thus, they may not be comparable. Further studies will have to investigate this matter further. Nevertheless, on a national level a two-stage stage revision appeared to be more successful in the eradication of infection, and this is why we think that this method should be applied unless the surgeon has significant experience in one-stage procedures. However, in the elderly patient with a low virulent infection, a well-performed one-stage revision can probably still be a good option. In the treatment for deep infection, there should be equilibrium between the optimal eradication of infection and optimal treatment for the patient. In elderly patients, optimal treatment can mean retention of the components and long-term suppression of the infection with antibiotics.[53] However, the aim of treatment in these patients is amelioration of symptoms and not eradication of infection and results in the literature of suppressive therapy only are conflicting.[54,55]
Also a longer re-implantation interval was associated with lower risk for re-revision for infection. Especially re-implantation intervals longer than 6 months were associated with low risk for re-revision. Such a long re-implantation interval may be diffi-

cult to use clinically,[8] but the results might imply that hasty re-implantations are best avoided and that repeat debridements are valuable. Meanwhile, the use of an antibiotic spacer technique improves quality of life in this period, delivers high local levels of antibiotics, reduces dead space, and can facilitate reimplantation.[56-59] However, even though our study showed an increased use of such devices, we failed to prove any statistical impact of them. We believe that all these implants should be scientifically scrutinized as any other orthopaedic implant before generally released onto the market. In the difficult group of elderly patients with a diagnosis of hip fracture and Gram-negative infections, many underwent permanent extraction of their prosthesis. Although the registry does not provide such information , additional reasons were probably also poor soft tissue conditions, poor bone stock and unwillingness of the patient to undergo further surgery. However, despite a dramatic increase in osteoporotic cervical hip fractures in Sweden over the past decade [60] and a general shift towards using total hip replacement for primary treatment [24] the percentage of permanent extractions for deep infected THR remains rather constant; about 10-20% of all revisions per year over the study period. Nevertheless, this area needs to be closely monitored in the future especially with regards to methicillin resistant strains,[1,9,13] which appear to be epidemic in elderly care at least in some parts of Sweden.[61] Thus, the indication for THR in these co-morbid patients must be weighed against the risk of sepsis. When present, even a permanent extraction, with the aim of eradication, might turn out to be technically difficult in severly osteoporotic bone. However, this procedure might in fact turn out to be a better option than suppressive treatment and the functional result will not necessarily worsen as a result of this procedure. This issue will be the focus of future studies.

In conclusion, relapse into infection and aseptic loosening are the main problems encountered after revision THA because of deep infection. A staged procedure (probably with a longer re-implantation interval than previously recommended) performed at an experienced center appears to be the best strategy to avoid these complications.
Furthermore, we observed a change in the type of bacteria cultured in revisions for infected THR in the study period with a shift towards more aggressive antibiotic treatment, probably due to appearance of more resistant species in infected THR. This is a worrying trend into which further research on causes and preventive measures is needed.

References

1. Lidgren L. Joint prosthetic infections: a success story. Acta Orthop Scand 2001; 72(6):553-6.
2. Söderman P, Malchau H, Herberts P, Johnell O. Are the findings in the Swedish National Total Hip Arthroplasty Register valid? J Arthroplasty 2000; 15(7):884-9.
3. Salvati E, Della Valle AG, Masri BA, Duncan CP. The infected total hip arthroplasty. Instr Course Lect 2003; 52:223-245.
4. Ostendorf M, Johnell O, Malchau H, Dhert WJ, Schrijvers AJ, Verbout AJ. The epidemiology of total hip replacement in The Netherlands and Sweden: present status and future needs. Acta Orthop Scand 2002; 73(3):282-6.
5. Sculco TP. The economic impact of infected joint arthroplasty. Orthopedics 1995; 18(9):871-3.
6. Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, Leslie K, Bullas R. The effect of elective total hip replacement on health-related quality of life. Journal of Bone and Joint Surgery Am 1993; 75(11):1619-26.
7. Dreghorn CR, Hamblen DL. Revision arthroplasty: a high price to pay. British Medical Journal 1989; 298(6674):648-649.
8. Charlton WP, Hozack WJ, Teloken MA, Rao R, Bissett GA. Complications associated with reimplantation after girdlestone arthroplasty. Clin Orthop 2003(407):119-26.
9. Bannister GC. Prevention of infection in joint replacement. Current Orthopaedics 2003; 16:426-433.
10. Espehaug B, Engesaeter LB, Vollset SE, Havelin LI, Langeland N. Antibiotic prophylaxis in total hip arthroplasty. Review of 10,905 primary cemented total hip replacements reported to the Norwegian arthroplasty register, 1987 to 1995. J Bone Joint Surg Br 1997; 79(4):590-5.
11. Persson U, Persson M, Malchau H. The economics of preventing revisions in total hip replacement. Acta Orthop Scand 1999; 70(2):163-9.
12. Hope PG, Kristinsson KG, Norman P, Elson RA. Deep infection of cemented total hip arthroplasties caused by coagulase- negative staphylococci. J Bone Joint Surg Br 1989; 71(5):851-5.
13. Mohanty SS, Kay PR. Infection in total joint replacements: why we screen MRSA when MRSE is the problem? J Bone Joint Surg Br 2004; 86-B:266-8.
14. Garvin KL, Hanssen AD. Infection after total hip arthroplasty. Past, present and future. J Bone Joint Surg Am 1995; 77(10):1576-1588.
15. Garvin KL, Fitzgerald RH, Jr., Salvati EA, Brause BD, Nercessian OA, Wallrichs SL, Ilstrup DM. Reconstruction of the infected total hip and knee arthroplasty with gentamicin-impregnated Palacos bone cement. Instr Course Lect 1993; 42:293-302.
16. Hanssen AD, Rand JA. Evaluation and treatment of infection at the site of a total hip or knee arthroplasty. Instructional Course Lectures, American Academy of Orthopaedic Surgeons 1999; 48.
17. Tsukayama DT, Estrada R, Gustilo RB. Infection after total hip arthroplasty. A study of the treatment of one hundred and six infections. J Bone Joint Surg Am 1996; 78(4):512-23.

18. Berry DJ, Chandler HP, Reilly DT. The use of bone allografts in two-stage reconstruction after failure of hip replacements due to infection. J Bone Joint Surg Am 1991; 73(10):1460-8.
19. English H, Timperley AJ, Dunlop D, Gie G. Impaction grafting of the femur in two-stage revision for infected total hip replacement. J Bone Joint Surg Br 2002; 84(5):700-5.
20. Wang JW, Chen CE. Reimplantation of infected hip arthroplasties using bone allografts. Clin Orthop 1997; 335(335):202-10.
21. McPherson EJ, Woodson C, Holtom P, Roidis N, Shufelt C, Patzakis M. Periprosthetic total hip infection: outcomes using a staging system. Clin Orthop 2002; 403:8-15.
22. Hanssen AD, Osmon DR. Evaluation of a staging system for infected hip arthroplasty. Clin Orthop 2002; 403:16-22.
23. Callaghan JJ, Katz RP, Johnston RC. One-stage revision surgery of the infected hip. A minimum 10-year followup study. Clin Orthop 1999(369):139-43.
24. Malchau H, Herberts P, Eisler T, Garellick G, Soderman P. The Swedish Total Hip Replacement Register. J Bone Joint Surg Am 2002; 84 Suppl 2:2-20.
25. Oyen WJ, Claessens RA, van der Meer JW, Rubin RH, Strauss HW, Corstens FH. Indium-111-labeled human nonspecific immunoglobulin G: a new radiopharmaceutical for imaging infectious and inflammatory foci. Clin Infect Dis 1992; 14(5):1110-8.
26. Glithero PR, Grigoris P, Harding LK, Hesslewood SR, McMinn DJ. White cell scans and infected joint replacements. Failure to detect chronic infection. J Bone Joint Surg Br 1993; 75(3):371-4.
27. Barrack RL, Harris WH. The value of aspiration of the hip joint before revision total hip arthroplasty [see comments]. J Bone Joint Surg Am 1993; 75(1):66-76.
28. Mulcahy DM, Fenelon GC, McInerney DP. Aspiration arthrography of the hip joint. Its uses and limitations in revision hip surgery. J Arthroplasty 1996; 11(1):64-8.
29. Kamme C, Lindberg L. Aerobic and anaerobic bacteria in deep infections after total hip arthroplasty: differential diagnosis between infectious and non- infectious loosening. Clin Orthop 1981(154):201-7.
30. Sanzen L, Carlsson AS, Josefsson G, Lindberg LT. Revision operations on infected total hip arthroplasties. Two- to nine-year follow-up study. Clin Orthop 1988(229):165-72.
31. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53:457-81.
32. Garellick G, Malchau H, Herberts P. Survival of hip replacements. A comparison of a randomized trial and a registry. Clin Orthop 2000(375):157-67.
33. Powles JW, Spencer RF, Lovering AM. Gentamicin release from old cement during revision hip arthroplasty. J Bone Joint Surg Br 1998; 80(4):607-10.
34. Dupont JA. Significance of operative cultures in total hip arthroplasty. Clin Orthop 1986(211):122-7.
35. Ringberg H, Sanzen L, Thoren A, Walder M. Bacteriologic evidence of infection caused by coagulase-negative staphylococci in total hip replacement. J Arthroplasty 1998; 13(8):935-8.

36. Gaine WJ, Ramamohan NA, Hussein NA, Hullin MG, McCreath SW. Wound infection in hip and knee arthroplasty. J Bone Joint Surg Br 2000; 82(4):561-5.
37. Lentino JR. Prosthetic joint infections: bane of orthopedists, challenge for infectious disease specialists. Clin Infect Dis 2003; 36(9):1157-61.
38. Hedin G. Staphylococcus epidermidis--hospital epidemiology and the detection of methicillin resistance. Scand J Infect Dis Suppl 1993; 90:1-59.
39. Gristina AG, Costerton JW. Bacterial adherence to biomaterials and tissue. The significance of its role in clinical sepsis. J Bone Joint Surg Am 1985; 67(2):264-73.
40. Neut D, van Horn JR, van Kooten TG, van der Mei HC, Busscher HJ. Detection of biomaterial-associated infections in orthopaedic joint implants. Clin Orthop 2003(413):261-8.
41. Tunney MM, Patrick S, Gorman SP, Nixon JR, Anderson N, Davis RI, Hanna D, Ramage G. Improved detection of infection in hip replacements. A currently underestimated problem. J Bone Joint Surg Br 1998; 80(4):568-72.
42. Archer GL. Alteration of cutaneous staphylococcal flora as a consequence of antimicrobial prophylaxis. Rev Infect Dis 1991; 13 Suppl 10:S805-9.
43. van de Belt H, Neut D, Schenk W, van Horn JR, van der Mei HC, Busscher HJ. Infection of orthopedic implants and the use of antibiotic-loaded bone cements. A review. Acta Orthop Scand 2001; 72(6):557-71.
44. Sanzen L, Walder M. Antibiotic resistance of coagulase-negative staphylococci in an orthopaedic department. J Hosp Infect 1988; 12(2):103-8.
45. Thornes B, Murray P, Bouchier-Hayes D. Development of resistant strains of Staphylococcus epidermidis on gentamicin-loaded bone cement in vivo. J Bone Joint Surg Br 2002; 84(5):758-60.
46. Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R, Badley EM. Differences between men and women in the rate of use of hip and knee arthroplasty. N Engl J Med 2000; 342(14):1016-22.
47. Fisman DN, Reilly DT, Karchmer AW, Goldie SJ. Clinical effectiveness and cost-effectiveness of 2 management strategies for infected total hip arthroplasty in the elderly. Clin Infect Dis 2001; 32(3):419-30.
48. Eisler T. On loosening and revision in total hip arthroplasty. Dept. of Orthopaedics, Stockholm Söder Hospital, Karolinska Institutet. Stockholm, 2003.
49. Morgan HD, McCallister W, Cho MS, Casnellie MT, Leopold SS. Impaction allografting for femoral component revision: clinical update. Clin Orthop 2004(420):160-8.
50. Kreder HJ, Deyo RA, Koepsell T, Swiontkowski MF, Kreuter W. Relationship between the volume of total hip replacements performed by providers and the rates of postoperative complications in the state of Washington. J Bone Joint Surg Am 1997; 79(4):485-94.
51. Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, Guadagnoli E, Harris WH, Poss R, Baron JA. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States medicare population. J Bone Joint Surg Am 2001; 83-A(11):1622-9.

52. Jackson WO, Schmalzried TP. Limited role of direct exchange arthroplasty in the treatment of infected total hip replacements. Clin Orthop 2000(381):101-5.
53. Pellegrini VD, Jr. Management of the patient with an infected knee arthroplasty. Instr Course Lect 1997; 46:215-9.
54. Tsukayama DT, Wicklund B, Gustilo RB. Suppressive antibiotic therapy in chronic prosthetic joint infections. Orthopedics 1991; 14(8):841-4.
55. Goulet JA, Pellicci PM, Brause BD, Salvati EM. Prolonged suppression of infection in total hip arthroplasty. J Arthroplasty 1988; 3(2):109-16.
56. Younger AS, Duncan CP, Masri BA, McGraw RW. The outcome of two-stage arthroplasty using a custom-made interval spacer to treat the infected hip. J Arthroplasty 1997; 12(6):615-23.
57. Magnan B, Regis D, Biscaglia R, Bartolozzi P. Preformed acrylic bone cement spacer loaded with antibiotics: use of two-stage procedure in 10 patients because of infected hips after total replacement. Acta Orthop Scand 2001; 72(6):591-4.
58. Joseph TN, Chen AL, Di Cesare PE. Use of antibiotic-impregnated cement in total joint arthroplasty. J Am Acad Orthop Surg 2003; 11(1):38-47.
59. Hsieh PH, Shih CH, Chang YH, Lee MS, Shih HN, Yang WE. Two-stage revision hip arthroplasty for infection: comparison between the interim use of antibiotic-loaded cement beads and a spacer prosthesis. J Bone Joint Surg Am 2004; 86-A(9):1989-97.
60. Johnell O. The socioeconomic burden of fractures: today and in the 21st century. Am J Med 1997; 103(2A):20S-25S; discussion 25S-26S.
61. Swedres 2002. A report on Swedish Antibiotic Utilisation and Resistance in Human Medicine. The Swedish Strategic Programme for the Rational Use of Antimicrobial Agents (STRAMA) and the Swedish Institute for Infectious Disease Control 2003.

Chapter 5

Waiting for a hip:

Waiting lists and waiting times in orthopaedics

Ostendorf M
Matthijssen MAH
Dzaferagic A
Schrijvers AJP
Dhert WJA
Verbout AJ

Medisch Contact 2000;55(12):416-8.

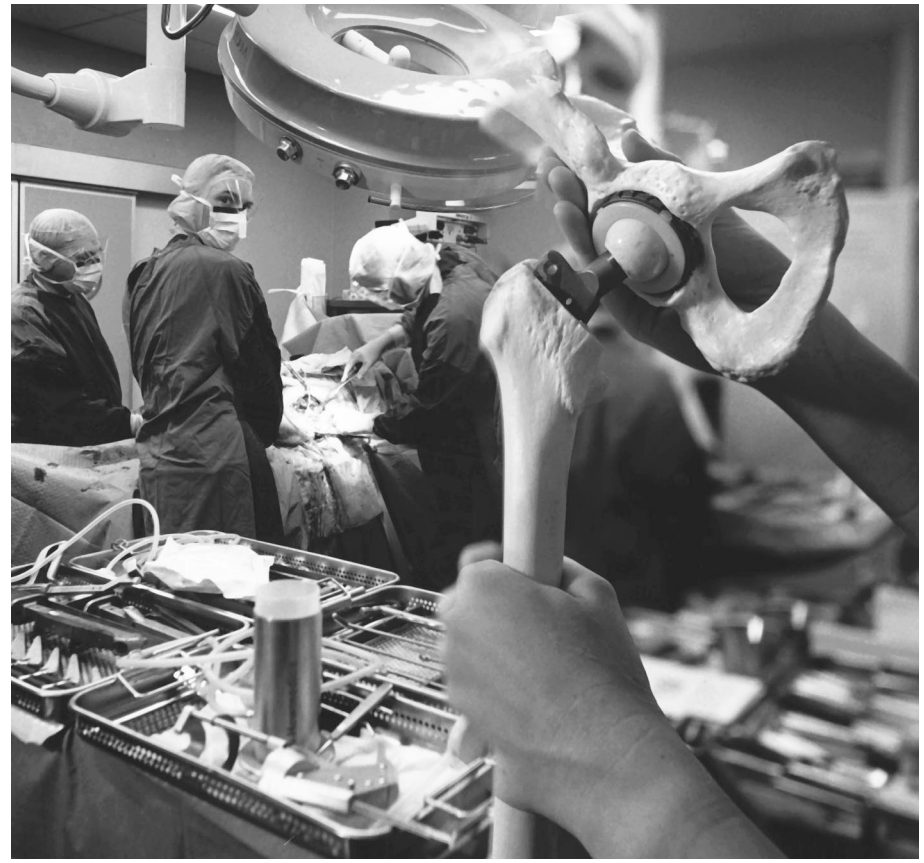
Long waiting times in health care are problematic for all parties involved. The patient has to wait long for adequate treatment, which leads to suboptimal functioning in daily life or even to the impossibility to work. For the treating physician, waiting times are frustrating because indicated treatments are postponed. For employers and health insurance companies, waiting times can cause high costs for absence through illness. The Minister of Health has decided to allocate extra incidental funding to shorten the waiting lists, but adds that these funds should have a measurable effect on the length of waiting times. Waiting times are nowadays made public to inform patients adequately. The University Medical Centre Utrecht (UMC-Utrecht) participates in a health insurance company's initiative to publish waiting times on the internet. In practice, it is more complicated to define waiting lists and waiting times than most of us assume. In this article we illustrate the complexity and limitations of waiting times and waiting list numbers.

Efficiency

Since 1998, the department of Orthopaedics of the UMC-Utrecht participates in the program 'Efficiency in Orthopaedics', which is initiated through Dutch Healthcare Research by order of the Ministry of Health. The incentive to start the program was the waiting list problem and the need to gain more insight into the aspects of efficiency improvement and quality policy in health care. The department of Orthopaedics of the UMC-Utrecht is, among other projects, involved in investigating efficiency of total hip and knee arthroplasty. The number of total hip arthroplasties (THA) has doubled during the past 17 years: from 6,750 operations in 1980 to 17,399 in 1997. Future scenarios predict an increase of 20,000 to 35,000 operations, depending on the fact whether only the aging and the growth of the population are taken into account, or that the trend of 6% increase in incidence per year continues. This large increase, in combination with the limited budgets for THA, has contributed to the long orthopaedic waiting lists that exist in many hospitals nowadays.

Discrepancy in predicted and real waiting times

In 1997 and 1998, 87 and 131 primary THAs were inserted in our hospital with a mean waiting time of 6.4 and 9.3 months, respectively. When investigating individual waiting times, we found that these waiting times are not normally distributed around the mean waiting time, but that the curve is skewed to the left. Thus, the median waiting time was shorter, namely 5.2 months in 1997.
A relatively high percentage of patients had been operated on within 3 months. These were mostly patients with increased morbidity (31%), for instance with rheumatoid arthritis or haemophilia and patients with a previous diagnosis of hip fracture. Further, we found a rather high percentage of patients that was operated on at relatively young age (< 50 years) (15%). These data showed that 78% of patients who were placed on the waiting list were also actually operated on in our hospital. Six percent were treated in another hospital and 16% were removed from the waiting list for other reasons. We found a discrepancy between the estimated and the real waiting time of patients who received a total hip arthroplasty in our hospital. The estimated waiting time (in months) is calculated by dividing the number of patients on the waiting list by the number of patients that is operated on per month. The estimated waiting time in July 1997, January 1998 and July 1998 was 12, 9 and 8 months, respectively. However, the real waiting time was 4.8, 6.4 and 5.4 months, respectively. We found a clear trend in the cumulative number of patients that was operated on per year. In 1997 one-third of patients received THA relatively soon (within 3 months). Around half of the patients waited between 3 and 9 months. A certain group of patients had to wait for a longer period, which strongly influenced the mean waiting time. These patients are mainly so-called illegitimate waiting list patients. These patients cannot undergo their operation yet because of other health problems or prefer to be operated on at a later stage (voluntary waiting time).



Photograph by Chris Timmers, Multimedia UMC Utrecht

Extra funding for THA

With the accidental funding that the Minister made available in 1998, a shift occurred in individual waiting times. It seems as if the less urgent cases, which had been on the waiting list for a longer time, had their turn now. In January 1999, considerably less patients were waiting for THA compared to January 1998 (80 vs 134 patients). However, the mean waiting time of a patient who entered the waiting list in January 1999 was not shorter than the waiting time of a patient who entered the waiting list in January 1998 (8 vs 7.2 months). The extra funding seems to have had more influence on the length of the waiting list than on the length of the waiting time. The distribution of the individual waiting times approaches a normal distribution. The shortening of the waiting list is also caused by a decrease in the number of patients that is placed on the waiting list in the outpatient clinic. The way patients are indicated for THA has not changed over the years, the decision to place a patient on the waiting list is still made in a weekly consultant meeting. Although the indication for THA is still recognized, patients are referred elsewhere because of the long waiting times or decide themselves to seek treatment elsewhere. As a result, the waiting list in our hospital is not growing, although the waiting list in another hospital is. Further, also the decision to put a patient on a waiting list is sometimes postponed, while the decision to operate is already made. Such a patient is –unjustly-not placed on the waiting list.

Validity of the waiting list

The number of patients on a waiting list and the mean waiting time are often used as performance indicators for hospitals.[1] Without further analysis, however, these parameters are of little significance. The validity of the waiting list and waiting time is limited because of 'pollution' of the waiting lists.[2] This concerns patients who are inappropriately placed on waiting lists, such as patients who are also on the waiting list of another hospital, patients who had died or who had a change in their clinical condition and no longer required the procedure. This gives rise to a source of uncertainty in reported waiting times or list lengths. Further, another limiting factor in the validity of the waiting list is the phenomenon that the treating physician already places the patient on the waiting lists who do not fulfil the indication criteria for treatment yet. On the other hand, patients who should be on the waiting list are not placed on it because of long waiting times and are referred to another hospital. Further, the characteristics of waiting lists differ per hospital. A hospital with an area of special interest will have shorter waiting times for the procedures involved, which is at the expense of other procedures.

'Iceberg' phenomenon

To shorten a long waiting time with a constant demand for the procedure, the supply of the procedure should be increased. The number of operations that can be performed depends on many factors, such as the availability of hospital beds, operation rooms and

personnel, nurses, surgeons, prosthesis budgets and rehabilitation possibilities. For an increase in the number of procedures, the whole cascade of factors needs to be adjusted. The extra funding for total hip and knee arthroplasties has contributed to the increase in the number of procedures in 1998 in the UMC-Utrecht, but the planned number of procedures was not reached. Causes were lack of operating time, available hospital beds and nurses. This has been called the iceberg phenomenon: every time a problem is solved, the next problem reaches the surface.[3]

An important question now is whether waiting times are reduced if the number of procedures – by improvement in efficiency or by the availability of extra funding-increases. A British study showed that although waiting time decreased because of the increased number of operations, the number of patients on the waiting list did not decrease because more patients were simultaneously placed on the waiting list. [1]

There ís a certain risk that the indication for a procedure changes because of the increased capacity. Thus, critical research into indications and guidelines for orthopaedic surgery for common diseases is necessary.[4] Although there exists a consensus statement for THA by the Quality Institute for Health Care,[5] this statement is not consequently applied by al surgeons.

A widening of the indication for THA cannot always be prevented. Modern medicine offers more possibilities and, moreover, patients do not consider all given alternatives acceptable. Medicine is evaluating together with changes in society, such as the increased individualization and emancipation of patients.[6]

Public waiting times

Currently, we perform a study in the UMC-Utrecht on the effect of waiting times on the health status of patients. It is a sensible idea to give patients more insight into different waiting times in an area, so that they can make informed choices.[7] However many patients prefer to receive a treatment from their 'own' doctor and are willing to wait longer for that reason.[8] Studies in the United States and Canada show that the medical success of a procedure is not related to the length of the waiting time. [9,10]

The need for THA will continuously increase, if only because of the demographic changes in the population. To decrease the size of waiting lists in a structural way, the number of procedures should be increased by additional funding and improved efficiency of the procedure. In several hospitals, the length of stay for THA has decreased significantly because of initiatives like short-stay wards and collaboration with rehabilitation units.[11] In the future it may be possible to improve efficiency by performing certain procedures in special clinics.

The management of long waiting times is a difficult problem. The increasing demand for care is an important causal factor in the waiting list issue and financing one obstacle will not solve the problem. Further, we showed that measuring and comparing waiting times to investigate the effect of incidental funding has shown to be unreliable. Waiting list figures should thus be interpreted carefully.

References

1. Newton JN, Henderson J, Goldacre MJ. Waiting list dynamics and the impact of earmarked funding. Bmj 1995; 311(7008):783-5.
2. Louwers JPA, Lekkerkerker J. Het ZBZ-project en wachtlijsten; Inzicht in vraag en aanbod met behulp van een input-output model. Technische Universiteit Eindhoven 1998.
3. Kroonen A. Analyse van wachtlijsten: minder mist geeft betere resultaten bij de bestrijding. Plan van Aanpak Wachttijden Zorgsector. De Nieuwe Buitensociëteit, Zwolle, 1998. pp. 7-12.
4. Nordberg M, Keskimaki I, Hemminki E. Is there a relation between waiting-list length and surgery rate? Int J Health Plann Manage 1994; 9(3):259-65.
5. Herziening consensus totale heupprothese. Quality Institute for Health Care CBO 1994.
6. De Tweede Kamer, Brief van de Minister en de Staatssecretaris van WVC: Financieel Overzicht Zorg 1994.
7. Structurele Aanpak van Wachttijden in de Zorgsector. 1998.
8. Rijken M, de Bruin A. Wachttijden in de gezondheidszorg: Ervaringen en meningen van mensen met chronische aandoeningen. NIVEL 1998.
9. Coyte PC, Wright JG, Hawker GA, Bombardier C, Dittus RS, Paul JE, Freund DA, Ho E. Waiting times for knee-replacement surgery in the United States and Ontario. N Engl J Med 1994; 331(16):1068-71.
10. Williams JI, Llewellyn Thomas H, Arshinoff R, Young N, Naylor CD. The burden of waiting for hip and knee replacements in Ontario. Ontario Hip and Knee Replacement Project Team. J Eval Clin Pract 1997; 3(1):59-68.
11. Barber TC, Healy WL. The hospital cost of total hip arthroplasty. A comparison between 1981 and 1990. Journal of Bone and Joint Surgery Am 1993; 75(3):321-5.

Chapter 6

Waiting for total hip replacement:

Avoidable loss in quality time and preventable deterioration

Ostendorf M
Buskens E
van Stel HF
Schrijvers AJP
Marting LN
Dhert WJA
Verbout AJ

Journal of Arthroplasty 2004;19(3):302-9.

Introduction

Over the past decade, the number of total hip arthroplasty (THA) procedures performed in the Netherlands has shown a steady increase.[1] The incidence of THA per 100,000 inhabitants increased from 71 procedures in 1986 to 112 procedures in 1997. Despite the steady growth, the surgical capacity has not kept up with the demand for the procedure. As a result, recurrent problems have occurred with lengthy waiting lists for THA. Similar problems have been reported in other countries with a publicly funded health system like the United Kingdom,[2] Canada[3,4] and New Zealand.[5] Some Dutch patients may have even gone abroad because of the long waiting lists.[6]
THA has proven to be an effective treatment to improve function and relieve pain in the hip secondary to severe osteoarthritis or other diseases affecting the hip joint.[7] Extended waiting time for surgery may prolong pain and difficulties in physical function.[8] Little is known about the burden of illness that patients experience while waiting for major joint arthroplasty.[4] For several years researchers agree that, in the eyes of payers, patients and even fellow providers, patient-centred outcome data should define clinical success and group expertise in general.[9]
The purpose of this study was to evaluate the effect of waiting times for THA on the health status of patients and the THA outcomes, using patient-centred questionnaires. We investigated the effect of waiting times on loss in quality-adjusted life years (QALYs) and additional burden perceived, and the effect of waiting times and pre-operative function scores on post-operative outcome scores.

Patients and Methods

Between April 1999 and September 2000, 161 patients were recruited from 3 different hospitals, 1 university hospital and 2 regional hospitals. Patients were asked to participate after the treating orthopaedic surgeon put their names on the waiting list for primary THA. Patients were excluded from the study if they were younger than 18 years, had rheumatoid arthritis, were unable to complete questionnaires, or had received a THA in the contralateral hip. Waiting time was calculated from the date the patient was placed on the waiting list to the date of surgery. An informed consent form and 4 questionnaires were mailed to patients initially consenting, together with a letter with additional information about the study. After final written consent was obtained, the same 4 questionnaires were sent to the patients pre-operatively and 3 and 12 months after surgery. If necessary, the missing questions were completed using a telephone call. To allay anxieties about discussing care and their experiences, patients were assured that the treating surgeons would not receive any information on patients' answers. The study received Ethics Committee approval in all hospitals.

Quality of life assessment

The health status of the patients was measured on 4 different measuring moments, 1] at inclusion, 2] pre-operatively and 3] three and 4] twelve months after the operation. We used two disease-specific questionnaires - the Oxford Hip Score and the Western Ontario and McMaster Osteoarthritis Index (WOMAC) - and 2 general health questionnaires - the SF-36 Health Status Questionnaire and the EuroQol questionnaire. To assess the effects of comorbidity on outcomes, the Index of Coexistent Disease (ICED) was scored using data from the medical records of the patients.[10]

Oxford Hip Score (OHS). The OHS[11] is a disease-specific measure consisting of 12 questions assessing pain and function of the hip in relation to different activities of daily life. Each question is answered by ticking a position on a 5-point ordinal scale. Responses are then totalled to obtain a score between 12 and 60. A low score indicates less burden in terms of pain and function, a high score indicates greater burden. The OHS was developed specifically to assess outcomes of hip arthroplasty and has been shown to be consistent, reproducible, valid and sensitive to clinical change.[11] The Dutch OHS has shown to be valid and reliable in measuring outcome in THA patients.[12]
Western Ontario and McMaster Universities Osteoarthritis Index. The WOMAC[13] is a well-known disease-specific instrument widely used for measuring outcome after THA.[9] Using a Likert scale, patients rate themselves on multiple items grouped in 3 domains: pain, stiffness, and difficulty in functioning. The maximum score is 20 points for pain, 8 for stiffness and 68 points for clinical function. A low score indicates lesser difficulty, and a high score indicates greater difficulty.
Short Form-36. The Medical Outcomes Study 36-Item short-form health survey (SF-36)[14] is a widely used measure of general health status.[15] The 36 items are grouped in 8 domains designed to fully represent the World Health Organization (WHO) definition of health: physical functioning, role limitations because of physical problems, role limitations because of emotional problems, social functioning, bodily pain, vitality, mental health and general health perceptions. The SF-36 was translated and validated for use in the Dutch population.[14]
EuroQol. The EuroQol[16] is a standardized generic instrument yielding an index score of utility for health-related quality of life. Patients describe their health status in 5 dimensions, each presented in 3 levels of severity. The domains are mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Scores include no problems, some/moderate problems or extreme problems, respectively. Furthermore, patients are asked to value their current health state on a thermometer-type visual analogue scale from 0 (worst imaginable health state) to 100 (best imaginable health state), the EQ_{VAS}. The score (utility) for quality of life obtained with the EuroQol, the $EQ\text{-}5D_{index}$, is a weighted overall value from the societal perspective, which is a weighted aggregated score for the five EuroQol dimensions of quality of life. Weights were deri-

ved on the basis of responses from a sample of 3000 households in the United Kingdom.[17] The EuroQol is recommended for use in combination with other, more-detailed generic measures, such as the SF-36.[18]

Data analyses.
Missing values for the WOMAC were imputed with the mean values of the remaining items of the subscale, conditional on 40% of the items having values.[19] Missing values for the SF-36 subscales were imputed with the mean values of the remaining items for the subscale, according to the SF-36 guidelines. Because the OHS and the EuroQol are short questionnaires with, respectively, 12 and 5 items, missing values were not imputed and questionnaires with missing values were not included in the analysis.
We calculated change scores and effect sizes for all outcome measures in the waiting time (inclusion score minus preoperative score). Effect sizes show the relative magnitude of change measured by different instruments.[20] For the current study, we calculated effect sizes as the difference between the inclusion and pre-surgical scores divided by the standard deviation of inclusion scores. Effect sizes of 0.2, 0.5 and 0.8 are regarded as indicating small, medium and large degrees of change, respectively.[21] Paired t-tests were used to compare differences in change scores.
We performed univariate and stepwise multiple linear regression analysis to estimate the effect of independent variables on change in scores or on absolute scores. The assumptions of multiple linear regression analysis were checked. All explanatory factors were checked for collinearity. Firstly, to predict deterioration during waiting time, we determined the association between the dependent variables (change in pain and function in the waiting period as measured with the WOMAC and Oxford Hip Score) and the independent variables (gender, age, waiting time, comorbidity (ICED), diagnosis, previous surgery in the same hip, the occurrence of other walking limitations (Charnley B and C),[22] living together or alone, job situation and education. Further, to predict improvement and outcome, we determined the relationships between the independent variables and post-operative outcome (respectively change in pain and function 3 months after the total hip replacement as measured by the WOMAC questionnaire and Oxford Hip Score and absolute WOMAC pain and function scores and Oxford Hip Score 1 year post-operatively). The independent variables used in model development were the same as mentioned earlier, with the additional use of the preoperative WOMAC and Oxford Hip scores.
To assess the difference in post-operative function scores between patients with high and low pre-operative function, we divided our patient cohort in 2 groups using the median pre-operative WOMAC physical function score (44 points) as a cut-off. Differences between groups were tested with independent t-tests. The statistical significance level of all procedures was established at $p< 0.05$. Statistical analyses were performed using SPSS software version 10.0 (SPSS Inc., Chicago, IL).

Results

Surgery was performed on 55 men and 106 women with a mean age of 68.4 years (standard deviation [SD], 9.7; range, 35.9-88.8). Thirteen additional patients were asked to participate in the study but refused. We compared this group with the patients who agreed to participate in the study. The only variables we could compare were age, gender and diagnosis. It turned out that the group that refused was significantly older (75.2 years, SD, 7.5; range, 62.1-84.7).
Most patients (136, 85%) were diagnosed with advanced primary osteoarthritis as the main cause of hip disease. In 13 patients (8%), the primary diagnosis was avascular necrosis. Post-traumatic arthritis was diagnosed in 6 patients (4%), osteoarthritis secondary to congenital dysplasia in 4 patients (2%) and Legg-Calvé-Perthes disease in 2 patients (1%). We evaluated the main outcomes of the study in the patient group with and without primary osteoarthritis as primary diagnosis. Because we found no difference in outcomes between both groups, we included all diagnoses in our analysis to have a more representative sample from which to extrapolate our results. The mean waiting time for operation was 6 months (SD, 3.6, range, 3 weeks-18 months). Forty-seven patients (29%) waited less than 3 months; 38 patients (24%) waited between 3 and 6 months; and 76 patients (47%) waited longer than 6 months. Three patients preferred another clinic for treatment because of the length of the waiting time.
We did not request predonated autologous blood from patients in any of the participating clinics. Consequently, this did not influence length of waiting times in the study. Fourteen patients underwent a second THA on the contralateral hip within a year from the first surgery and two patients received a total knee arthroplasty. These patients were excluded from the follow-up analysis because of the effect of the second surgery on outcome scores.
One hundred and twenty patients (75%) were affected in one hip (Charnley category A), 34 patients (21%) in both hips (Charnley category B) and 7 patients (4%) had other conditions directly impeding mobility (Charnley category C). The levels of comorbidity according to the ICED were none in 58 patients (36%), mild in 74 patients (46%), moderate in 23 patients (14%) and severe in 6 patients (4%).
Eighteen patients (11%) had undergone previous surgery on the hip joint. Of these, 8 patients underwent an internal fixation for a femoral neck fracture, 6 patients an osteotomy of the femur, 2 an osteotomy of the femur and a shelf procedure and 2 a core decompression of the femoral head. One hundred and forty nine of the THAs performed (92%) involved a fully cemented prosthesis, 11 were uncemented (7%) and one was a hybrid prosthesis (1%). The majority of patients received some form of higher education (114, 71%). Nineteen patients (12%) were still employed, the others (142, 88%) were retired or unable to work. Furthermore, most patients were living with someone else (115, 71%) and 46 (29%) were living alone.
At the preoperative, 3-month and 1-year measurements, we had, respectively, 143, 142, and 124 patients with complete questionnaires. At the 1 year measurement 3 patients had died, 2 also received a total knee arthroplasty and 14 received a contrala-

teral hip arthroplasty. We analyzed the characteristics of patients who were excluded from the analysis because they did not return questionnaires. We found that 3 months after surgery, patients who did not respond were a little older (72.2 years vs 68.0 years, NS) and had significantly lower scores on the SF Mental Health score (65.8 vs 75.8, $p < 0.05$) and the EQ_{VAS} score at inclusion (49.4 vs 62.3, $p < 0.02$). Their decreased mental wellbeing and their lower estimation of their own health might be an explanation for not returning their questionnaires. However, because only 18 from the 161 patients were excluded from the analysis we do not expect this group to have a pronounced effect on the results. Table 1 shows the results for the OHS, WOMAC, and SF-36 subscales and EQ_{VAS} and EQ-$5D_{index}$ at the different measuring moments. Most outcome scores show a limited deterioration during the waiting period. The pre-operative scores for the SF-36 are substantially below the norms for older individuals in either the Netherlands or the United States.[14,23]
We noted a large improvement in scores at 3 months, and more so 1 year post-operatively, not only in the OHS, WOMAC and the physical domains of the SF-36 but also in the emotional role and social functioning scores of the SF-36. Table 2 shows the mean change scores and effect sizes for the OHS, WOMAC, SF-36 subscales and EQ_{VAS} and EQ-$5D_{index}$ during the waiting period. Although the effect sizes found were all small changes, the differences in the disease-specific scores and domains were statistically significant. A striking finding was the considerable improvement in the mean EQ-$5D_{index}$ score of 0.42 at 1 year after surgery (Table 1, Figure 1). For 100 patients not treated, this implies a loss of 3.5 QALYs per month that they remain on the waiting list. Accordingly, with an average waiting time of 6 months, a loss of 21 QALYs per 100 patients can be anticipated.

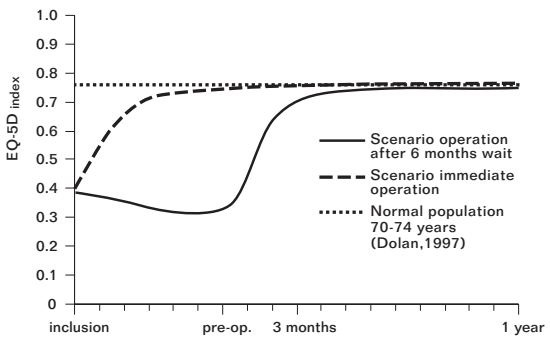


Figure 1
Different scenarios for EQ-$5D_{index}$ scores over time. The black line shows the development in mean EQ-$5D_{index}$ scores at the different measuring moments in the study population. The grey dotted line shows the mean EQ-$5D_{index}$ score over time if patients would be operated immediately after placement on the waiting list. The black dotted line shows the mean EQ-$5D_{index}$ score for the general population in the age group 70-74 years (0.76) (Dolan 1997).

Table 1.
Mean scores (SD) for OHS, WOMAC, SF-36 subscales, EQ_{VAS} and EQ-$5D_{index}$ at all measuring moments.*

Method of assessment	Inclusion scores N = 161	Preoperative scores postoperative N = 143	Three-month follow-up postoperative N = 142	One-year follow-up N = 124
OHS	41.3 (7.9)	43.1 (7.4)	22.1 (8.9)	19.2 (8.0)
WOMAC				
Pain	11.2 (3.7)	11.8 (3.7)	3.8 (4.0)	3.7 (4.4)
Stiffness	4.8 (1.7)	5.0 (1.6)	2.7 (1.7)	2.2 (1.8)
Function	41.7 (11.0)	43.9 (11.1)	18.6 (14.1)	16.4 (15.7)
SF-36				
Bodily pain	31.3 (17.4)	27.1 (16.9)	63.3 (25.7)	70.1 (27.2)
Physical function	23.4 (17.0)	19.5 (15.2)	53.0 (23.7)	58.4 (26.6)
Role physical	12.7 (25.7)	9.3 (21.9)	38.9 (42.4)	52.9 (46.2)
Role emotional	51.7 (46.4)	48.1 (44.5)	74.5 (38.5)	65.2 (44.2)
Mental health	74.5 (20.3)	71.7 (20.8)	80.1 (19.0)	78.5 (20.5)
Vitality	54.1 (22.1)	53.9 (20.7)	65.3 (20.5)	67.9 (21.5)
Social function	56.1 (28.3)	53.2 (29.8)	76.3 (26.0)	78.1 (26.3)
General health	61.8 (20.5)	61.6 (21.2)	68.2 (20.6)	65.3 (20.8)
EQ_{VAS}	60.7 (20.5)	57.4 (21.9)	74.8 (16.9)	74.4 (17.8)
EQ-$5D_{index}$	0.39 (0.31)	0.33 (0.32)	0.71 (0.26)	0.75 (0.28)

* For the Oxford Hip Score and the WOMAC, higher scores represent more pain and limitation. For the SF-36 and EQ-5D, higher scores represent better health status.

Deterioration in OHS and WOMAC scores during waiting times

We found a high variability in inclusion and pre-operative measurements. For instance, 38 patients showed improvement in OHS during the waiting time, and 20 others had the same scores as at the inclusion measurement. However, most (85) patients had a decreasing OHS during waiting period. In fact, this effect was clearly visible in all disease-specific scores, and overall we saw a significant negative difference between the inclusion and pre-operative measurements (Table 2). In multiple regression analysis, length of waiting time was the only significant predictor for deterioration in scores, but the associated R-square values were very low (R^2 is 0.07 for OHS, 0.05 for WOMAC pain score and 0.03 for WOMAC function score, respectively).

Table 2
Mean change scores (SD) and effect sizes for OHS, WOMAC, SF-36 subscales, EQ_{VAS} and EQ-$5D_{index}$ in waiting time (inclusion minus preoperative scores)*

	Mean change score (SD)	Effect size	p Value
OHS	-2.1 (6.0)	-0.25	< 0.001
WOMAC			
Pain	-0.9 (3.1)	-0.26	0.008
Stiffness	-0.2 (1.4)	-0.09	0.048
Function	-2.4 (8.4)	-0.22	0.004
SF-36			
Bodily pain	3.5 (15.2)	0.20	< 0.001
Physical function	3.8 (13.9)	0.22	< 0.001
Role physical	3.6 (23.8)	0.13	NS (0.06)
Role emotional	3.6 (44.4)	0.08	NS (0.1)
Mental health	2.8 (12.3)	0.15	0.032
Vitality	-0.4 (15.6)	0.02	NS (0.9)
Social function	3.2 (25.0)	0.12	NS (0.1)
General health	0.2 (17.5)	0.03	NS (0.7)
EQ_{VAS}	2.1 (23.0)	0.10	NS (0.3)
EQ-$5D_{index}$	0.05 (0.28)	0.15	0.027

* For the Oxford Hip Score and the WOMAC, higher scores represent more pain and limitation.
For the SF-36 and EQ-5D, higher scores represent better health status.
Effect sizes show the relative magnitude of change measured by different instruments (the difference between the inclusion and pre-surgical scores divided by the standard deviation of inclusion scores).
Effect sizes of 0.2, 0.5 and 0.8 are regarded as indicating small, medium and large degrees of change.

Predictors for improvement in OHS and WOMAC scores. Significant predictors for improvement in OHS 3 months postoperatively were male gender ($p = 0.004$) and high WOMAC function scores (low function scores preoperatively, $p < .001$), with a total R^2 of 0.15. Significant predictors for improvement in WOMAC pain and function scores were Charnley A (no other walking limitations) ($p = 0.008$ and $p = 0.004$) and high OHS (high pain and low function score) ($p < 0.001$ for both scores), with a total R^2 of 0.14 and 0.15 for WOMAC pain and function score, respectively. Thus, patients with low pre-operative scores showed more improvement 3 months post-operatively. Waiting time was not a determinant for change in OHS, WOMAC pain or WOMAC function 3 months post-operatively.

Predictors for OHS and WOMAC scores postoperatively. To investigate the effect of independent variables on outcome scores 1 year after surgery, we calculated univariate and multivariate regression models for OHS, WOMAC pain score and WOMAC function score 1 year after THA. Predictors for high (worse) Oxford hip score ($R^2 = 0.19$) were worse Oxford hip scores preoperatively ($p = 0.005$) and the occurrence of other walking limitations (Charnley B and C) ($p < 0.001$). Predictors for high (worse) WOMAC function scores 1 year postoperatively ($R^2 = 0.29$) were worse preoperative WOMAC function scores, lower education and the occurrence of other walking limitations (Charnley B and C) ($p < 0.001$, $p < 0.001$ and $p = 0.002$, respectively). Worse preoperative WOMAC pain scores were also predictive ($R^2 = 0.40$) for high (worse) WOMAC pain score ($p < 0.001$), aside from female gender ($p < 0.001$) and other walking limitations (Charnley B and C) ($p < 0.001$).

Table 3
Disease-specific scores in high and low preoperative function groups preoperatively and 1 year after THR*

	High preop. function Mean (SD)	Low preop. function Mean (SD)	Difference (high-low) Mean	p value
Oxford Hip Score				
Preoperative	37.6 (5.4)	48.5 (5.1)	-10.9	<0.001
1 year	17.7 (7.4)	20.7 (8.3)	-3.0	0.047
Difference (1 year – preop)	-19.9	-27.8	7.9	<0.001
WOMAC pain				
Preoperative	9.7 (3.0)	13.9 (2.6)	-4.2	<0.001
1 year	2.0 (2.7)	4.9 (4.8)	-3.0	<0.001
Difference (1 year – preop)	-7.7	-8.9	1.2	NS (0.15)
WOMAC function				
Preoperative	33.8 (7.8)	52.3 (5.6)	-18.5	<0.001
1 year	9.7 (11.7)	21.2 (15.9)	-11.5	<0.001
Difference (1 year – preop)	-24.1	-31.1	7.0	0.011

* For the Oxford Hip Score and the WOMAC, a higher score represents more pain and limitation.
At 1 year after surgery, patients with higher preoperative function still had significantly less pain and better function compared with those with lower preoperative function. Both groups of patients improved significantly but the low-function group improved more in Oxford Hip Score and WOMAC function score than the patients in the high-function group.

The outcome scores have a large variability (Table 1). Consequently, some patients with short waiting times had scores and outcomes worse than those with longer waiting time and vice versa. We did not find statistical differences in the absolute outcome scores between waiting groups (< 3 months, 3-6 months, > 6 months) at any of the measuring moments. In multiple regression analysis, waiting time was not a determinant for OHS, WOMAC pain or WOMAC function scores at 1 year post-operatively.
We compared the OHS and WOMAC pain and function scores between the patient groups with high and low preoperative physical function.[24] At 1 year after surgery, patients with higher pre-operative function still had significantly less pain and better function compared with those with lower preoperative function (Table 3). Both groups of patients improved significantly but the low-function group improved more in OHS and WOMAC function score than the patients in the high-function group.

Discussion

The present study clearly shows improvement in disease-specific pain and function scores as well as in general health scores as a result of THA, but it also shows deterioration in outcome scores as a result of the waiting time for the procedure.
In our prospective study, we noted a small but significant deterioration in the OHS, WOMAC, SF-36 bodily pain, physical function and mental health score and the EQ-$5D_{index}$ during the time patients spent waiting for surgery. The large and significant improvement found at 3 months after surgery was maintained at 1 year after surgery. A substantial loss of quality adjusted life years occurred with postponed surgery.

We found that the patients who declined to participate were significantly older than the patients who did participate (75.2 vs 68.4 years). Completing the numerous questionnaires may have been too large a burden for these elderly patients. Because it concerns a small number and because age was not a predictive factor in regression analyses for postoperative improvement and absolute disease-specific scores, we do not believe this has led to selection bias.
Patients who did not return questionnaires 3 months after surgery were a little older and had significantly lower scores on the SF Mental Health score and the EQ_{VAS} score at inclusion. Their decreased mental well-being and their lower estimation of their own health might explain the nonresponse. However, because only 18 from the 161 patients were not included in the analysis we do not expect this group to have a pronounced effect on the results.
THA compares favourably with other interventions with respect to gain of quality-adjusted life years. For instance, patients with occlusive artery disease showed an increase of 0.24 in the EQ-$5D_{index}$ score 1 year after primary stent placement compared with 0.42 in our patient group 1 year THA.[25] Resource allocation, however, does not reflect this observation.[26]
We did not find statistical differences in the absolute outcome scores between waiting groups (< 3 months, 3-6 months, > 6 months) at any of the measuring moments. However, waiting time was the only predictive factor for deterioration in scores in multiple regression analysis. Identifying the types of patients who experience intensified pain or functional decline while waiting for surgery is important to minimize the overall burden of illness while waiting for major total joint arthroplasty.[4] Apart from waiting time, we found no other predictive variables for deterioration in scores.
A prospective study from Canada showed no negative impact of waiting time on the amount of pain and dysfunction experienced.[4] However, this study only considered changes of at least 10% in WOMAC pain and function or SF-36 bodily pain and physical function scores to be considered a change in health status. The average waiting time in this study was comparable (4.5 months) to the waiting time in our study (6 months). Moreover, the authors used a logistic regression model to investigate the effect of waiting time and other variables on deterioration in pain and function scores. This has less power compared to linear regression analysis, the method used in our study. A study from New Zealand did not find deterioration in disease specific health or general quality of life with the duration of wait, but this may have been an effect of the cross-sectional design of the study.[5]

We found no effect of length of waiting times on any of the outcome scores 3 months or 1 year post-operatively. Nevertheless, our results corroborate other studies on surgical outcomes of major joint arthroplasty. We found that patients with greater pain and dysfunction pre-operatively, that is, at a later point in the natural history of functional decline of the disease process, did not improve to the level achieved by those with higher pre-operative function 1 year post-operatively.[24,27] Traditional orthopaedic practice, to delay surgery as long as possible, should be re-evaluated in view of these findings, especially in elderly patients.[24,28] Although we found only a small deterioration in outcome scores in an average waiting time of 6 months, it can be speculated that length of waiting time becomes a determinant for lower postoperative scores with even longer waiting times. Apart from this effect, clearly, if functionally impaired patients can achieve consistently excellent relief of symptoms, disproportionately long waiting times impose an avertable burden of disability. [3]
Apart from preoperative disease-specific scores and other walking limitations, female gender and lower education level were determinants for respectively higher pain score (WOMAC pain) and lower function levels (WOMAC function score) 1 year post-operatively. Similar results have previously been reported. Others also found gender and education to be predictors for lower outcome scores after total hip replacement.[24,29]
A limitation of our study was that definition of waiting time only included the period from the moment the patient was put on the waiting list to the surgery and not the

time between the general practitioner's referral and orthopaedic consultation. Patients may wait a long time before they consult the general practitioner for their hip complaints, and further delay occurs in the period before patients are referred to an orthopaedic surgeon.[30] Ideally, the entire waiting time from general practitioner referral to surgery should be monitored. However, information on waiting periods other than the official waiting time on the surgical waiting list is rarely available.[4] Further research on the deterioration process in patients with osteoarthritis of the hip is needed to estimate the best time for THA.

In conclusion, we found that patients' health scores deteriorate while waiting for THA. Moreover, a substantial and avoidable loss of quality-adjusted life years results from waiting for surgery. Patients experience weeks and months of rapidly reversible pain and disability while awaiting hip arthroplasty.[3]
Although our data do not show a direct effect of waiting time on post-operative outcomes, patients in a later phase of their disease process do not improve to the level achieved by the patients with better pre-operative function.

References

1. Ostendorf M, Johnell O, Malchau H, Dhert WJ, Schrijvers AJ, Verbout AJ. The epidemiology of total hip replacement in The Netherlands and Sweden: present status and future needs. Acta Orthop Scand 2002; 73(3):282-6.
2. Rigge M. Quality of life of long wait orthopaedic patients before and after admission: a consumer audit. Qual Health Care 1994; 3(3):159-63.
3. Williams JI, Llewellyn Thomas H, Arshinoff R, Young N, Naylor CD. The burden of waiting for hip and knee replacements in Ontario. Ontario Hip and Knee Replacement Project Team. J Eval Clin Pract 1997; 3(1):59-68.
4. Kelly KD, Voaklander DC, Johnston DW, Newman SC, Suarez-Almazor ME. Change in pain and function while waiting for major joint arthroplasty. J Arthroplasty 2001; 16(3):351-9.
5. Derrett S, Paul C, Morris JM. Waiting for elective surgery: effects on health-related quality of life. Int J Qual Health Care 1999; 11(1):47-57.
6. Sheldon T. Dutch patients travel to Spain for orthopaedic surgery. Bmj 2001; 322(7302):1565C.
7. Rissanen P, Aro S, Sintonen H, Slatis P, Paavolainen P. Quality of life and functional ability in hip and knee replacements: a prospective study. Qual Life Res 1996; 5(1):56-64.
8. Coyte PC, Wright JG, Hawker GA, Bombardier C, Dittus RS, Paul JE, Freund DA, Ho E. Waiting times for knee-replacement surgery in the United States and Ontario. N Engl J Med 1994; 331(16):1068-71.
9. Ritter MA, Albohm MJ. Overview: maintaining outcomes for total hip arthroplasty. The past, present, and future. Clin Orthop 1997(344):81-7.
10. Greenfield S, Apolone G, McNeil BJ, Cleary PD. The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. Med Care 1993; 31(2):141-54.
11. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire in the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78 (2):185-190.
12. Hoefnagels NH, Dhert WJ, Gosens T, Bulstra SK, de Vet HCW, van Langelaan EJ, Okhuysen-Vonk S, Geesink RGT. De Oxford Heup Score, een Nederlandstalige patiënten-vragenlijst ter evaluatie van het resultaat van een totale heupoperatie gebaseerd op de vertaling van "the 12-item questionnaire on the perceptions of patients about THR". Tijdschrift voor Orthopaedie 2002; 9:5-16.
13. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. Journal of Orthopaedic Rheumatology 1988; 1:95-108.

14. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998; 51(11):1055-68.
15. Ware JE, Jr., Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. Med Care 1995; 33(4 Suppl):AS264-79.
16. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990; 16(3):199-208.
17. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35(11):1095-108.
18. Brooks R. EuroQol: the current state of play. Health Policy 1996; 37(1):53-72.
19. Bellamy N. WOMAC Osteoarthritis Index User Guide IV. Queensland, Australia, 2000.
20. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27(3 Suppl):S178-89.
21. Cohen J. Statistical Power Analyses for the Behavioral Sciences. Vol. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
22. Charnley J. Low Friction Arthroplasty of the Hip. Berlin Heidelberg New York: Springer-Verlag, 1979.
23. McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994; 32(1):40-66.
24. Fortin PR, Clarke AE, Joseph L, Liang MH, Tanzer M, Ferland D, Phillips C, Partridge AJ, Belisle P, Fossel AH, Mahomed N, Sledge CB, Katz JN. Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. Arthritis Rheum 1999; 42(8):1722-8.
25. Bosch JL, van der Graaf Y, Hunink MG. Health-related quality of life after angioplasty and stent placement in patients with iliac artery occlusive disease: results of a randomized controlled clinical trial. The Dutch Iliac Stent Trial Study Group. Circulation 1999; 99(24):3155-60.
26. Williams A. Seeking priorities in health care: an economist's view. J Bone Joint Surg Br 1991; 73(3):365-367.
27. MacWilliam CH, Yood MU, Verner JJ, McCarthy BD, Ward RE. Patient-related risk factors that predict poor outcome after total hip replacement. Health Serv Res 1996; 31(5):623-38.
28. Garellick G, Malchau H, Herberts P, Hansson E, Axelsson H, Hansson T. Life expectancy and cost utility after total hip replacement. Clin Orthop 1998(346):141-51.
29. Young NL, Cheah D, Waddell JP, Wright JG. Patient characteristics that affect the outcome of total hip arthroplasty: a review. Can J Surg 1998; 41(3):188-95.
30. Smith T. Waiting times: monitoring the total postreferral wait. Bmj 1994; 309(6954): 593-6.

Chapter 7

Patient-reported outcome in total hip replacement:

A comparison of five instruments of health status

Ostendorf M
van Stel HF
Buskens E
Schrijvers AJP
Marting LN
Verbout AJ
Dhert WJA

Journal of Bone and Joint Surgery Br 2004;86 (6):801-8.

Introduction

Total hip arthroplasty (THA) is an effective treatment which improves function and relieves pain in the hip secondary to severe osteoarthritis or other diseases which affect the joint.[1] In an environment of limited resources, health status questionnaires are of particular importance when comparing the cost-to-benefit ratio of medical interventions.[2] In the evaluation of THA, especially the Western Ontario and McMaster osteoarthritis index (WOMAC) and the SF-36 health status questionnaire (SF-36) have been recommended as patient centred questionnaires.[3] However, the 12-item Oxford hip score (OHS) and the EuroQol questionnaire have also been suggested, the first because of its brevity and site-specificity and the second because it allows comparisons to be made with the effect of other health-care interventions.[4,5] The psychometric characteristics of these four questionnaires have been reported extensively in the literature. Our aim was to define a minimal set of outcome measures to assess health status in THA. We therefore compared the baseline characteristics and sensitivity to change of the OHS, the WOMAC, the SF-36, the SF-12 health status questionnaire (SF-12), which is derived from the SF-36, and the EuroQol questionnaire (EQ-5D). We also investigated the pre- and post-operative distribution of the EuroQol scores. Finally, the SF-36 and EuroQol scores were compared with published data on age and gender-matched norms (personal communication).[6,7]

Patients and Methods

Between April 1999 and September 2000, 147 patients were recruited from three different hospitals, one university hospital and two regional hospitals. They were asked to participate in the study after their orthopaedic surgeon had put them on the waiting list for primary total hip replacement. Patients were excluded from the study if they were under the age of 18 years, suffering from rheumatoid arthritis, were unable to fill out the questionnaires or if they had previously undergone a THA on the contralateral hip. Before and at three and 12 months after operation four questionnaires were mailed to consenting patients together with a letter with additional information about the study and an informed consent form. If necessary, missing questions were answered by telephone. To allay anxieties about discussing their care and their experiences with surgery, patients were assured that their surgeons would not receive any information on patients' answers to any questions. Approval was obtained from the Ethical Committees of all the participating hospitals.

Assessment of quality of life

The health status of the patients was measured before and at three and 12 months after operation. We used two disease-specific questionnaires, the OHS and the WOMAC, and two general-health questionnaires, the SF-36 and the EQ-5D. We calculated the SF-12 scores from the answers to the SF-36 questionnaire as described by Kosinski.[8] In order to assess the effects of comorbidity on outcomes, the Index of Co-

Existent Disease (ICED) was scored using data from the medical records of the patients.[9]

Oxford hip score.[10] The OHS is a disease-specific measure consisting of 12 questions, which assess pain and function of the hip in relation to different activities of daily life. Each question is answered by ticking a position on a five-point ordinal scale. Responses are then totalled to obtain a score between 12 and 60. A low score indicates lesser difficulty and a high score greater difficulty. The OHS was developed specifically to assess the outcomes of hip replacement surgery and has been shown to be consistent, reproducible, valid and sensitive to clinical change.[10,11] The Dutch OHS has shown to be valid and reliable in measuring outcome in patients who have had THA.[12]

Western Ontario and McMaster Universities (WOMAC) osteoarthritis index.[13] The WOMAC is a well-known, disease-specific measure which is widely used for measuring outcome after THA.[3] Using a Likert scale, patients rate themselves on multiple items which are grouped in three domains, pain, stiffness and difficulty in function. The maximum score is 20 points for pain, 8 for stiffness and 68 points for clinical function. A low score indicates lesser difficulty and a high score greater difficulty.

SF-36. This is a widely used measure of general health status.[14] The 36 items are grouped on eight scales which are designed to represent fully the WHO definition of health: physical functioning, role limitations due to physical problems, role limitations due to emotional problems, social functioning, bodily pain, vitality, mental health and perceptions of general health. Aaronson et al.[6] translated the test into Dutch and tested its validity and reliability.

EuroQol. This is a standardized generic instrument for describing and valuing health-related quality of life, and identifies 243 possible health states.[15] Patients describe their own health state on five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. One of three levels of severity is chosen for each dimension: no problems, some/moderate problems or extreme problems for the dimension in question. A single weighted utility score, the EQ-$5D_{index}$, is calculated from the five dimensions.[7] Perfect health and death have utility scores of one and zero, respectively, and states worse than death (< 0) are possible. Patients are also asked to value their current state of health on a thermometer scale from 0 ('worst imaginable') to 100 ('best imaginable'). We have re-scaled this score, the EQ-$5D_{vas}$, to zero to one for comparison with the EQ-$5D_{index}$. The EuroQol therefore generates two overall values (utilities) for the quality of life (one from a social perspective, the EQ-$5D_{index}$, and the other from the patient's perspective, the EQ-$5D_{VAS}$). It is recommended for use in combination with other more detailed generic measures, such as the SF-36.[15]

SF-12. The SF-12 Health Survey, a short form with 12 items, is a generic measure of health status which was developed to provide an alternative to the SF-36 for the purpose of monitoring large samples from general and specific populations. The SF-12 is

attractive because it is short, easy to administer, has proven reliability and validity, and can be printed on one page. It has norm-based scores, which means it has a mean of 50 and an SD of 10 in the general population.[16] We did not use the actual questionnaire in the study, but we calculated the SF-12 physical (PCS-12) and mental summary scores (MCS-12) from the responses to the SF-36 in order to compare the SF-12 scores and changes in scores with the other questionnaires.[8]

Analyses of data

Missing values for the WOMAC were approximated as long as 40% of the items had values, using the mean values of the remaining items of the subscale. Missing values for the SF-36 subscales were also approximated according to the SF-36 guidelines.[17] Because the OHS, the EQ-5D and the SF-12 are short questionnaires with 12, five and 12 items, respectively, missing values were not approximated and questionnaires with missing values were not included in the analysis. We assessed the pre- and post-operative floor and ceiling effects of all questionnaires as a measure of the validity of their content. Patients who achieve the best possible score on a questionnaire cannot demonstrate improvement on a subsequent application of the same questionnaire even if they have improved clinically.[18] This is referred to as the ceiling effect. The floor effect is the opposite.

Statistical analysis. To assess construct validity, we calculated Spearman correlation coefficients to examine the relationship between instruments at the pre-operative measurement. We calculated changes in scores and effect sizes for all the measures of outcome (one-year post-operative score minus the pre-operative score). Effect sizes show the relative magnitude of a change as measured by different instruments.[19] For this study we calculated effect sizes as the difference between the pre-operative and one year post-operative scores, divided by the SD of the pre-operative scores. Effect sizes of 0.2, 0.5 and 0.8 were regarded as indicating small, medium and large degrees of change respectively.[20] The Wilcoxon signed-rank and the Mann-Whitney-U test were used to compare the change in scores among the different groups of patients. Spearman correlation coefficients were calculated to examine whether change in one measure corresponded with a change in other measures. Each patient was matched to his or her expected population norm for the SF-36 and EuroQol based on a five-year age interval and gender (personal communication).[6,7]

Results

Of the 147 patients, 114 patients had complete questionnaires at the one-year measurement. Three had died, two had a revision for infection, one had a femoral fracture, 2 had undergone total knee replacement and 12 had undergone contralateral hip replacement. Those who had undergone a second operation after their total hip replacement were excluded from the follow-up analysis because of the effect of the second operation on the outcome scores. Thirteen patients did not return their questionnaires. We used only the results of the 114 patients with completed questionnaires in the analysis. Details of the patients are given in Table 1. We compared the characteristics of patients who did not return their questionnaires with our analyzed group. We found that those who did not respond at the one-year follow-up were more likely to have a post-operative complication (Chi-squared test $p \leq 0.05$) and had significantly lower scores on the WOMAC pain score (3.2 vs 7.0, $p < 0.05$) and the SF physical functioning score (34.2 vs 57.0, $p < 0.05$) at three months postoperatively.

All outcome scores showed a significant improvement at one year after operation ($p < 0.001$) apart from the SF-36 General Health score ($p = 0.09$) (Table 2). Very large effect sizes were found for the disease-specific measures and the physical domains of the SF-12, SF-36 and the EQ-$5D_{index}$ (1.3 to 3.0). Effect sizes for the other domains of the SF-36 and EQ-$5D_{VAS}$ were small to moderate (0.1 to 0.8).

We compared changes in scores for patients in Charnley class A (affected in one hip) and Charnley classes B and C (affected in both hips/other conditions directly impeding mobility)[21] (Table 3). Patients in Charnley class A showed a greater change in their OHS, WOMAC pain and function, the physical domains of the SF-36 and the EQ-$5D_{VAS}$. An interesting finding was that for the Charnley B and C group, the effect size for the OHS more than doubled the effect sizes of WOMAC pain and physical function. We did not find a significant correlation between outcome scores and the ICED co-morbidity score.

We calculated pre-operative correlations and the correlations of change among the questionnaires in order to assess the construct validity of the questionnaires. There was a large number of significant correlations between the outcome scores at the pre-operative measurement. We found especially high correlations between the OHS, WOMAC, the physical domains of the SF-36 and EQ-$5D_{index}$ (Table 4). We also found high correlations of change between the disease-specific scores and the disease-specific domains and a low correlation of SF general health with the OHS and the WOMAC pain and stiffness (Table 5). The physical and mental scores of the SF-12 (PCS and MCS) showed moderate to high correlations and correlations of change with, respectively, the physical and mental domains of the SF-36.

We compared the mean SF-36 and EQ-5D scores before and at one year after operation with the age- and gender-matched SF-36 and EQ-5D scores found in general population samples.[6,7] (Figure 1 and 2). The scores at one year after operation approached the scores for the general population for both questionnaires, although we still

Table 1.
Characteristics of patients before THR

	Number	Percentage
Female	71	62.3
Diagnosis		
Primary OA	95	83.3
Avascular necrosis	9	7.0
Post-traumatic	5	4.4
Dysplasia	3	1.3
Legg-Calvé-Perthes	2	1.8
Charnley class		
A	97	85.1
B	13	11.4
C	4	3.5
Comorbidity (ICED)		
None	39	34.2
Mild	53	46.5
Moderate	19	16.7
Severe	3	2.6
Previous surgery		
None	100	87.7
Internal fixation	7	6.1
Femur osteotomy	4	3.5
Femur osteotomy and shelf procedure	2	1.8
Core decompression of femoral head	1	0.9
Prosthesis		
Cemented	103	90.4
Uncemented	11	9.6
Mean age (SD; range)	67.6 years (10.1; 35.9 to 88.8)	
Mean waiting time (SD; range)	6 months (3.6; 3 weeks to 18 months)	

found significant differences for the physical function, and physical and emotional role domains of the SF-36. The mental health, vitality and general health scores were even higher than the age- and gender-matched scores from the general population (Figure 1). We found large floor effects at the baseline for the SF-36 role physical

Table 2
Mean scores (SD) and effect sizes for OHS, WOMAC, SF-12, SF-36 subscales, EQ-$5D_{VAS}$ and EQ-$5D_{index}$ preoperatively and 1 year after THR. Higher OHS and WOMAC scores indicate poorer results, while higher SF-36 and EQ-5D scores show better results.

	Mean score preop (SD)	Mean score 1 year (SD)	Effect size	Floor/ceiling effect pre-op. (%)		Floor/ceiling effect post-op. (%)	
OHS	42.5 (7.9)	19.0 (7.7)	3.0	0	0	0	13.5
WOMAC							
Pain	11.7 (3.5)	3.6 (4.3)	2.3	0.9	0	0	36.2
Stiffness	4.9 (1.7)	2.1 (1.7)	1.7	5.6	0	1.0	16.2
Function	42.7 (11.4)	15.6 (15.0)	2.4	0	0	0	7.8
SF-12 PCS	30.5 (8.3)	45.6 (9.6)	1.8	0	0	0	0
SF-12 MCS	41.4 (12.5)	49.7 (12.2)	0.7	0	0	0	0
SF-36							
Bodily pain	28.5 (16.5)	72.1 (26.3)	2.6	11.2	0	1.9	32.4
Physical function	22.0 (16.9)	60.7 (26.0)	2.3	9.5	0	1.0	1.9
Role physical	11.0 (24.0)	55.3 (47.1)	1.9	78.3	2.8	33.7	41.3
Role emotional	51.7 (43.5)	68.7 (43.0)	0.4	37.7	36.8	24.0	60.6
Mental health	73.5 (19.4)	79.4 (19.8)	0.3	0	5.7	0	12.4
Vitality	55.1 (21.3)	68.4 (21.0)	0.6	0	0	0	4.8
Social function	55.7 (29.7)	79.2 (25.7)	0.8	10.3	12.1	2.9	46.7
General health	63.8 (19.8)	66.7 (20.4)	0.1	0	0.9	0	4.8
EQ-$5D_{VAS}$	0.59 (0.22)	0.75 (0.18)	0.7	0	0	0	4.9
EQ-$5D_{index}$	0.35 (0.31)	0.76 (0.27)	1.3	0	0	0	39.8

(78%) and role emotional (38%) domains (Table 2). At one year after operation most outcome scores showed a pronounced ceiling effect, mainly large (> 30%) although some were moderate (> 10%). The floor effects for SF-36 role physical (34%) and role emotional (24%) remained (Table 2).
We found several features in the EQ-5D scores. First, there was a binomial distribution of the pre-operative EQ-$5D_{index}$ score, which changed to a skewed normal distribution at one year after operation. Scores were especially scarce between 0.25 and 0.5. Furthermore, there were 15 patients with pre-operative EQ-$5D_{index}$ scores less than zero. Particularly notable was the pre-operative discrepancy and post-operative agreement between the EQ-$5D_{VAS}$ and EQ-$5D_{index}$. The pre-operative discrepancy was greatest in patients with a low EQ-$5D_{index}$ score (Figure 3).

Chapter 7

Table 3
Change scores (SD) for OHS, WOMAC, SF-12, SF-36 subscales, EQ-5D$_{VAS}$ and EQ-5D$_{index}$ preoperatively and twelve months after THR for: A. Charnley class A patients (N=97), B/C. Charnley class B/C patients (N=17).Change scores and effect sizes were calculated for all questionnaires comparing preoperative scores and scores after 1 year. Higher OHS and WOMAC scores indicate poorer results, while higher SF-12, SF-36 and EQ-5D scores show better results.

	A.			B/C.			
	Mean change score preop vs. 1 year (SD)	Effect size preop vs. 1 year (SD)	p value	Mean change score preop vs. 1 year (SD)	Effect size preop vs. 1 year (SD)	p value	p value Difference A and B/C
OHS	24.7 (8.7)	3.1	< .001	17.9 (11.6)	2.4	< .001	=.026
WOMAC							
Pain	8.8 (4.3)	2.6	< .001	4.8 (3.3)	1.1	< .001	< .001
Stiffness	2.9 (1.8)	1.7	< .001	2.3 (1.5)	1.7	< .001	NS (.11)
Function	29.7 (13.3)	2.6	< .001	15.5 (10.0)	1.4	< .001	< .001
SF-12 PCS	17.3 (10.6)	2.1	< .001	8.2 (7.8)	0.9	=.002	=.002
SF-12 MCS	8.8 (11.4)	0.7	< .001	6.4 (8.8)	0.6	=.015	NS (.59)
SF-36							
Bodily pain	-48.1 (23.5)	2.9	< .001	-22.5 (18.9)	1.4	= .001	< .001
Physical function	-42.5 (22.6)	2.5	< .001	-18.8 (18.8)	1.1	= .003	< .001
Role physical	-49.5 (48.7)	1.7	< .001	-17.2 (32.6)	0.8	= .048	= .012
Role emotional	-19.1 (41.6)	0.4	< .001	-10.4 (51.2)	0.3	NS (.47)	NS (.64)
Mental health	-5.8 (16.0)	0.3	< .001	-8.5 (18.0)	0.5	NS (.08)	NS (.68)
Vitality	-15.0 (20.1)	0.6	< .001	-7.3 (13.3)	0.4	= .024	NS (.16)
Social function	-26.2 (30.0)	0.8	< .001	-13.3 (27.2)	0.4	NS (.1)	NS (.12)
General health	-3.8 (19.9)	0.2	NS (.06)	-0.3 (0.4)	0.0	NS (.93)	NS (.42)
EQ-5D$_{VAS}$	-0.18 (0.23)	0.7	< .001	-0.05 (0.23)	0.2	NS (.30)	= .043
EQ-5D$_{index}$	-0.43 (0.29)	1.3	< .001	-0.30 (0.37)	1.0	= .008	NS (.27)

Table 4
Spearman inter-correlations between preoperative questionnaires

	OHS	WOMAC Pain	Womac Stiffness	WOMAC Function	SF-12 PCS	SF-12 MCS	EQ-5D$_{VAS}$	EQ-5D$_{index}$
WOMAC								
Pain	0.76**							
Stiffness	0.63**	0.64**						
Function	0.88**	0.74**	0.67**					
SF-12 PCS	-0.53**	-0.46**	-0.26**	-0.45**				
SF-12 MCS	-0.49**	-0.37**	-0.35**	-0.54**	0.31**			
SF-36								
Bodily pain	-0.73**	-0.70**	-0.53**	-0.71**	0.67**	0.48**	0.43**	0.68**
Physical function	-0.62**	-0.53**	-0.42**	-0.67**	0.66**	0.60**	0.32**	0.55**
Role physical	-0.39**	-0.34**	-0.20*	-0.40**	0.53**	0.38**	0.18	0.41**
Role emotional	-0.29**	-0.20*	-0.21*	-0.33**	-0.01	0.79**	0.15	0.42**
Mental health	-0.31**	-0.20*	-0.25**	-0.33**	0.06	0.82**	0.28**	0.48**
Vitality	-0.39**	-0.34**	-0.32**	-0.40**	0.27**	0.69**	0.53**	0.51**
Social function	-0.57**	-0.45**	-0.30**	-0.57**	0.48**	0.68**	0.30**	0.55**
General health	-0.23*	-0.17	-0.16	-0.32**	0.31**	0.38**	0.44**	0.38**
EQ-5D$_{VAS}$	-0.37**	-0.39**	-0.32**	-0.38**	0.38**	0.36**		
EQ-5D$_{index}$	-0.64**	-0.57**	-0.42**	-0.62**	0.42**	0.56**	0.38**	

*p < 0.05 (two-tailed)
**p < 0.01 (two-tailed)

Table 5
Spearman correlations of change (preoperatively-1 year postoperatively)

	OHS	WOMAC Pain	Womac Stiffness	WOMAC Function	SF-12 PCS	SF-12 MCS	EQ-5D$_{VAS}$	EQ-5D$_{index}$
WOMAC								
Pain	0.68**							
Stiffness	0.48**	0.64**						
Function	0.66**	0.75**	0.61**					
SF-12 PCS	0.38**	0.44**	0.30**	0.57**				
SF-12 MCS	0.50**	0.33**	0.33**	0.51**	0.29**			
SF-36								
Bodily pain	-0.48**	-0.61**	-0.43**	-0.66**	0.62**	0.35**	0.38**	0.37**
Physical function	-0.41**	-0.53**	-0.39**	-0.72**	0.68**	0.42**	0.42**	0.41**
Role physical	-0.33**	-0.41**	-0.36**	-0.51**	0.75**	0.35**	0.35**	0.27**
Role emotional	-0.26**	-0.15	-0.20*	-0.25**	0.04	0.71**	0.27**	0.31**
Mental health	-0.23*	-0.10	-0.16	-0.22*	0.05	0.72**	0.28**	0.36**
Vitality	-0.45**	-0.41**	-0.44**	-0.50**	0.30**	0.64**	0.45**	0.50**
Social function	-0.41**	-0.39**	-0.21*	-0.49**	0.35**	0.57**	0.24*	0.39**
General health	-0.08	-0.06	-0.08	-0.26**	0.22*	0.29**	0.40**	0.22*
EQ-5D$_{VAS}$	-0.43**	-0.36**	-0.28**	-0.45**	0.37**	0.43**		
EQ-5D$_{index}$	-0.51**	-0.47**	-0.34**	-0.49**	0.31**	0.46**	0.45**	

*p < 0.05 (two-tailed)
**p < 0.01 (two-tailed)

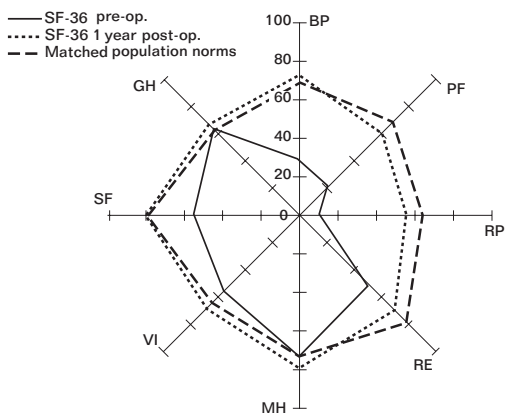


Figure 1

Radar graph of the scores on the SF-36 health domains at the preoperative measurement, 1 year after the operation and scores according to age and gender matched population norms.

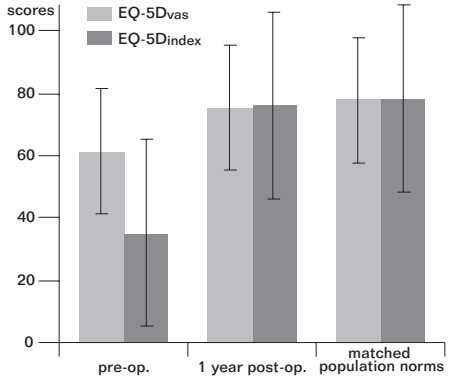


Figure 2

$EQ\text{-}5D_{VAS}$ and $EQ\text{-}5D_{index}$ score at the preoperative measurement, 1 year after the operation and scores according to age- and gender matched population norms.

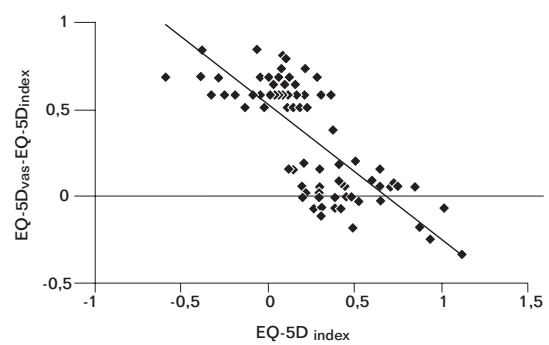


Figure 3

Graph of the difference of $EQ\text{-}5D_{VAS}$ - $EQ\text{-}5D_{index}$ vs $EQ\text{-}5D_{index}$ score. There is substantial disagreement between the score at $EQ\text{-}5D_{index}$ levels less than 0.5, the diagonal line is a linear trendline through the data points.

Discussion

For some years it has been agreed that patient reported outcome data should define clinical success and expertise in the eyes of payers, patients and even fellow providers.[3] When limited resources must be distributed to different healthcare interventions, it is important to have information on the relative gain in health status achieved. We compared two disease-specific and three generic questionnaires in the evaluation of THA with respect to their floor and ceiling effects, content validity and sensitivity to change. We found that patients who did not return their questionnaires at the one-year follow-up were more likely to have had a post-operative complication and had significantly lower scores on the WOMAC pain and the SF physical functioning scores at three months after operation. The fact that these patients experienced a complication after surgery may be an explanation for their not returning the questionnaires. Patients with a post-operative complication who were included in our analyzed group had similar outcome scores at follow-up at one year as the patients without complications. We therefore do not expect the non-responding group to have a pronounced effect on the results at one year after operation.

All outcome scores had improved significantly at one year after operation except for the SF-36 general health domain. In our study, the OHS, together with pain and physical function domains of the WOMAC and the SF-36, showed the largest effect sizes after THA compared with other dimensions of health status. These were thus identified as being of particular importance in the assessment of outcome after THA.[11,22] We found greater effect sizes for patients with unilateral hip involvement (Charnley A) compared

with patients who had other conditions which directly impeded mobility (Charnley B and C). In this last group, the effect size of the OHS more than doubled the effect sizes of WOMAC pain and physical function. This emphasizes the site-specific properties of the OHS. However, the effect size of the OHS in the Charnley B and C group is significantly smaller than that in the Charnley A group. This suggests that there is still an influence of the other conditions which impede mobility upon the OHS,[23] which is in contrast with the findings of Dawson et al.[22] who did not find a significant difference in the OHS in a group with unilateral hip symptoms and a group with diffuse symptoms.
Our main concern was to identify the optimal set of health-outcome measures in THA. In a standard evaluation study of THA the completion of four questionnaires is a heavy burden for patients, especially when the OHS and the WOMAC measure similar symptoms as shown by their high correlations and correlations of change. We prefer the OHS since it is a shorter, more site-specific and responsive questionnaire. It also shows fewer ceiling effects after operation and seems to be the most appropriate disease-specific questionnaire of the two tested in this study. In a study by Dunbar et al.[18] on patients who underwent total knee replacement, the Oxford knee score was also preferred, as opposed to the WOMAC and Lequesne scores, because of its higher feasibility, validity and reliability. The SF-36 captures additional, important quality-of-life domains that are influenced by a THA.[24] The SF-36 could probably be replaced by the SF-12 because of the large changes that occur between the domains which belong to the physical component summary (PCS) and the mental component summary (MCS) scores. However, this is at the cost of losing detailed information about separate health-status domains. For large, cross-sectional surveys, the SF-12 is more appropriate because of its high feasibility.[18] It should be noted that we calculated the SF-12 scores from the SF-36 in order to reduce the number of questionnaires which patients had to complete. We did not expect the scores obtained in this way to be different from those we would have found if the patients had completed the SF-12 questionnaire itself.
We found considerable floor and ceiling effects in the questionnaires, especially for the physical and emotional role domains of the SF-36. This can be explained by the limited answering options (yes or no) for questions about limitations in role functioning. In the new version of the SF-36 (SF-36v2) these scales are replaced by five-point Likert scales, which improve precision and range for both domains.[25] The many ceiling effects post-operatively reflect the excellent outcome which can be achieved with THA in patients with severe hip disease. This is also reflected in the scores of the SF-36 and the EQ-5D, which approach general population norms at one year after operation. The OHS, WOMAC and SF-36 have shown good psychometric qualities in several studies in patients undergoing THA.[10,26] There are only a few published studies on the outcome of THA in which the EQ-5D was used.[5] The developers of the EQ-5D recommend it as a complementary instrument, not as a substitute for other instruments.[15] Hurst et al.,[27] in a study of rheumatic patients, showed that the EQ-5D is simple to use, is valid, responsive to change and reliable. The EQ-5D, especially

the EQ-$5D_{index}$, showed good pre-operative correlations and correlations of change with the other questionnaires. It also showed a similar return to age- and gender-matched population norms as the SF-36.
However, our observations of the EQ-5D scores highlighted similar problems as those found by Wolfe and Hawley,[28] who studied rheumatic patients, and Fransen and Edmonds[29] who studied patients with osteoarthritis of the knee. The binominal distribution of the EQ-$5D_{index}$ score can be explained by the crude scaling of the EQ-5D. For example, in the mobility domain, the two extreme categories are 'no problems' and 'unable/confined to bed'. Since almost no patients will be in the 'unable' category, the only movement which can be detected will be between 'no problems' and the presence of 'problems'.[28] The pre-operative difference and the post-operative agreement between the EQ-$5D_{VAS}$ (patient centred value) and the EQ-$5D_{index}$(social value) was remarkable (Table 2).[27-29] This pre-operative difference can be explained by what has been called 'response shift'. Response shift is the change in internal standards, in values, or in the perception of quality of life, which are created by changes in health state.[30] In clinical practice, this means that most patients perceive their quality of life as being better than from the perspective of the general population. This is either because patients become accustomed to their disease or because their expectations about their health status have changed. Another explanation for the difference in pre-operative EQ-5D values is that patients with severe hip disease interpret their hip condition as a mechanical problem which has to be fixed, not as part of their general health. This is also reflected in the general health domain of the SF-36, the only domain of the SF-36 which did not improve significantly at one year after THA, and which is relatively high when compared with other domains.[31] Several patients had a pre-operative EQ-$5D_{index}$ less than zero. That is, from a social perspective their health state was worse than death. This represents the fact that normal individuals asked to consider such an existence would regard themselves as better off dead.[27] The self-assessed health of patients on the EQ-$5D_{VAS}$ scale diverges from the social view in more severely disabled patients, which raises important ethical and practical questions about the use and interpretation of utility values, for example in cost-utility analyses. The peculiarities in the EQ-5D scores may justify further research into the EQ-5D, especially in the scaling of the EQ-$5D_{index}$ and the valuation of severe health states in relation to death. Further investigation may also be helpful when smaller changes in health status are investigated, for instance in the follow-up of THA.
Based on the results of our study, we recommend the OHS and SF-12 in the assessment of THA. The SF-36 may be used in smaller studies in which a more detailed description of health domains is needed and when smaller changes in health status are being investigated. The EQ-5D is useful in situations in which utility values are needed to calculate cost-effectiveness or quality-adjusted life years (QALYs), such as in the assessment of new techniques in THA.

References

1. Rissanen P, Aro S, Sintonen H, Slatis P, Paavolainen P. Quality of life and functional ability in hip and knee replacements: a prospective study. Qual Life Res 1996; 5(1):56-64.
2. Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, Leslie K, Bullas R. The effect of elective total hip replacement on health-related quality of life. J Bone Joint Surg Am 1993; 75(11):1619-26.
3. Ritter MA, Albohm MJ. Overview: maintaining outcomes for total hip arthroplasty. The past, present, and future. Clin Orthop 1997(344):81-7.
4. Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement surgery. Quality in health care 1996; 5:81-88.
5. Dawson J, Fitzpatrick R, Frost S, Gundle R, McLardy-Smith P, Murray D. Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. J Bone Joint Surg Br 2001; 83(8):1125-9.
6. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998; 51(11):1055-68.
7. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35(11):1095-108.
8. Kosinski M. Scoring the SF-12 Physical and Mental Summary Measures. Medical Outcomes Trust Bulletin 1997; 5(5):3-4.
9. Greenfield S, Apolone G, McNeil BJ, Cleary PD. The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. Med Care 1993; 31(2):141-54.
10. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78(2):185-90.
11. Fitzpatrick R, Morris R, Hajat S, Reeves B, Murray DW, Hannen D, Rigge M, Williams O, Gregg P. The value of short and simple measures to assess outcomes for patients of total hip replacement surgery. Qual Health Care 2000; 9(3):146-150.
12. Hoefnagels NH, Dhert WJ, Gosens T, Bulstra SK, de Vet HCW, van Langelaan EJ, Okhuysen-Vonk S, Geesink RGT. De Oxford Heup Score, een Nederlandstalige patiëntenvragenlijst ter evaluatie van het resultaat van een totale heupoperatie gebaseerd op de vertaling van "the 12-item questionnaire on the perceptions of patients about THR". Tijdschrift voor Orthopaedie 2002; 9:5-16.

13. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. Journal of Orthopaedic Rheumatology 1988; 1:95-108.
14. Ware JE, Jr., Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. Med Care 1995; 33(4 Suppl):AS264-79.
15. Brooks R. EuroQol: the current state of play. Health Policy 1996; 37(1):53-72.
16. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, Bullinger M, Kaasa S, Leplege A, Prieto L, Sullivan M. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998; 51(11):1171-8.
17. How to score the SF-36 Health Survey. Medical Outcomes Trust 1994.
18. Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty. Results of a survey of 3600 patients from The Swedish Knee Arthroplasty Registry. J Bone Joint Surg Br 2001; 83(3):339-44.
19. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27(3 Suppl):S178-89.
20. Cohen J. Statistical Power Analyses for the Behavioral Sciences. Vol. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
21. Charnley J. Low Friction Arthroplasty of the Hip. Berlin Heidelberg New York: Springer-Verlag, 1979.
22. Dawson J, Fitzpatrick R, Murray D, Carr A. The problem of 'noise' in monitoring patient-based outcomes: generic, disease-specific and site-specific instruments for total hip replacement. J Health Serv Res Policy 1996; 1(4):224-31.
23. McMurray R, Heaton J, Sloper P, Nettleton S. Measurement of patient perceptions of pain and disability in relation to total hip replacement: the place of the Oxford hip score in mixed methods. Qual Health Care 1999; 8(4):228-33.
24. Lieberman JR, Dorey F, Shekelle P, Schumacher L, Kilgus DJ, Thomas BJ, Finerman GA. Outcome after total hip arthroplasty. Comparison of a traditional disease-specific and a quality-of-life measurement of outcome. J Arthroplasty 1997; 12(6):639-45.
25. Ware JE. Comparisons among SF-36 Surveys. Qualimetric Inc., 2003.
26. Boardman DL, Dorey F, Thomas BJ, Lieberman JR. The accuracy of assessing total hip arthroplasty outcomes: a prospective correlation study of walking ability and 2 validated measurement devices. J Arthroplasty 2000; 15(2):200-4.
27. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997; 36(5):551-9.
28. Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. Br J Rheumatol 1997; 36(7):786-93.

29. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology (Oxford) 1999; 38(9):807-13.
30. Sprangers MA, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. Soc Sci Med 1999; 48(11):1507-15.
31. McGuigan FX, Hozack WJ, Moriarty L, Eng K, Rothman RH. Predicting quality-of-life outcomes following total joint arthroplasty. Limitations of the SF-36 Health Status Questionnaire. J Arthroplasty 1995; 10(6):742-7.

Chapter 8

Summary and conclusions

In this chapter, the results of the studies that were performed for this thesis are summarized by addressing the four questions that were formulated in Chapter 1 (Introduction and aims).

Question 1: What is the demographic profile of patients receiving THA in The Netherlands and Sweden and what is the expected demand for THA caused by the aging population in both countries?

From earlier work by Okhuysen et al. it was demonstrated that the incidence of THA had increased rapidly in the last two decades in the Netherlands. In **Chapter 2**, we reported an increasing incidence of THA in both The Netherlands and Sweden, which could only for a small part be explained by changes in size and age-profile of the population. Although there are comparable population characteristics and health care systems in The Netherlands and Sweden, we found a 20% higher incidence of THA in the Netherlands, as compared to Sweden, even after correction for age- and population structure. This might be partly explained by relatively lower budgets for THA in Sweden, or by differences in indications between these countries. The most important result was that the increase in number of THAs until 2020, based on aging of the population alone, was expected to be 28% in Sweden and 44% in the Netherlands. It is concluded from this study that a large increase in THA rate will be necessary in order to satisfy the future needs for THA, both in The Netherlands as well as in Sweden, which is important for both policy makers and orthopaedic surgeons. Our predictions on future THA rates in The Netherlands and Sweden should even be considered as conservative, because the demand for THA in Sweden and The Netherlands has not been satisfied yet, as can be concluded from the still existing waiting times for the procedure. Further, the indications for THR will probably still widen if the outcome continues to improve. Lastly, as patients become more emancipated, and more aware of the existing treatment methods, they will be less inclined to accept their disability and prefer treatment in an earlier stage of their disease. The increasing demands for medical treatments caused by the aging of the population in developed countries, in the presence of limited health care budgets and a smaller employed population, will result in new deficiencies in the provision of medical procedures like THA.

Question 2: To what extent are indication guidelines for THA applied in clinical practice, and what are the differences in indications for THA between The Netherlands and Sweden?

In **Chapter 3**, we report on a survey of Dutch and Swedish orthopaedic surgeons regarding their indications for THA. We found good compliance in both countries to existing guidelines for indication for THA with respect to pain and walking ability, but moderate compliance with respect to function, findings at physical examination and radiographic signs of joint damage. Although most surgeons agreed on several

indications for THA, there was no clear consensus among surgeons regarding the indications for THA in either country. We found no systematic differences in indication ratings of pain, function and physical examination between Dutch and Swedish surgeons. Swedish surgeons had a stricter interpretation of modifying factors in the decision for THA, which might partly explain the lower incidence of THA in Sweden as compared to The Netherlands that is described in **Chapter 2**. Uniformity of indications is important to create similar chances of access to a procedure for all patients. The results of this survey are important, especially in this era when appropriateness of surgery and geographic variations in rates of procedures are under the scrutiny of administrators, health insurance companies and a more informed patient population.

Question 3: What are the demographic and bacteriological profiles and surgical practice in revision THA surgery for deep infection, and what is the risk for repeated revision for both aseptic loosening and recurrent infection?

After investigating 960 first revisions for deep infection from the Swedish National Hip Register in **Chapter 4**, some interesting conclusions could be drawn. Re-revision rates were moderate with - at 10 years follow-up - 10% re-revision for infection, 11% re-revision for aseptic loosening and 21% re-revision for any cause. We found an increased risk for re-revision for infection in patients who had wound complications after primary THA, in patients revised in a rural clinic and in patients who underwent a one-stage revision. Patients undergoing two-stage revision with an implantation interval over 6 months had lower risk for re-revision. Further, another finding was an increased re-revision risk for aseptic loosening in patients under 65 years, in patients who underwent a previous minor procedure after primary THA, and in patients who received bone-grafts at re-implantation of the prosthesis.
Our results also suggest increasing clinical problems with low-virulent peri-prosthetic infections and a development of increasing bacterial resistance in infected THA. The issue of increasing antibiotic resistance in THA infections needs meticulous attention because it severely threatens the success of the procedure and causes severe suffering in affected patients. Therefore, also the possible causal relationship between the use of systemic prophylaxis/antibiotic containing bone cement and the development of resistant strains deserves further investigation.
Further, this study shows the excellent ability of a national register to assess outcome of relatively rare complications of THA by providing large numbers of patients and long follow-up in a cost-effective way.

Question 4: Does waiting for THA affect pre-operative health status and post-operative outcome of the operation?

In **Chapter 5 and 6**, we assessed several aspects of orthopaedic in-patient waiting lists and the effect of waiting for THA on pre-operative health status and outcome. In the study described in Chapter 5, we found a limited validity of the waiting list for THA in our university hospital (22% of the patients on the list never underwent the planned procedure in our hospital) and inaccurate prediction of actual waiting times for the procedure. Solving the waiting list problem proved to be a multi-factorial issue, which could not be solved by incidental extra funding only. The limited validity of waiting lists and waiting time estimates we found is alarming, especially because such information is used in policy making and as performance indicator for hospitals. However, without further analysis, these parameters have little significance.
Confusion over terminology, differences in measurement approaches and a general lack of awareness of the relative effectiveness of different approaches to managing waiting lists and waiting times, all hamper real progress in this area.
In **Chapter 6**, we demonstrated that health scores deteriorated significantly during an average waiting time of 6 months for THA. During this waiting period, a considerable loss in quality-adjusted life years occurred, simply by postponing surgery. If functionally impaired patients can achieve consistently excellent relief of symptoms, disproportionately long waiting times impose an avertable burden of disability.
Although we did not find any direct effect of waiting times on postoperative outcome of THA, patients in a later phase of the disease process did not improve to the level achieved by patients with a better preoperative function.
In a subsequent study regarding the outcome of THA in relation to waiting times, we assessed the performance of several disease-specific and general patient oriented outcome questionnaires (**Chapter 7**). We found high correlations and correlations of change in scores between the OHS, WOMAC, physical domains of the SF-12 and SF-36, and the EQ-$5D_{index}$. Patients with unilateral hip disease showed more change in OHS, WOMAC pain and function, the physical domains of the SF-36 and the EQ-$5D_{VAS}$, as compared to patients with bilateral hip disease or other conditions directly impeding mobility. In the last group, the effect size for the OHS more than doubled the effect sizes of WOMAC pain and physical function. Furthermore, we found binomial distribution of the EQ-$5D_{index}$ score and a pre-operative discrepancy and post-operative agreement between the EQ-$5D_{VAS}$ and the EQ-$5D_{index}$. These peculiarities may justify further research into the EQ-5D, especially in the scaling of the EQ-$5D_{index}$ and the valuation of severe health states. We recommend using the OHS and SF-12 in the evaluation of THA. The EQ-5D is useful in situations where utility values are needed to calculate cost-effectiveness or quality adjusted life years (QALYs), such as in the evaluation of new techniques in THA.

Conclusions and recommendations

From the studies described in this thesis, investigating different aspects of outcome assessment in total hip arthroplasty, the following conclusions and recommendations can be made:

To ensure good quality of life and adequate mobility for future elderly generations, adequate measures need to be taken to respond adequately to the increasing demand for THA, for instance with regard to future budgets for THA, the use of cost-effective prosthesis systems and training of surgeons and other medical personnel.
The revision burden of THA in The Netherlands is relatively high compared to the Scandinavian countries. Apart from figures from a discharge registry and an implant registry with compliance less than 60% and no possibilities for follow-up, there is no registration system for THA in The Netherlands. In spite of problems with liability, compliance, cost and confidentiality with the introduction of a register, new efforts should be undertaken within the Dutch orthopaedic community to start nation-wide orthopaedic registries. Such registries should serve both as a database to study outcome after surgery and as a benchmarking tool for all participating clinics.
The moderate to good compliance to existing recommendations for the indication for THA found in our survey of Dutch and Swedish orthopaedic surgeons shows the applicability of these recommendations in clinical practice. However, we found that patients in a later phase of the disease process did not improve to the level achieved by patients with a better preoperative function. Therefore, traditional orthopaedic practice, to delay surgery as long as possible, might need to be re-evaluated, especially for elderly patients. Clinical guidelines on THA, based on review of clinical evidence and involving cost-effectiveness and patient preference, should be established in both The Netherlands and Sweden. Apart from consensus statements, no such guidelines exist in either country.
Because of the multiple risk factors found for re-revision for both septic and aseptic loosening, and the relatively low annual number of revisions for infection, the findings of our study on revisions for infection from the Swedish National Hip Registry might be a reason to centralize these difficult treatment procedures in specialized clinics. This would also improve adequate monitoring of the type and resistance patterns of bacteria found in these infections. Development of resistant bacteria is a worrisome trend in medicine in general and especially in the vulnerable group of patients with an artificial joint. Future research in prevention and treatment of this devastating complication is needed to maintain the high percentages of patients with excellent outcome after THA.
Long waits for THA cause considerable loss in quality-adjusted life years for patients involved and thus impose an avertable burden of disability. With regard to the waiting list problem that still exists in orthopaedic surgery in The Netherlands and Sweden, especially for joint replacements, new approaches for managing waiting lists

should be followed. These include reducing demand for the procedure by audits of waiting lists and reassessment of patients on the lists. Further, prioritization of patients on the waiting list based on clinical urgency is a strategy used in other countries. Lastly, reorganization of care patterns such as redirection of referrals to clinicians with shorter waiting lists, reduction of missed appointments and reduction of specialist physician follow-up visits are possibilities to improve access to a medical procedure like THA.

In this thesis, several aspects of outcome assessment in total hip arthroplasty were investigated. We should strive for continuous quality improvement of the procedure, by following and adapting practice guidelines based on patient oriented outcome measures and population studies, to use available health care resources in an optimal way and to guarantee best quality care to our patients.

Samenvatting en conclusies

In dit hoofdstuk worden de resultaten van de eerder beschreven hoofdstukken samengevat door de vragen welke werden gesteld in Hoofdstuk 1 (Introductie en doelstellingen) van dit proefschrift, te beantwoorden.

Vraag 1: Wat is het demografische profiel van de groep patiënten die een totale heupprothese (THP) ontvangt in Nederland en Zweden en wat is de verwachte vraag naar totale heupprotheses door de vergrijzing van de bevolking in beide landen?

Eerder werk van Okhuysen et al. heeft aangetoond dat de incidentie van totale heupprothese operaties snel gestegen is de laatste twee decennia. In **Hoofdstuk 2** werd zowel in Nederland als in Zweden een toenemende incidentie van totale heupprothese operaties aangetoond, die slechts gedeeltelijk kon worden verklaard door veranderingen in grootte en leeftijdsstructuur van de bevolking. Hoewel de bevolkingskenmerken en gezondheidszorgsysteem vergelijkbaar zijn in Nederland en Zweden, vonden we dat er in Nederland relatief 20% meer totale heupprothese operaties plaatsvonden ten opzichte van Zweden, zelfs na correctie voor leeftijds- en bevolkingsstructuur. Dit zou verklaard kunnen worden door de relatief lagere budgetten voor THP operaties in Zweden, of door verschillen in indicatiestelling tussen beide landen. Het belangrijkste resultaat was, dat door de vergrijzing van de bevolking het aantal protheses per jaar tot en met 2020 in Zweden met 28% stijgt en met 44% in Nederland. Uit deze studie wordt geconcludeerd dat zowel in Nederland als in Zweden een grote toename in aantallen operaties nodig is om aan de vraag naar THP operaties te kunnen voldoen, hetgeen belangrijke informatie is voor zowel beleidsmakers als orthopaedisch chirurgen. Onze voorspellingen voor toekomstige aantallen benodigde operaties in beide landen moeten als conservatief worden beschouwd omdat nog niet aan de vraag naar totale heupprothese operaties voldaan is, zoals blijkt uit de nog steeds bestaande wachtlijsten voor de ingreep. Daarnaast zullen de indicaties voor de operatie waarschijnlijk nog verbreden als de resultaten van de operatie verder verbeteren. Omdat patiënten mondiger worden en zich meer bewust zijn van de bestaande behandelingsmogelijkheden, zullen ze minder geneigd zijn om hun beperkingen te accepteren en de voorkeur geven aan behandeling in een vroeger stadium van het ziekteproces. Mede door beperkte budgetten en een kleinere beroepsbevolking zal de toenemende vraag naar medische behandelingen, veroorzaakt door vergrijzing van de bevolking in ontwikkelde landen, nieuwe tekorten veroorzaken in de verstrekking van medische ingrepen zoals totale heupvervanging.

Vraag 2: In hoeverre worden indicatierichtlijnen voor totale heupprothese operaties toegepast in de klinische praktijk, en wat zijn de verschillen in indicatiestelling voor totale heupprothese operaties tussen Nederland en Zweden?

In **Hoofdstuk 3** wordt een onderzoek naar de indicatiestelling voor THP operaties onder Nederlandse en Zweedse orthopaedisch chirurgen besproken. Het bleek dat

zowel Zweedse als Nederlandse orthopaeden zich goed houden aan de consensus richtlijnen zoals die in beide landen bestaan, met name met betrekking tot indicaties als pijn en verminderd loopvermogen en in mindere mate met betrekking tot functie, bevindingen bij lichamelijk onderzoek en radiologisch bewijs van gewrichtsschade. Hoewel de meeste orthopaeden het eens waren over een meerderheid van de indicaties was er in geen van beide landen een duidelijke consensus tussen chirurgen met betrekking tot de indicatiestelling voor de plaatsing van een totale heupprothese. Er waren geen systematische verschillen tussen indicatiestelling met betrekking tot pijn, functie of bevindingen bij lichamelijk onderzoek tussen Zweedse en Nederlandse chirurgen. Zweedse chirurgen waren strikter in hun interpretatie van bepalende factoren (zoals leeftijd en co-morbiditeit) in de beslissing voor operatie, hetgeen gedeeltelijk de lagere incidentie van totale heupprothese operaties in Zweden zou kunnen verklaren zoals beschreven in **Hoofdstuk 2.** Uniformiteit in indicatiestelling is belangrijk om voor alle patiënten een vergelijkbare toegankelijkheid tot operatie te creëren. De resultaten van dit onderzoek zijn belangrijk, vooral in deze tijd waarin een al dan niet terechte indicatiestelling en variaties in aantallen operaties kritisch bekeken worden door overheden, verzekeraars en een goed geïnformeerde patiëntenpopulatie.

Vraag 3: Wat zijn demografische en bacteriologische kenmerken bij revisiechirurgie van geïnfecteerde totale heupprotheses en wat zijn de risicofactoren voor een nieuwe revisie voor infectie en aseptische loslating van de prothese?

Na onderzoek van 960 primaire revisies van geïnfecteerde totale heupprotheses uit het Zweedse Hip Registry in **Hoofdstuk 4** konden enkele interessante conclusies worden getrokken. Re-revisie percentages waren redelijk met –bij 10 jaar follow-up- 10% re-revisies voor infectie, 10% re-revisies voor aseptische loslating en 21% re-revisies wanneer alle oorzaken voor re-revisie in beschouwing worden genomen. Er werd een verhoogd re-revisie risico voor hernieuwde infectie gevonden bij patiënten die wondproblemen hadden na hun primaire THP operatie, bij patiënten die hun revisie operatie in een kleiner ziekenhuis hadden ondergaan en bij patiënten die een one-stage revisie hadden ondergaan. Ook werd er een verhoogd re-revisie risico voor aseptische loslating gevonden bij patiënten jonger dan 65 jaar, bij patiënten die eerder een kleine ingreep ondergingen in hun heupgewricht na de primaire THP operatie, en bij patiënten die een bottransplantaat ontvingen bij re-implantatie van de prothese. Onze resultaten suggereren ook toenemende problemen met prothese infecties door bacteriën met een lage virulentie en een ontwikkeling van toenemende resistentie van bacteriën bij geïnfecteerde totale heupprotheses. Mede daarom verdient de mogelijke relatie tussen het gebruik van profylactische systemische antibiotica en antibioticumhoudend botcement en de ontwikkeling van resistente bacteriën nader onderzoek. Verder laat deze studie de uitstekende mogelijkheden van een nationaal registratiesysteem zien om, door de grote aantallen patiënten en lange follow-up, op een kosten-effectieve manier de resultaten van relatief zeldzame complicaties bij totale heupprothese operaties te evalueren.

Vraag 4: Beïnvloedt het wachten op een totale heupprothese operatie de pre-operatieve gezondheidstoestand dan wel de uitkomst van de operatie?
In Hoofdstuk 5 en 6 werden verschillende aspecten van orthopedische wachtlijsten geëvalueerd, evenals het effect van de wachttijd voor een THP operatie op de pre-operatieve gezondheidstoestand en het resultaat van de operatie. In Hoofdstuk 5 werd een beperkte validiteit van de wachtlijst gevonden in ons academisch ziekenhuis (22% van de patiënten op de wachtlijst werd nooit in ons ziekenhuis geopereerd) evenals onjuiste voorspellingen van de eigenlijke wachttijd voor de ingreep. Het oplossen van de wachtlijstproblematiek bleek een multi-factorieel probleem wat niet alleen door extra tijdelijke financiële middelen kon worden opgelost. De beperkte validiteit van wachtlijsten en wachttijden is alarmerend, omdat deze informatie gebruikt wordt door beleidsmakers en als performance indicator voor ziekenhuizen. Zonder verdere analyse hebben deze parameters echter weinig betekenis. Verwarring over terminologie, verschillen in meetmethodes en verschillende manieren om wachttijden en wachtlijsten aan te pakken belemmeren echte vooruitgang op dit terrein.
In **Hoofdstuk 6** werd aangetoond dat verschillende gezondheidsscores significant verslechterden tijdens een wachttijd van gemiddeld 6 maanden voor een totale heupprothese operatie. Door het uitstellen van de operatie trad er een behoorlijk verlies in kwaliteit van leven (quality-adjusted life years (QALYs)) op. Als functioneel beperkte patiënten over het algemeen een uitstekend resultaat bereiken na de operatie, betekent dat dat onevenredig lange wachttijden een vermijdbare bron van disfunctioneren zijn. Hoewel er geen direct effect gevonden werd van de wachttijd op de gezondheidstoestand ná de operatie, verbeterden patiënten in een latere fase van het ziekteproces (dus met slechtere functiescores voor de operatie) niet tot het niveau wat patiënten met een betere functie voor de operatie bereikten.
In een vervolgstudie naar de uitkomst van THP operaties werden de prestaties van verschillende ziekte-specifieke en algemene, patiëntgeoriënteerde vragenlijsten onderzocht (**Hoofdstuk 7**). Er werden hoge correlaties gevonden tussen de Oxford Hip Score (OHS), WOMAC, fysieke domeinen van de SF-12 en SF-36 en de EQ-$5D_{index}$ en hoge correlaties tussen de pre- en postoperatieve verschillen van deze vragenlijsten. Patiënten met éénzijdige heupklachten lieten meer verandering zien in OHS, WOMAC pijn en functie scores, de fysieke domeinen van de SF-12 en SF-36 en de EQ-$5D_{VAS}$ vergeleken met de patiënten met heupklachten beiderzijds of andere klachten die de mobiliteit beïnvloeden. In de laatste groep was de effect size (maat voor verandering in score) voor de OHS meer dan dubbel zo groot als de effect sizes van de WOMAC pijn en functie scores. Daarnaast werd er een binomiale verdeling gevonden van de EQ-$5D_{index}$ score en een pre-operatieve discrepantie tussen de EQ-$5D_{VAS}$ en de EQ-$5D_{index}$. Deze bevindingen rechtvaardigen verder onderzoek naar de EQ-5D, vooral met betrekking tot de schaalverdeling en de waardering van slechte gezondheidsscores. Het gebruik van de OHS en de SF-12 wordt aanbevolen in de evaluatie van totale heupprothese operaties. De EQ-5D is nuttig in situaties waar utility values nodig zijn om kosten-effectiviteit of QALYs uit te rekenen, bijvoorbeeld bij de evaluatie van nieuwe technieken in totale heupprothese operaties.

Conclusies en aanbevelingen

Uit de resultaten zoals beschreven in dit proefschrift kan het volgende worden geconcludeerd:
Om een goede kwaliteit van leven en een goede mobiliteit te kunnen bieden aan toekomstige generaties ouderen, moeten er adequate maatregelen genomen worden om op een juiste manier te kunnen reageren op de toenemende vraag naar totale heupprothese operaties. Deze maatregelen zouden kunnen bestaan uit realiseren van voldoende budgetten voor de ingreep, het gebruik van kosten-effectieve prothesesystemen en opleiding van voldoende chirurgen en ander medisch personeel.
Het aantal revisies van totale heupprotheses is relatief wat hoger in Nederland dan in Zweden. Behalve getallen uit de landelijke medische registratie en een implantaten registratiesysteem met een compliance van minder dan 60% zonder mogelijkheden voor follow-up, is er geen registratiesysteem voor totale heupprotheses in Nederland. Ondanks problemen met wettelijke verantwoordelijkheid, compliance, kosten en privacy zouden er nieuwe pogingen moeten worden ondernomen binnen de Nederlandse orthopaedie om tot landelijke orthopaedische registratiesystemen te komen. Zulke registers zouden zowel als database kunnen dienen om het resultaat van de operatie te onderzoeken als wel als graadmeter (benchmarking tool) voor deelnemende klinieken kunnen dienen.
Het feit dat zowel Zweedse als Nederlandse orthopaedisch chirurgen zich redelijk tot goed hielden aan bestaande consensus richtlijnen voor de indicatiestelling bij totale heupprothese operaties, laat de toepasbaarheid van deze richtlijnen in de klinische praktijk zien. We vonden echter dat patiënten in een latere fase van het ziekteproces (dus met slechtere functiescores voor de operatie) niet tot het niveau verbeterden wat patiënten met een betere functie voor de operatie bereikten. Daarom zou de traditionele orthopedische praktijk, namelijk de operatie zolang mogelijk uitstellen, wellicht opnieuw geëvalueerd moeten worden, vooral voor oudere patiënten. Verder zouden zowel in Zweden als in Nederland richtlijnen moeten worden opgesteld die gebaseerd zijn op klinisch bewijs, voorkeuren van patiënten en kosten-effectiviteit. Dergelijke richtlijnen bestaan noch in Zweden, noch in Nederland, behalve de reeds bestaande consensus verklaringen.
Vanwege de meerdere risicofactoren voor hernieuwde revisies voor zowel infectie als aseptische loslating van de prothese, als de relatief lage aantallen revisies die per jaar voor infectie van de prothese worden uitgevoerd, zouden de bevindingen uit onze studie van geïnfecteerde revisies uit het Zweedse National Hip Registry een aanbeveling kunnen zijn om deze moeilijke operaties in gespecialiseerde klinieken uit te voeren. Dit zou ook het vervolgen van typen bacteriën en zich ontwikkelende resistentiepatronen kunnen verbeteren. De ontwikkeling van resistente bacteriën is een zorgwekken-

de trend in de geneeskunde in het algemeen en vooral in de kwetsbare groep van patiënten met een kunstgewricht. Verder onderzoek naar het voorkomen en de behandeling van deze ernstige complicatie is nodig om de hoge percentages patiënten met een uitstekend resultaat na een totale heupprothese operatie te kunnen blijven realiseren.
Lange wachttijden voor een totale heupprothese operatie veroorzaken een behoorlijk verlies in kwaliteit van leven (quality-adjusted life years (QALYs)) en zijn dus een vermijdbare bron van disfunctioneren. Met betrekking tot het wachtlijstprobleem wat nog steeds bestaat binnen de orthopedische chirurgie, met name voor gewrichtsvervangende operaties, zouden er nieuwe maatregelen genomen moeten worden om wachtlijsten te hanteren. Voorbeelden zijn het reduceren van de vraag naar operaties door het opschonen van wachtlijsten en prioriteit geven aan patiënten op de wachtlijst die klinisch gezien het meest een operatie nodig hebben, maatregelen die ook in andere landen toegepast worden. In de laatste plaats kan reorganisatie van zorgpatronen, zoals het verwijzen naar chirurgen met een kortere wachtlijst, het verminderen van door patiënten gemiste afspraken en het reduceren van het aantal controlebezoeken bij de specialist tot de mogelijkheden behoren.

In dit proefschrift zijn verschillende aspecten van outcome assessment bij totale heupprothese operaties aan de orde gekomen. Om beschikbare middelen zo goed mogelijk te gebruiken en onze patiënten de best mogelijke zorg te bieden, moet er gestreefd worden naar een continue kwaliteitsverbetering van de operatie. Deze kwaliteitsverbetering moet gebaseerd zijn op richtlijnen die volgen uit onderzoek met patiëntgeoriënteerde methodes in grote, representatieve groepen patiënten.

Acknowledgements

This thesis could not have been written without the help and support of many people. I would like to express my deep gratitude for their help in this project. Especially the following people have contributed, directly or indirectly, to great extent to the manuscript.

Wouter Dhert
Henk van Stel
Erik Buskens
Ria Matthijssen
Thomas Eisler
Rogier Donders
Ingeborg van der Tweel
Professor Henrik Malchau
Professor Guus Schrijvers
Professor Ab Verbout
Professor Peter Herberts
Kajsa Erikson
Marieke Vos
Annette Lengkeek
Marc Nijhof
André Fleer
Professor Olof Johnell
Louis Marting
Çumhur Öner
My fellow research residents
The sponsors

All the participating patients

Karin Hardus
Jorrit-Jan Verlaan

My parents, John and Joke Ostendorf

Roeland

Chapter 8

Curriculum Vitae

The author of this thesis was born as one of a twin on February 16th, 1975 in Almelo, The Netherlands. In 1993, she graduated from high school (Gymnasium, SG 'Jerusalem', Venray) and started to study medicine the same year at the University of Utrecht. In 1997, she started a research project at the Department of Orthopaedics (supervisor: Prof. dr. H. Malchau) and the Department of Biomaterials (supervisor: Prof. dr. P. Thomsen) of Gothenburg University, Sweden. In April 1998, she started a PhD project on several aspects of total hip arthroplasty in The Netherlands and Sweden at the Department of Orthopaedics of the University Medical Center Utrecht (head: Prof. dr. A.J. Verbout). The present work has resulted in several publications, a number of presentations at international conferences and this thesis. Since July 2004 she is doing a residency at the Department of Surgery at the Meander Medical Center in Amersfoort (head: Dr. G.H. Verberne) after which she will start her orthopaedic surgery training at the University Medical Center Utrecht.