Who should have total hip replacement?

Use of patient-reported outcome measures in identifying the indications for and assessment of total hip replacement

Meridith Emilie Greene

UNIVERSITY OF GOTHENBURG

Department of Orthopaedics
Institute of Clinical Sciences
at Sahlgrenska Academy
University of Gothenburg

2015

megreene@mgh.harvard.edu
meridithemilie@gmail.com

Massachusetts General Hospital
Harris Orthopaedic Laboratory
55 Fruit Street
GRJ 1125
Boston, MA 02114
USA

Typeset by Team Media Sweden AB
Cover illustration and similar illustrations by Pontus Andersson, Pontus Art Production
Printed in Gothenburg by Salomonsson Grafiska AB

e-publishing: http://hdl.handle.net/2077/38458
ISBN 978-91-628-9343-9

Contents

Abbreviations

Abbreviation	*Definition*
AUDIT	World Health Organization alcohol use disorders identification test
BMI	Body mass index
EQ-5D-3L	Three-level version of the EuroQol group's health-related quality of life measure
EQ-5D-5L	Five-level version of the EuroQol group's health-related quality of life measure
HADS	Hospital anxiety and depression scale
HHS	Harris hip score
HJR	Harris Joint Registry
HRQoL	Health-related quality of life
ICD-10	International Classification of Diseases 10th revision
JSW	Joint space width
MGH	Massachusetts General Hospital
MID	Minimal important difference
OA	Osteoarthritis
PRO	Patient-reported outcomes
PROM	Patient-reported outcome measure
SHAR	Swedish Hip Arthroplasty Register
THR	Total hip replacement
UCLA activity	University of California Los Angeles activity score survey
VAS	Visual analogue scale

Abstract

Background

Total hip replacement (THR) is a successful treatment for end-stage hip osteoarthritis (OA). Patients commonly seek this treatment to improve physical function, diminish pain, and ultimately to increase health-related quality of life (HRQoL). In recent years, patients have been asked to self-assess these areas using patient-reported outcomes measures (PROMs) both before and after treatment. Combining PROMS with national registers allows identification of factors that may influence how a patient will do after treatment. Detection of factors influencing poor outcomes after elective THR is important for understanding how to improve the effectiveness of this treatment.

Objectives

These works aimed to identify patient factors that contribute to better or worse patient-reported outcomes (PROs) after THR and to identify the most influential patient factors on surgical recommendation. In doing so, new PROMs were explored, as were various methodologies for investigating these types of data.

Patients and Methods

The first four papers utilized patients from the national Swedish Hip Arthroplasty Register (SHAR) while the last two papers include patients from the Harris Joint Registry (HJR). The influence of comorbid conditions, education, marital status, mental health, OA severity, and preoperative health states on surgical recommendations and patient-reported HRQoL, pain, and satisfaction after THR was explored. A new version of the EQ-5D survey was investigated as was how best to treat the relationship between the preoperative and postoperative EQ-5D index scores.

Results

On average, PROs improved after THR. Those who started with worse scores tended to improve similar amounts to those with better preoperative scores; however, due to their starting point, they did not achieve scores that were as high after surgery. Individuals with greater musculoskeletal comorbidities, with low or medium levels of education, and a history of preoperative antidepressant use, were identified as being patients who began and ended with worse PROs. The patient's joint space width had the greatest influence on THR recommendations. The new version of the EQ-5D survey appeared to better measure HRQoL in both preoperative and postoperative patients. Less ceiling effects were seen and substantial utilization of the new answer options occurred particularly before THR surgery.

Conclusions

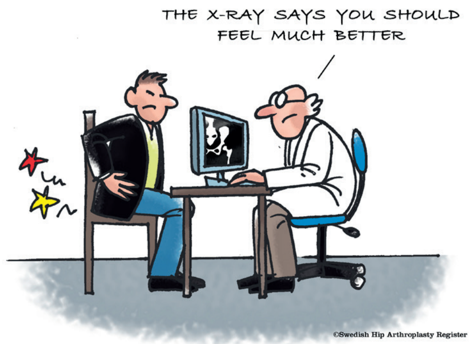
Patients at risk for poor outcomes can be identified through preoperative reporting of musculoskeletal comorbidities and their medical record. Clinicians are not discouraged from treating these patients, but rather are encouraged to discuss individual risk factors to aid in the decision-making process for the patient.

Background and Introduction

Total Hip Replacement

Osteoarthritis (OA) is a joint disease common in aging individuals.[78] In a Swedish population, hip OA ranged from less than 1% in patients younger than 55 and up to 10% in those over 85.[19] Because symptomatic OA results in chronic pain and functional disability, patients experience diminished health-related quality of life (HRQoL). If these symptoms persist despite non-surgical interventions like physical therapy and pain medication, total hip replacement (THR) is commonly recommended. THR is a highly effective treatment for patients suffering from end-stage OA of the hip.[78] Components are placed in the femur and the acetabulum of the pelvis as a means to replace the articulating ball-and-socket hip joint. The success of THR has been so great that it was named 'the operation of the century'.[61] This clinician-assessed surgical success however was traditionally based upon implant material and design performance assessed via radiographic analysis by surgeons and through surgeon-assessed functional status or survivorship of the implanted components. Survivorship or success was defined as an implant system remaining in a patient with no revision or exchange of components, rather than improvement in the patient's pain or functional status. While development of new implants continues, data suggest that many hip implants consistently have greater than 95% survivorship at 10 years.[29] Despite technical surgical success, a proportion of patients have persistent pain, diminished physical function, and/or dissatisfaction after 'the operation of the century'.[2,15,67,69]

Hip OA is a painful debilitating condition, but it is not life threatening. Treatment of OA with THR is common and safe, but like any surgical procedure, not without risks (e.g. the risk of fatal pulmonary embolism after THR ranges from 0.2% to 5%[32]). Because THR is most commonly elective and intended to improve HRQoL not to prevent death, the patient's functional improvement and satisfaction need also to define successful THR rather than just the implant survivorship. A promising way to improve upon 95% survivorship of a particular implant system is to shift focus to patient-reported outcomes (PROs) such as HRQoL, pain, and satisfaction. The patient may not be enthusiastic about an implant remaining in their hip for ten years if they are living with constant pain and inhibited function. Patient-reported outcome measures (PROMs) allow the patient's voice to be heard and become a part of the treatment process. Inclusion of PRO in assessing THR will allow for further improvement in this surgical treatment perhaps making it also the operation of the twenty-first century.



Patient-Reported Outcomes

A PRO is any account of a person's health status reported directly by the individual without interpretation by another person. While the term can be misleading, PROs are not limited to outcomes after an intervention, but rather can be reported at any point in time and represent the individual's personal assessment of their feelings or functional ability with respect to their health at that moment. Patient-reported outcome measures (PROMs) are the standardized instruments designed to measure specific elements, known as constructs and domains, of a person's health status. PROs are assessed using PROMs as a means to standardize the evaluation of a particular area of health, condition, or treatment rather than using qualitative interviews. The FDA encourages measurement of PROs for clinical trials that assess new medical devices and products because the patient's perspective is a critical piece of determining medical treatment efficacy.[80]

General versus Specific Measures

Common areas measured with PROMs in THR patients are HRQoL, general health and wellbeing and symptoms such as pain, functional impairment, stiffness, and activity. PROMs can be printed on paper or administered through electronic systems where the patient inputs their responses directly into a computer-generated survey. Patients complete surveys in the clinical office or at home via mailed forms or through a secure emailed internet hyperlink.[84]

PROs are measured using two types of PROMs: general and specific health measures. Both types of PROMs share valuable information about a patient's health status, but each provides a different look at the patient's condition. General health measures broadly assess health across subpopulations, medical conditions, or treatment groups. While general health measures do provide information on the individual level, they typically are intended to provide a more global look at health and allow comparisons between populations or treatment groups thus providing greater generalizability. Broad continued use of general measures adds to the cumulative knowledge of health and quality of life outcomes and can establish the relative burden of different diseases and the relative merit of different interventions.[74] Treatment policy or resource allocation decision makers tend to be more interested in differences between subjects rather than within-person changes of a particular treatment type; making general health measures particularly important for setting healthcare standards.

Specific measures, alternatively, are designed to target defined diagnostic groups, particular populations, body parts, or organ systems. They are typically utilized to observe changes in or responsiveness of a particular condition to a treatment on the individual patient level. Specific measures are most commonly administered at two or more time intervals to determine the within-patient change. Investigators implementing specific measures typically tailor the survey to the intervention of interest to understand specific patient concerns and identify small clinically important changes after treatment.[74] If well designed, specific measures provide a high level of specificity, but as a trade-off, have low generalizability outside the targeted population. To mitigate this, many studies which implement PROMs utilize both general and specific measures.

PRO Collection Challenges

Implementation of PROMs in any medical practice requires additional effort from the medical office staff and the patient. An organized system for distribution, collection, and retention of patient-reported surveys is critical to make proper use of the data. In order to enhance the rate of patient compliance, the questionnaire needs to be as brief as possible, while also providing enough valuable information to justify the collection effort. An extensive questionnaire consisting of multiple general health measures as well as several disease-specific surveys may provide a broad profile of the patient's health, but result in low levels of compliance due to the burden on the patient. When collecting PRO data on the national level, a short survey is critical to maintain high levels of patient compliance because all patients receiving THR are asked to participate. Numerous survey questions may be a deterrent for some patients resulting in low rates of compliance and diminished generalizability for national register-based observational studies. Cohort studies and clinical trials on the other hand, have a bit more leeway with the number of questions a patient can be asked. Participants in targeted prospective studies provide informed consent agreeing to complete the collection of selected survey questions. Therefore, the patient has an understanding of the time and effort necessary and consents to participation.

When selecting PROMs, it is also important to choose surveys which have been validated and their reliability tested to ensure that the questionnaire items are universally understood and measuring the same construct across all patients. Without validation and input from patients on their interpretation of survey questions, the investigator may believe they are collecting different information than the patient is providing (Table 1).

Table 1. Summary of Acceptability Criteria for a Validated PROM[97]

Validity	Content validation*	How well the content of survey items meets the criteria of experts
	Criterion validation	How well a scale correlates to the 'gold-standard' measurement of the area of interest
	Construct validation	How well a relationship between behaviors or attitudes is explained
Reliability	Repeatability	How reproducible the scale's results are under different conditions
	Internal consistency	How well items within the same domain correlate to one another
Responsiveness*		How well a scale can measure meaningful change in a clinical state[63]

**Not all survey development theorists find this necessary.[97]*

Patient-reported experience measures (PREMs) are important when clinics aim to assess or improve the patient experience within the healthcare setting. However, when asking a patient about satisfaction with their outcomes after treatment, the investigator wants to ensure the patient provides this rather than receiving satisfaction with the experience at the clinic. These subtle differences between PREMs and PROMs can influence results, thus confirmation of face validity of survey items is essential to confirm measurement of the area of interest.

Ideal intervals for questionnaire administration must also be established. Depending upon what information the clinician is interested in collecting, the questionnaire may need to be administered more or less frequently. One clinician maybe interested in health status immediately following a procedure while others may be more interested in how the patient is doing after the average recovery period. Similarly, different treatments are intended to provide relief from symptoms for varying amounts of time. Clinics interested in understanding how well an intervention has worked will need to administer their questionnaire both before and after the treatment. In order to understand how well a particular intervention is working, surveying patients at consistent intervals may provide a clearer picture of how the treatment influences changes over time.

PRO Interpretation Challenges

Interpretation of the patient-reported data can be challenging. Because there are many different health measures commonly found in the literature for THR patients, generalizability and direct comparisons between centers, regions, or nations can be limited. Even when the same instrument is used, scoring may vary between populations. The EQ-5D index is a weighted measure of HRQoL based upon responses to the five dimensions of the instrument. Several national value sets specific to their cultural norms exist based on time-trade-off or visual analogue scale (VAS) studies conducted on that country's general population. Because of cultural differences, populations may value one area of health higher than another. To account for these differences, national value sets weight the patient's responses differently. Therefore, comparisons of EQ-5D indices across nations cannot be done in a one-to-one fashion; trends may need to be considered rather than absolute index values.

There are two conflicting concerns about patient response trends that could influence the sensitivity or reliability of a PROM. First, end-aversion bias suggests that respondents are reluctant to select answers in the extremes because individuals do not want to make absolute judgments like 'always/never' or 'best/worst'.[97] In some cultures where individualism is not encouraged, responses in the extremes may be rare, ultimately causing one population to appear very different from another. Alternatively, ceiling and floor effects occur when the respondents answer predominantly in the extreme. Responses of this sort do not allow room to measure improvement or degradation over time or after treatment. Ceiling and floor effects also make it very difficult to distinguish between those who see good improvement versus those who see very good improvement and vice versa. Any continuous scale with end-points, such as a VAS or index, has the capacity to have ceiling and floor effects. The goal of an instrument though should be to provide enough levels between those end-points to minimize floor and ceiling effects. An overwhelming use of either response trend, end-aversion or floor/ceiling effects, may suggest the instrument is not sensitive or reliable to measure the area of interest in that population.

A common question, which arises with the presentation of PRO data, is whether changes measured correspond to clinically relevant improvement or degradation. Unfortunately, this is sometimes not a straightforward question to answer. Clinicians and policy makers are interested in the minimal important difference (MID) provided by a particular treatment as a means to assess efficacy or differences between groups. Several methodologies exist to calculate MIDs however these calculations differ greatly, the patient's opinion is not always included, and consensus of which to used does not exist.[57] When measuring subjective domains such as HRQoL or pain, assessment must come from the individual rather than dictated by the clinician. The significance of a MID is dependent upon the population used to calculate the value. MIDs calculated from individual responses may not translate to changes measured on the population level. For example, if a MID were established at the patient level and the average change for a population is below that MID value, then the distribution of change is more important than the average. A narrow distribution of change likely indicates that the treatment may not have been effective, but if the distribution of change was broad, it is likely that the treatment was productive or deleterious for some portion of the population.[12] Universal MID values for PROMs are theoretically appealing, but without a strong understanding of the implications of the MID on the patient versus the population level, they can be misleading. Some may argue that small changes on the population level are not clinically relevant, thereby dismissing a particular PROM

Table 2. Patient- and Procedure-related Data Classified by Levels of Registry Data

Data Type	Level I Data	Level II Data	Level III Data	Level IV Data
Patient-related	Personal ID Sex Diagnosis Ethnicity° Death	ASA score^{+} Height Weight Surgeon-defined Charnley Class°	PROMs Sick leave* Functional recovery*	
Procedure-related	Date of surgery Type of procedure Laterality Hospital ID Surgeon ID° Reoperation and/or revision	Prophylactic measures^{+} Surgical technique$^{\Delta}$ Surgical approach Implant details Fixation method Anesthesia type° Blood loss° Incision length° Local Complications	Adverse events* Costs*	Radiographs°

$^{+}$ *SHAR only*
° *Harris Joint Registry (HJR) only*
$^{\Delta}$ *Aggregated hospital level in the SHAR and surgery specific in the HJR*
* *Data obtained via linkage studies*

as unimportant. However, upon closer inspection, subgroups within the larger population may ultimately show highly significant differences in the benefit or lack of improvement from a particular treatment.

Swedish Hip Arthroplasty Register

The national Swedish Hip Arthroplasty Register (SHAR) is a prospective THR data repository that collects level I, II, and III data (Table 2).

The aim of the SHAR is to capture all THR cases nationally with the purpose of describing the epidemiology and the clinical outcomes of THR in Sweden and to efficiently identify any problems associated with the procedure. Complete prospective, national collection of surgical, component, follow-up, and patient-reported data provides an indispensable tool for clinical care. By following the national THR population over time both before and at regular intervals after treatment, the register is able to attain statistical power which is not possible in a single hospital or randomized trial. Rare complications associated with surgical techniques or implants are identified more quickly due to the huge sample size from the national register. The SHAR is one of the 12 full member registers of the International Society of Arthroplasty Registers (ISAR). Full ISAR membership requires over 80% compliance of national hospitals (coverage) and that those reporting provide a minimum completeness of 90% of the total joint replacement procedures from each medical unit.[48] In 2011, the SHAR reported 100% coverage with all hospitals conducting THR reporting to the register with 98% of all THRs reported.[29]

Benefits of National Prospective Observational (Register) Studies

Register studies remove biases common in epidemiological studies. Selection bias is mitigated by the complete collection of the THR patient population within the country ('completeness'). Information or recall bias is minimized due to the prospective nature of the surgical and patient-reported data collection. While data entry errors may occur, these are minimal.[29] Finally, because health is all encompassing, not all health-related confounders may be collected within the SHAR. Linkage studies, which merge additional interdisciplinary official national registers with the SHAR, provide additional risk factors and confounders for exploration allowing deeper understanding of outcomes after THR treatment.

The ability to conduct comprehensive post-market surveillance is greatly enhanced by registers. Development of new implant designs and materials for arthroplasty is ongoing and ever changing. Ideally, new technologies and surgical techniques would be introduced to the market in a step-wise fashion; starting with a small closely followed cohort to determine early safety, followed by larger multicenter monitoring, and finally investigated on a large scale in a register study.[68] Implementation of new technology or surgical techniques in this way may identify problems in a limited number of cases which could then be mitigated or eliminated from the market altogether. Step-wise introduction can eliminate catastrophic failure rather than allowing early introduction of new technologies nationwide before they are vetted. Because of the statistical power, analysis of implants and techniques in a register allows for stratification of possible cofounders to identify whether differences are related to the implant or technique in question. These observational register studies are not designed to determine causation, but rather to provide evidence-based monitoring to identify problems. For this reason, observational national register studies work in concert with cohort and randomized trials, where causation may ultimately be determined.

Traditionally, total joint replacement registers are used to monitor component performance with survivorship defined by revision. Kaplan-Meier and Cox regression analyses are typically used to identify sub-optimal implants due to high rates of revision. Used in this way, registers are useful in assessing surgical techniques and specific component efficacy. The pitfall of using revision as the only endpoint or outcome of THR is that the patient's voice is not heard and neither their satisfaction nor their HRQoL is taken into account when assessing this primarily elective procedure. Surgical technique and component reliability are essential elements of THR surgery, but the PROs are equally important when evaluating efficacy of the treatment as mentioned earlier. With the introduction of PROMs, the SHAR became an effective tool to assess not only surgical techniques and component performance, but equally important patient satisfaction, their pain before and after treatment, and their HRQoL.

Introduction of PROMs to the SHAR

The SHAR began the PROM program in 2002, which was gradually adopted and has been active nationwide since 2008. Preoperatively, 86% of patients complete the set of questionnaires while the response rate at one year follow-up is 90%.[83] In order to prevent the influence of clinic staff on patient responses, the follow-up questionnaire is completed by the patient at home. They are asked to complete the EQ-5D, the musculoskeletal co-morbidity Charnley classification survey, a VAS for pain, and after surgery, a VAS for satisfaction with their outcomes after treatment. The questionnaire is administered to the patients preoperatively (excluding satisfaction) and at 1, 6, and 10 years postoperatively. The EQ-5D consists of five health dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The patient chooses from three answer options for each dimension: no problems, moderate problems, or extreme problems. From their responses, a weighted health index is calculated representing the patient's HRQoL. Index scores correspond to health states ranging from perfect health to death and to states worse than death. In addition to the five dimensions, the patient is also asked to complete a VAS of their impression of their overall health on that day from zero to one hundred (EQ VAS). The Charnley classification survey assesses whether the patient has unilateral hip disease (class A), bilateral hip disease (class B), or hip disease as well as other conditions which negatively influence their ability to walk (class C). The patient rates the level of their pain on the pain VAS from zero (no pain) to one hundred (worst imaginable pain), and after treatment, the patient is asked to rate their satisfaction with the outcomes from treatment from zero (complete satisfaction) to one hundred (complete dissatisfaction) on the satisfaction VAS. The combination of surgical data and patient-reported data makes it possible to establish whether specific risk or protective factors contribute significantly to the patient's life after THR.

Harris Joint Registry

The Harris Joint Registry (HJR) is a local total hip and knee replacement registry, maintained by the Harris Orthopaedic Laboratory at Massachusetts General Hospital (MGH). The HJR collects all four levels of data (Table 2). The collection of level IV radiographic images in a local registry is more logistically feasible to implement than on the national level. The PROM protocol in the HJR comprises the EQ-5D, the Charnley classification survey, a pain VAS, a satisfaction VAS, the Harris hip score (HHS),[66] and the University of California Los Angeles (UCLA) activity score[3] as the standard of care for all THR patients. The HHS is a disease-specific survey measuring the outcomes of THR from zero to one hundred where 44% of the score is associated with pain. The UCLA activity score rates the patient's activity level on a scale from one (inactive) to ten (regular participation in impact sports). Radiographs and PROMs are obtained preoperatively (excluding the satisfaction VAS) and at standard clinical follow-up intervals at 6 to 10 weeks (radiographs only) and 1, 3, 5, 7, and 10 years.

On January 1, 2012, the standard PROM protocol in the arthroplasty clinic at MGH was updated. All new patients without a history of THR complaining of hip symptoms interested in discussing THR with the surgeon received three additional PROMs. The arthroplasty service expanded the extended PROM protocol in September 2012 to include any patient interested in discussing primary THR whether they had received a contralateral joint replacement or not. The new protocol added the Hospital Anxiety and Depression Scale (HADS),[106] the World Health Organization's Alcohol Use Disorders Identification Test (AUDIT),[87] and the Aberdeen Participation survey.[77] Any individual completing the new surveys is enrolled in the program and is asked to complete the surveys again at subsequent follow-up visits.

Benefits and Limitations of a Local Registry

The HJR is not a hospital-wide joint replacement registry. It targets the arthroplasty clinic at MGH and captures 96% of targeted primary procedures.[6] Therefore, it provides a useful tool for the participating clinicians and researchers, but cannot indicate how the institution as a whole is doing with respect to outcomes or surgical techniques. The ability to identify very rare outcomes is substantially less in a local registry the size of the HJR than with a national registry. However, because of the size, the HJR is able to collect all four levels of data where storage and organization of large files like radiographs is not a problem. The limited number of surgeons contributing to the registry will also limit the catalogue of implant data collected by the HJR as many surgeons have their preferred implant manufacturers and systems thus limiting conclusion that can be drawn about rarely used implants.

A major challenge for the HJR is continued follow-up of all registered patients. If a patient were dissatisfied with treatment at MGH and required revision surgery or contra-lateral treatment, it is conceivable that the patient may go to a different hospital for treatment. Unless the hospital was affiliated with the Partners Healthcare system, this would not be captured by the HJR. Therefore, success rates in the HJR are over estimated and the generalizability of its data is minimized. This challenge will be the same for any institutional registry until a national system for tracking THR procedures is established based on a unique patient identifier such as social security numbers in the United States.

The ability of the HJR to collect both PROMs and radiographs and easily associate these with surgical and demographic data is very powerful. Trends of implant use can be tracked with respect to radiographic and PROs and feedback can be provided to clinicians. Subsequent clinician improvement or degradation may then also be tracked over time.[39]

PROMs in the Swedish Hip Arthoplasty Register and Harris Joint Registry

EQ-5D

The EuroQol group's patient-reported measure the EQ-5D is a generic HRQoL survey used by both the SHAR and the HJR.[25] The survey consists of five dimensions measuring different areas of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. In the original version of the survey, the respondent chooses from three levels that define each dimension: no problems, some or moderate problems, and extreme problems. The EuroQol group developed a new version of the survey giving the respondent five levels of responses from which to choose: no, some, moderate, severe, and extreme problems.[43]

The response options for the three-level can result in 243 (3^5) unique health states, which in turn, can be translated into a weighted health index. Different countries have different index value sets that reflect response norms for the given population. Until recently when a Swedish version became available, the SHAR used the British value set to score the EQ-5D index. The HJR uses the United States value set for reporting the EQ-5D index. The American three-level index (derived from time trade-off responses) can range from -0.109 to 1.00 where 1.00 corresponds to perfect health, 0 corresponds to death, and negative indices correspond to health states perceived to be worse than death.[1,91] The new five-level version has 3,125 (5^5) unique health states possible. Currently the five-level survey does not have a unique value set to calculate an index score, but a 'cross-walk' from the three-level does exist.[20,101]

Unique health states are defined by a particular combination of responses to each of the five dimensions. For the three-level survey *no problems* in all dimensions would be notated as 11111, while *extreme problems* in all five dimensions would be notated as 33333. For the five-level survey, a response of *no problems* in all dimensions is once again notated as 11111, while *extreme problems* in all five dimensions is notated as 55555, and so on.

The final component of both versions of the EQ-5D survey is a vertical VAS assessing the patient's subjective rating of their overall health status that day on a scale from zero to the best possible rating of 100 (EQ VAS). While both the EQ-5D index and the EQ VAS are measures of HRQoL, they measure different elements of HRQoL and should be considered separately.

The EQ-5D is a brief survey making it appealing to both patients and clinicians. Because it is a general health measure, it can be used to compare populations and cost effectiveness across different disease and treatment groups. However, the EQ-5D index in particular has been criticized in the literature.[8,11,37,52,58,65,98] Because the index is bounded, it can be useful for looking at a snapshot of a population at a particular point in time, but unfortunately, if one were interested in measuring change over time or after a particular intervention, floor or ceiling effects may cause limitations. For example, if an individual had a high EQ-5D index prior to treatment, they would have very little room for improvement resulting in a ceiling effect. Conversely, an individual with a relatively low HRQoL would have a much greater capacity for improvement. Thus making the magnitude of change highly dependent upon where the patient began on the scale. Another challenge with the EQ-5D index is that despite describing it as a continuous scale between the bounds, the index for some value sets behaves more ordinal in nature with patients clustering at certain index values. In a population of OA patients eligible for THR in Sweden, British value set indices of 0.1 and 0.7 were very common.[83]

Each of these challenges, the bounded index possibly leading to floor or ceiling effects and the multimodal distribution of indices need to be accounted for when performing statistical analyses of the EQ-5D index which rarely happens. An additional challenge with the original EQ-5D-3L version was whether with three response options, the survey was sensitive enough to pick up changes in fairly healthy populations such as those eligible for THR due to the aforementioned ceiling effects.

Correlation and regression are the most common methods used to analyze EQ-5D data. Neither correlation nor regression alone are able to handle bi- or multi-modal distributions of EQ-5D indices. It is important to find the right structural relationship between the pre- and postoperative EQ-5D indices when investigating this outcome measure.

Pain VAS

The pain VAS is implemented pre- and postoperatively in the SHAR and the HJR. The Swedish version of this survey ranges from zero to 100 where 100 is the respondent's worst imaginable pain. For the HJR, the scale ranges from zero to 10, but follows the same trend as the Swedish version where a rating of 10 corresponds to the respondent's worst imaginable pain. Zero on both scales represents no pain.

Satisfaction VAS

The postoperative satisfaction VAS is the last common PROM between the SHAR and the HJR. Like the pain VAS, the satisfaction VAS is displayed horizontally, and in Sweden, it ranges from zero to 100 while at MGH, it ranges from zero to 10. For each version, zero corresponds to complete satisfaction and the high end of the scale corresponds to the greatest level of dissatisfaction with the outcomes from treatment.

Harris Hip Score

The hip-disease specific Harris hip score developed by Dr. William H. Harris of the HOL in 1969 is a standard survey given to all hip patients in the arthroplasty clinic at MGH.[40] The score was not originally designed for a

THR population, but is one of the most broadly used outcome measures in the THR literature. The scale has a maximum of 100 points consisting of four domains: pain (up to 44 points), hip function (up to 47 points), deformity (up to 4 points), and range of motion (ROM) (up to 5 points). The original Harris hip score was staff-administered, but has since been converted into a self-reported survey.[66] The deformity domain was originally included to account for patients who had major deformities due to traumatic arthritis. Because this domain rarely applies to standard THR patients, it was set as a constant, and therefore, the lowest possible self-administered Harris hip score is 4. The ROM domain was also standardized for the self-administered survey providing up to 5 points to the overall score (possible points are 0, 3, or 5). Because it is unreasonable to ask the patient to define their ROM, the allotted points for this domain are established based upon the response combination to the shoes/socks and sitting questions. Traditionally, postoperative Harris hip scores below 70 indicated poor hip outcomes. It can be seen in the literature that fair outcomes had scores from 70 to 80, good outcomes had scores from 80 to 90, and scores from 90 to 100 were considered excellent outcomes. However, categorization of scores is misleading and should be a practice of the past. Because outcome scores are so dependent upon their case mix and their preoperative score, they should not be categorized in this way. For this reason, the Harris hip score in paper VI was treated as a continuous variable.

Despite the extensive use of the Harris hip score, the survey has critics. The score shows high rates of ceiling effects in THR patients. For this reason, its usefulness for measuring relevant changes after THR is questioned.[102] At its introduction to the literature in 1969, the Harris hip score was not properly vetted through what are now considered standard psychometric tests for health questionnaires looking at validity, reliability, and responsiveness (Table 1). It was compared to two rating systems common at the time the Larson and Shepard systems, but only for score distributions.[59] Given that there are high rates of ceiling effects in THR patients today, the content validity of this measure could be questioned. As pointed out by Wamper and colleagues, the Harris hip score probably had very good content validity in the population for which it was designed, but indications for THR have changed since 1969 and it may not measure as much as was originally intended.[102] Groups have however reported good construct validity for the Harris hip score with comparisons to the Western Ontario and McMaster Universities Osteoarthritis Index, the Short Form 36, and the Nottingham Health Profile.[30,94,95] Söderman and Malchau found the staff-administered version of the score to be reliable after testing and retesting.[94]

University of California Los Angeles (UCLA) Activity Score

The UCLA activity score is a standard survey administered to all hip and knee patients in the arthroplasty clinic at MGH. It consists of a single question asking the respondent to identify their most appropriate activity level. The score ranges from 1 (wholly inactive; dependent on others; cannot leave residence) to 10 (regularly participate in impact sports such as jogging, tennis, skiing, acrobatics, ballet, heavy labor, or backpacking).[3] Like the Harris hip score, this measure was originally presented in a paper investigating a specific patient population, and as it is presented in the paper, no psychometric tests were performed during the design or implementation of the survey.

Since the introduction of the UCLA activity score, groups have looked at some of the psychometric qualities of the survey. Naal and collaborators concluded that the UCLA activity score was reliable, feasible, and valid for use in THR patients.[73] However, they drew these conclusions based on only weak or moderate correlations with hip disease specific measures commonly used for THR patients and with references to Zahiri and colleagues who used investigator administered UCLA activity score surveys.[73,105] Zahiri's group did ask the patient to rate their activity, but this was done on a VAS 'relative to other people' rather than with the UCLA activity score itself. Ultimately these measures were correlated, but correlations were weak.[105] Many agree that some measure of activity is important in assessing THR outcomes and success, but no gold-standard exists.[7,73,105] In order to minimize the burden on the patient, the UCLA activity score was the brief survey selected to do this in patients at MGH.

Aberdeen Participation Survey

The Aberdeen participation survey is one of the included instruments in the new PROM protocol for the arthroplasty clinic at MGH.[77] This survey consists of nine questions investigating how the respondent's hip condition influences participation in activities of daily living. According to the International Classification of Functioning, Disability, and Health three areas of health outcomes should be explored when using PROMs: Impairment, activity limitation, and participation restriction.[76] Pollard and colleagues developed a measure for each domain which could work either in conjunction with one another for patients with arthritis or as stand-alone measures.[76,77] Impairment and activity were already covered in the standard PROM protocol in the HJR with the Harris hip score and the UCLA activity score and therefore only the Aberdeen participation survey was implemented so as not to over burden patients with redundant questions. Scores range from 9 to 45 where 9 represents an individual with no apparent participation restriction, and those with 45 have extreme participation restriction due to their joint disease. At present, no cut points have been published establishing ranges for low, medium, or high participation restriction.

PROM Summary

Due to the national coverage of the SHAR the PROM protocol was purposefully kept brief (11 questions) to minimize the burden on patients and increase the response rate.[83] The HJR puts a greater burden on the patient with 20 questions for the original protocol and up to 53 questions with the addition of the surveys for new patients. The HJR predominantly collects PROMs electronically when the patient comes to the clinic for follow-up while the SHAR uses paper forms mailed to the patients at their designated follow-up intervals. The HJR hopes to transition to an email based system where PROMs are collected whether the patient returns for follow-up or not; however this has not successfully been implemented as of yet. It is likely that the HJR will have to minimize the number of surveys administered or questions asked in order for the email system to be successful. Results presented in paper VI suggest that some surveys may not contribute significantly to predicting who will be recommended for THR or who will decide to move forward with the treatment, but those measures may prove to be useful in predicting who will have successful outcomes, and therefore have not been removed from the protocol yet.

Patient-reported Comorbidity Screening Instruments in the SHAR and HJR

Charnley Classification Survey

The patient-reported Charnley classification survey is used by both the SHAR and the HJR. The questions in this survey identify the musculoskeletal comorbidity status of a patient based on the classifications defined by Sir John Charnley.[14] Individuals with unilateral hip disease are classified as A. Those with bilateral hip disease are classified as B, and anyone with multiple joint disease or other problems that inhibit the individual's walking ability are classified as C. Some have argued that class B should be divided into two separate groups accounting for those who have one side or the other already treated, but this has not been sufficiently supported in the literature. It is also possible that the surgeon can assign a Charnley classification to a patient based on their clinical assessment, and therefore, readers of THR literature should be cognizant of which version of this musculoskeletal comorbidity classification system was implemented.

Hospital Anxiety and Depression Scale (HADS)

The HADS survey is part of the new PROM protocol in the arthroplasty clinic at MGH. All new preoperative hip and knee patients are enrolled in the new PROM protocol and receive this survey at their first visit to the clinic and will again receive it at all subsequent visits. The survey was developed for patients in non-psychiatric hospital departments.[106] It is broken into two pieces assessing anxiety and depression separately and providing a summary score for each.[106] There are fourteen questions; half dedicated to the anxiety subscale and the other half to the depression subscale. Scores on both subscales range from zero to 21. Scores up to 7 are indicative of 'non-cases', scores from 8 to 10 are doubtful cases, and scores of 11 or greater are definite cases with low rates of false positives.[106]

This survey was added to the HJR PROM protocol as a means for the arthroplasty clinicians to screen for patients who may be experiencing anxiety or depressive disorders. Patients with depression tend to have less pain reduction and are less satisfied after surgical treatment.[81,89] By screening for these patients before surgery, the clinician can discuss this risk with the patient before undergoing THR.

Alcohol Use Disorders Identification Test (AUDIT)

The WHO AUDIT survey is one of the measures included in the new PROM protocol for the arthroplasty clinic at MGH. The survey screens respondents for risky alcohol use implementing up to ten questions. If the respondent were to indicate that they do not drink alcohol on the first question, the respondent answers two more questions and the survey ends. For those who do consume alcohol, the system administers the complete ten-question survey. The scores can range from zero to 40. Individuals whose score is from zero to 8 are regarded as safe alcohol users, 8 to 15 may have a medium level of alcohol problems, and scores above 16 may indicate a high level of alcohol problems.[87]

Aims

Study Objectives

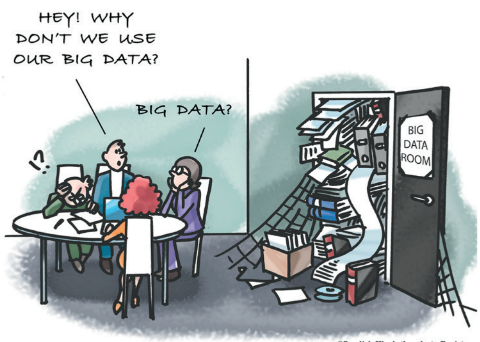
These works aim to investigate and describe several patient factors associated with PROs after THR as well as identify differences among individuals who are indicated and opt to undergo THR and those who do not. The specific objectives were to:

- Explore how socioeconomic, marital, and comorbid health statuses are associated with patient-reported HRQoL, pain, and satisfaction with THR one year after surgery.
- Understand whether mental health status and treatment of mental health conditions are associated with patient-reported HRQoL and pain before and after treatment of OA with THR as well as if they are associated with the patient's satisfaction with the outcome of THR one year after treatment.
- Investigate multiple models to improve the analysis of EQ-5D index profiles for use in clinical outcomes studies both preoperatively and postoperatively.
- Validate whether the new five-level version of the EQ-5D survey will provide a more discriminating measure of patient-reported HRQoL in THR patients by adding intermediate response options to the previous three-level version.
- Calculate the probability that a patient is indicated and will be recommended for THR and whether they will move forward with the procedure after considering demographics and radiographic signs of arthritis as well as patient-reported HRQoL, pain, function, mental health, alcohol use, and participation in daily activities.

Patients

Swedish Hip Arthroplasty Patients

Primary THR patients with a diagnosis of OA from the SHAR were the focus of the first four papers. Participation in the pre- and postoperative PROM program was required and patient age at surgery, gender, and Charnley classification noted. Data from the SHAR was merged with Swedish National Patient Register, the Prescribed Drug Register at the National Board of Health and Welfare and Statistics Sweden via the unique patient identifier. Linkage of these national registers provided additional information about medical comorbidities, antidepressant drug prescriptions and utilization, education attainment, and marital status.



©Swedish Hip Arthroplasty Register

The inclusion criteria for the first four papers were similar. Individuals in the SHAR had to have complete preoperative and 1 year postoperative PROMs. These included EQ-5D, Charnley classification survey, pain VAS, and satisfaction VAS (at one year). They could not have a revision within 1 year of their surgery (excluding paper I), and for bilateral patients, only the first hip with complete pre- and postoperative PROMs was included in the analyses.

Paper I

Individuals included in paper I had surgery between January 2002 and December 2007. These cases were merged with the Swedish National Patient Register to obtain any other diagnoses beyond the patient's hip OA as a means to calculate three of the International Classification of Disease-based comorbidity measures: Elixhauser, Charlson, and the Royal College of Surgeons (RCS) Charlson.

Paper II

Those included in paper II had surgery between January 2005 and December 2007. These cases were merged with the Swedish National Patient Registry to obtain comorbid conditions, and the cases were also merged with data from Statistics Sweden to obtain the individuals' highest level of education and the patients' marital status. The Charlson's comorbidity index was calculated for all patients up to two years before THR.

Paper III

Patients had surgery between July 2006 and December 2007 in paper III. These cases were merged with the Prescribed Drug Register to determine which THR patients purchased antidepressant medications up to a year before surgery. The Prescribed Drug Register began recording all prescription purchases in Sweden in July 2005 which is what limited the THR patient inclusion criteria.

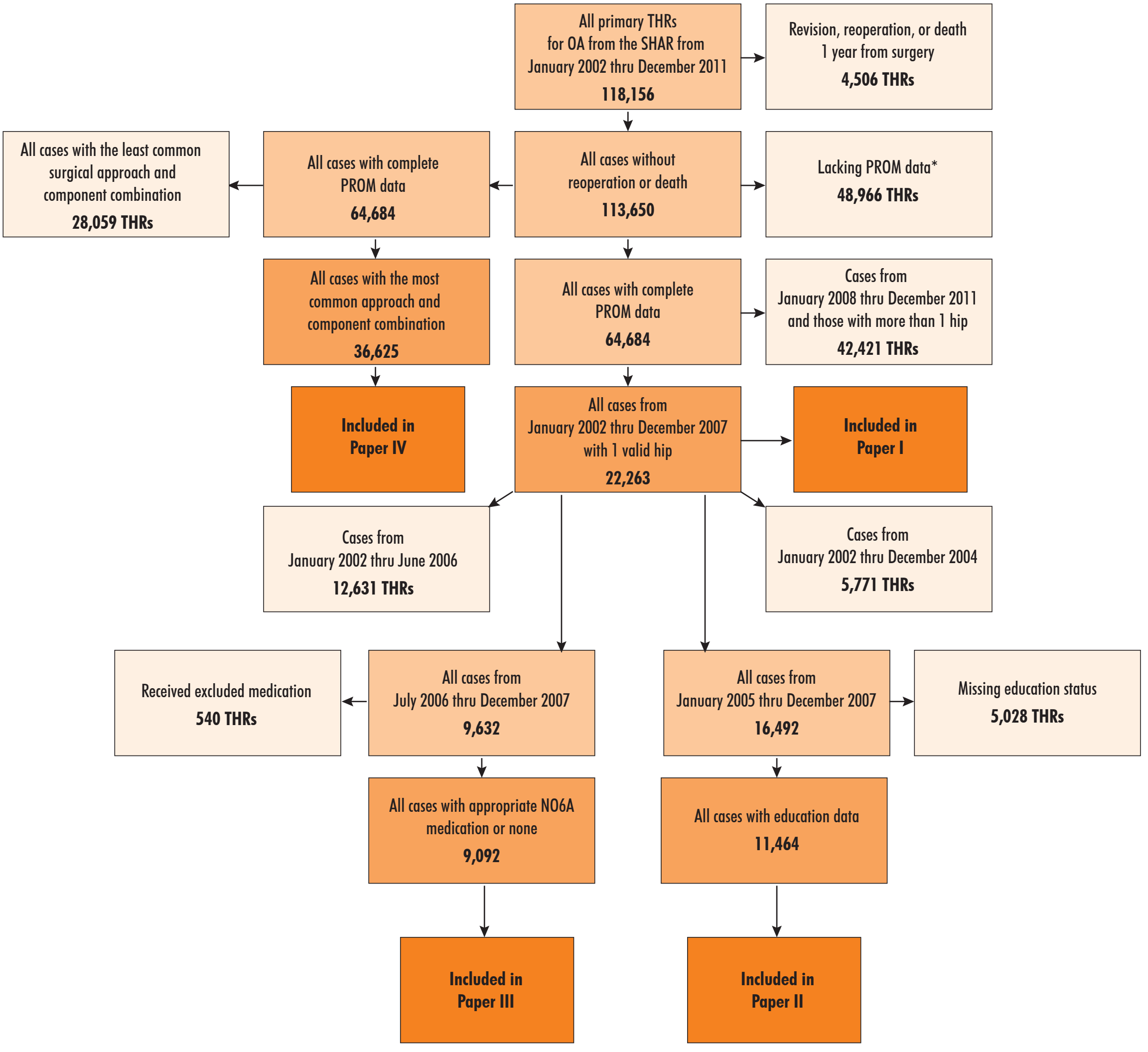
Paper IV

Inclusion criteria were most broad for paper IV where all THR patients operated between January 2002 and December 2011 with pre- and postoperative PROMs and no revisions or death within the first year after surgery were included in the analysis.

Table 3. Patient Population Counts for Each Paper

Paper Number	Number of Patients	Patient Source
I	22,263	SHAR
II	11,464	SHAR
III	9,092	SHAR
IV	36,625	SHAR
V	127	MGH
VI	325	MGH

Figure 1. Patient Selection from the Swedish Hip Arthroplasty Register



**The SHAR PROM program began in 2002 at 11 hospitals. Participation gradually increased until 2008 when it was active nationwide.*

Massachusetts General Hospital Patients

Paper V

Individuals were prospectively recruited for the validation of the EQ-5D-5L survey presented in paper V. Patients complaining of hip problems who had yet to undergo THR and those who were 1 to 6 years post THR surgery without a revision were invited to participate. The patient-reported HRQoL of the patients who agreed to participate did not differ from those of the patients who did not. Fifty preoperative and seventy postoperative participants were required to compare response trends from the EQ-5D-3L survey to the EQ-5D-5L version.

Paper VI

All patients complaining of hip problems participating in the new PROM protocol in the arthroplasty clinic at MGH between January 2012 and December 2013 were considered for the analysis in paper VI. They could not have had an earlier THR on the side for which they were visiting the clinic, and the clinician had to determine that the problem they were encountering was in fact due to their hip and not referred pain due to another musculoskeletal problem.

Massachusetts General Hospital Bulfinch Building in Boston. Contained within this building is the Ether Dome; the location of the first public use of ether as a surgical anesthetic in 1846.

Methods

Papers I, II, and III

The general study structure was similar for papers I, II, and III. The influence of one or more patient factors on PRO 1 year after surgery were investigated using SHAR data. Linkage to other national health and demographic data from additional national registers in Sweden facilitated these works. All national register data was prospectively collected according to their own protocols and therefore these were all observational studies.

Table 4. SHAR Linkage Studies

National Register Used	Paper I	Paper II	Paper III
Swedish Hip Arthroplasty Register	X	X	X
Swedish National Patient Register	X	X	
Statistics Sweden		X	
Prescribed Drug Register			X

Four national databases were utilized for papers I through III. Patients from the SHAR were linked to information in the other databases via a national patient identification number.

Paper IV

While the study aims were different for paper IV, the data utilized for illustrative purposed was collected in the same way as papers I through III from the SHAR. As a means to investigated alternative ways to present changes in EQ-5D index data we aimed to find the 'right' structural relationship between the pre- and postoperative EQ-5D indices to obtain the best estimation of the effect of the preoperative score on the postoperative score. Four models were investigated. The first was a null model which only had an intercept, next was a single line model, then a 2 line model with single transition point, and finally we looked at 3 line model with 2 change points.

Paper V

Individuals who agreed to participate in the validation of the EQ-5D-5L survey, which was detailed in paper V, were asked to complete both the old and new versions of the survey to determine if the newer version was equally or more sensitive for determining the patient's HRQoL. There were at least two weeks between the survey version administrations and half of the enrolled patients did the EQ-5D-3L first and the other half did the EQ-5D-5L first. At the point of recruitment in the arthroplasty clinic at MGH, the first survey was completed either on a tablet or at a touchscreen kiosk. The patient then selected their preferred method for completion of the second survey either by a paper form in the mail or via a secure link sent to their email. Individuals who failed to complete the second survey in a timely manner were contacted by phone to confirm that they were interested in continued participation. This usually motivated the patient to complete the second survey.

Paper VI

In paper VI, once the pre-surgery individuals who participated in the new PROM protocol were identified in the HJR, several additional data points were collected from either the registry or the medical record: age, gender, marital status, ethnicity, education, and body mass index (BMI). Anterior/posterior (AP) pelvis radiographs were obtained when available and AP hip images were used if the pelvis image did not exist in the HJR. The minimal joint space width (JSW) was measured on the hip of interest and the severity of OA was graded according to Tönnis.[99] Where 0 was no OA, 1 was mild OA, 2 was moderate OA, and 3 was severe OA. The office visit notes were reviewed for all patients and the surgeon's recommendation was documented. These recommendations were categorized in three ways; THR was recommended, THR was not recommended now, or THR was not recommended at all. Reasons for delaying a THR recommendation included the need to control other risk factors such as weight loss or smoking or drug use cessation, their symptoms were not bad enough to warrant surgery yet and non-operative treatment was recommended, or further work up was necessary to determine if their hip was in fact the cause of their problems. THR was not recommended to individuals who had risk factors that made major surgery too dangerous or the patient's problems were not due to their hip.

Statistical methods

Papers I, II, and III

The first three papers implemented linear regression analyses where PROs (EQ-5D index and EQ VAS, pain VAS, and patient satisfaction with the outcomes of THR) were the dependent variables. The various papers explored different patient demographic variables as well as preoperative HRQoL and pain as the dependent variables. Assessment of coefficients and confidence intervals determined the level of association of each significant variable on the outcomes. Each of the first three papers included patient-reported Charnley classification in the tested models in addition to the demographic variables of most interest: paper I looked at the influence of the International Classification of Diseases (ICD)-based comorbidity measures (Elixhauser, Charlson, and RCS Charlson); paper II explored the influence of the patient's highest level of education, their marital status, age, and gender; and in paper III the models accounted for age, gender, self-reported anxiety and depression, and whether the patient took antidepressant medication up to 1 year before THR surgery.

The regression analyses used in paper I included the three ICD-10-based comobidity measures, Charnley classification, and the preoperative score of the outcome in question as the independent variables. No other patient demographic variables were included in the final analysis for two reasons. First, gender and age each contributed less than 1% to the predictive power of the models, and second, we wanted to find the greatest predictive power contributed by the ICD-10-based comorbidity measures. Therefore, gender and age were excluded from these analyses.

Papers II and III implemented some subtle differences in their statistical methods. Paper II, looking at the influence of education attainment and marital status on the outcomes of interest, used Bayesian model averaging to identify the significant predictors of each outcome parameter allowing the models to include only significant independent variables with posterior probabilities of 0.50 or greater.[33,47,54,79] This process identified both the EQ-5D index as well as the EQ VAS as independent predictors; therefore, each model included both measures of HRQoL. Paper III, investigating the influence of antidepressant prescription usage on PROs, also implemented Bayesian model averaging to select the influential variables for each regression model. Paper III however tested each model with two-line linear regression splines to determine if a change point should be implemented accounting for patients with a low or high preoperative health status as detailed in paper IV. The EQ-5D index was the only model that benefited from using the piecewise linear regression splines with a change point at a preoperative EQ-5D index of 0.051.

Paper IV

Paper IV differed from the first three in that it was a methodological investigation of how to treat the preoperative EQ-5D index variable when conducting linear regression modeling for outcomes research. The paper explored four regression models to determine which model best predicted outcomes in an OA population from the SHAR as the example. This methodology is useful for modeling the pre-treatment EQ-5D index

Table 5. Statistical Tests Utilized for Each Project

Paper	Statistical Tests						
	Linear Regression	Bayesian Model Averaging	Piecewise Splines	Correlations	McNemar's Test	Random Forest	Flexible Discriminant Analysis
I	X						
II	X	X					
III	X	X	X				
IV	X		X				
V	X			X	X		
VI	X					X	X

in populations where HRQoL is expected to improve and that the HRQoL before treatment will influence HRQoL after treatment.

Paper V

In paper V where responses to EQ-5D-5L survey were compared to those from the EQ-5D-3L version, response trends were compared on a case by case basis. Ceiling and floor effects were investigated using McNemar's test for each dimension and for the surveys as a whole. To test convergent validity of both versions of the survey, Spearman's rank correlation coefficient was calculated between the EQ VAS scores between the two survey versions and with each of the five dimensions of the corresponding survey version. Finally, the change in EQ VAS from one version to the next was modeled in two linear regression models. First, against the response trends (same, new, or different) for each of the fives dimension and second, against the time between the administrations for both the preoperative and postoperative groups where both models controlled for the order with which the versions were administered.

Paper VI

In the final paper, thirteen different algorithms were tested to determine which had the best predictive power to determine the probability of whether a patient would be recommended for THR and also the probability that the patient would move forward with the surgery. The thirteen tested algorithms could be classified in three ways: linear classification, nonlinear classification, and classification trees and rule-based models. Predictive power was determined by four measures of accuracy. The area under the curve (AUC) was compared for each model, as were the sensitivity, the specificity, and the negative and positive predictive values as a means of identifying the best model for our dataset.

Centre of registries, Västra Götaland, Gothenburg, Sweden. The house has 27 National Quality Registries (SHAR included) and is one of six centers in Sweden that support and help develop additional registers.

Summary of Papers

Paper I

Standard comorbidity measures do not predict patient-reported outcomes 1 year after total hip arthroplasty: Charnley class better predictor of outcomes

This study compared the predictive capacity of three measures of comorbidity (Elixhauser's, Charlson's, and Royal College of Surgeons [RCS] Charlson's) at 1, 2, and 5 years before surgery to the musculoskeletal measure of comorbidity the Charnley classification for patient-reported HRQoL, pain, and satisfaction 1 year after THR.

Results

Of the measures of comorbidity considered, only the Charnley classification was significantly associated with all PROs. In fact, the RCS Charlson comorbidity score had no significant associations with any of the PROs (Table 6).

Regardless of the timeframe used to calculate the ICD-10-based comorbidity scores, Charnley classification and both preoperative measures of HRQoL were the strongest predictors for the postoperative HRQoL. Charnley classification and the preoperative pain VAS were the strongest predictors for the postoperative pain VAS in all time-frame models. Finally, irrespective of the timeframe, Charnley classification was the strongest predictor for the postoperative satisfaction VAS with marginal influence from the Elixhauser comorbidity score. For all outcomes, there was a gradual increase in the model's predictive power with the length of the timeframe considered when calculating the ICD-10-based comorbidity measures.

Paper II

Education attainment is associated with patient-reported outcomes: Findings from the Swedish Hip Arthroplasty Register

This study investigated the influence of a patient's highest level of education, as a surrogate for socioeconomic status, and their marital status on patient-reported HRQoL, pain, and satisfaction 1 year after THR.

Results

The majority of patients were married (Figure 2) and had medium or high levels of education (Figure 3).

BMA was used to determine which variables to include in each outcome analysis. Charnley classification, preoperative EQ-5D index, preoperative EQ VAS, and educa-

Table 6. Influence of Variables on Patient-reported Outcomes

Measure	EQ-5D index		EQ VAS		Pain VAS		Satisfaction VAS	
	coeff.	CI_{95}	coeff.	CI_{95}	coeff.	CI_{95}	coeff.	CI_{95}
Intercept	0.785	0.778, 0.791	71.727	70.938, 72.516	4.752	3.83, 5.674	13.243	12.82, 13.667
Charnley Class								
A	ref		ref		ref		ref	
B	-0.064	-0.073, -0.054	-4.392	-5.196, -3.588	4.187	3.444, 4.929	2.959	2.120, 3.797
C	-0.112	-0.119, -0.105	-8.95	-9.509, -8.391	5.848	5.335, 6.36	5.431	4.854, 6.007
Charlson*	-0.023	-0.035, -0.011	-3.407	-4.400, -2.414	0.688	-0.23, 1.607	0.88	-0.158, 1.918
RCS Charlson*	0.003	-0.008, 0.015	-0.016	-1.012, 0.980	0.245	-0.675, 1.166	-0.074	-1.115, 0.967
Elixhauser*	-0.010	-0.015, -0.004	-0.886	-1.353, -0.420	0.522	0.091, 0.954	0.628	0.141, 1.116
Preoperative value	0.147	0.137, 0.156	0.174	0.162, 0.186	0.096	0.082, 0.111	-	-
R^2	0.107		0.108		0.037		0.018	

** Each of the ICD-10-based comorbidity measures were calculated 1 year before THR* tion level were the four variables consistently included in all regression models (Table 7). None of the variables were excluded from all analyses, but a couple of variables, marital status and preoperative pain scores, were only included in one regression model, the postoperative pain VAS.

Patients with high levels of education reported higher levels of HRQoL, lower levels of pain, and greater satisfaction than their lesser educated peers. Charnley class C had the greatest detrimental effect on both measures of HRQoL and pain 1 year after THA, but age for those over 26 years had the greatest detrimental effect on satisfaction followed by Charnley class C.

Figure 2. Marital Status of Patient Population

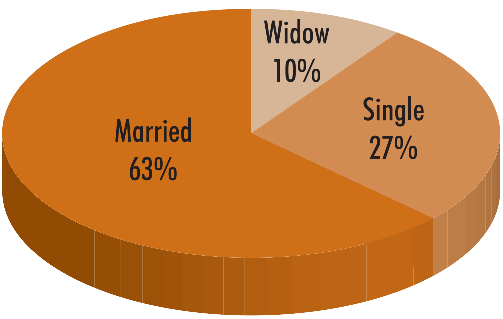


Figure 3. Education Level of Patient Population

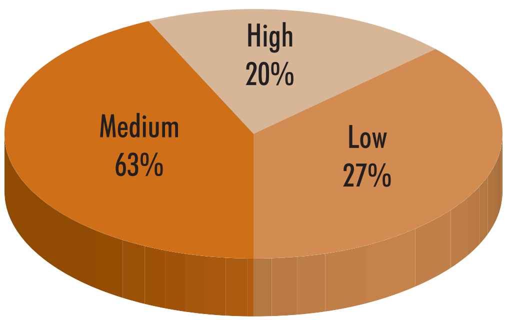


Table 7. Variables Included in Each Regression Analysis as Selected by BMA

	Patient-reported Outcomes			
Predictors	EQ-5D Index	EQ VAS	Pain VAS	Satisfaction VAS
Age			X	X
Gender	X			X
Comorbidity	X	X		
Charnley Classification	X	X	X	X
Preop EQ-5D index	X	X	X	X
Preop EQ VAS	X	X	X	X
Preop pain VAS			X	
Hip order		X		X
Marital Status			X	
Education level	X	X	X	X

Linear regression results with the slope coefficients (coeff.) and the 95% confidence intervals (CI_{95}) for each variable.

Figure 4. Influence of Categorical Variables on Postoperative EQ-5D Indices

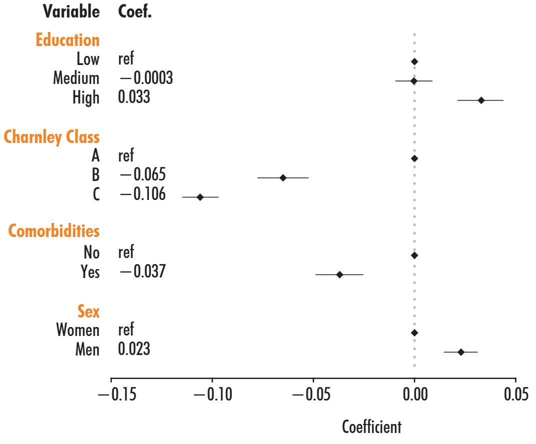


Table 8. Influence of Continuous Variables on Postoperative EQ-5D Indices

	Coeff.	CI_{95}
Pre EQ-5D Index	0.106	0.092, 0.120
Pre EQ VAS	0.001	0.001, 0.001

Figure 4 and Table 8
Preoperative HRQoL had significant influence on postoperative HRQoL. Men and those with high levels of education had more favorable postoperative scores, while those with musculoskeletal comorbidities did worse.

Figure 5. Influence of Categorical Variables on Postoperative EQ VAS Scores

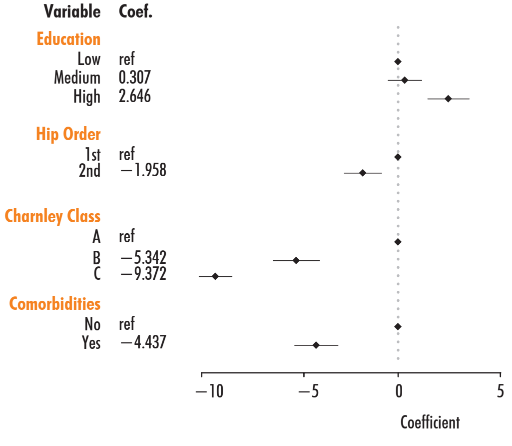


Table 9. Influence of Continuous Variables on Postoperative EQ VAS Scores

	Coeff.	CI_{95}
Pre EQ-5D Index	5.151	3.943, 6.360
Pre EQ VAS	0.138	0.121, 0.154

Figure 5 and Table 9
Similar to the postoperative EQ-5D index, preoperative HRQoL influenced postoperative EQ VAS scores. Highly educated patients reported higher scores and those with musculoskeletal comorbidities lower scores. Gender did not influence EQ VAS reporting, but having the second hip done resulted in slightly lower HRQoL than those only having the first.

Categorical variables with a coefficient of zero acted as the reference. Confidence intervals which include zero indicate patient demographics that were not significantly different from the reference category.

Figure 6. Influence of Categorical Variables on Postoperative Pain VAS Scores

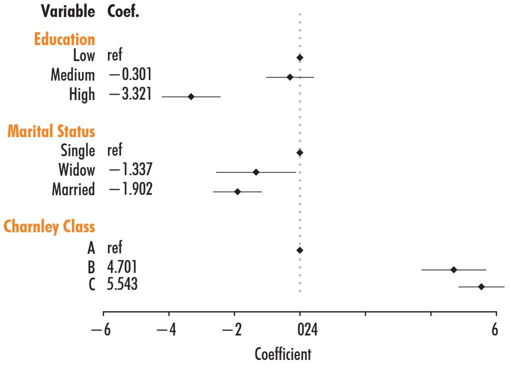


Table 10. Influence of Continuous Variables on Postoperative Pain VAS Scores

	Coeff.	CI_{95}
Age (years)	0.087	0.044, 0.129
Pre EQ-5D Index	-3.058	-4.334, -1.782
Pre EQ VAS	-0.047	-0.063, -0.031
Pre Pain VAS	0.041	0.018, 0.065

Figure 6 and Table 10
Postoperative pain was less in individuals with high education, high preoperative HRQoL, and low preoperative pain. Individuals with musculoskeletal comorbidities and those who were single tended to have greater postoperative pain.

Figure 7. Influence of Categorical Variables on Postoperative Satisfaction VAS Scores

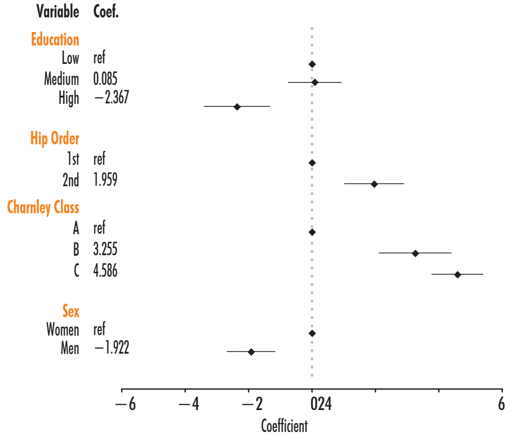


Table 11. Influence of Continuous Variables on Postoperative Satisfaction VAS Scores

	Coeff.	CI_{95}
Age (years)	0.170	0.121, 0.219
Pre EQ-5D Index	-2.363	-3.673, -1.054
Pre EQ VAS	-0.060	-0.078, -0.042

Figure 7 and Table 11
Men, those with high levels of education, individuals classified as Charnley A, and those with high HRQoL before surgery were the most satisfied with the outcomes from their THR surgery.

Paper III

Does the use of antidepressants predict patient-reported outcomes following total hip replacement surgery?

This study looked at the influence of the use of antidepressant medication up to one year before THR on patient-reported HRQoL, pain, and satisfaction 1 year after surgery.

Results

Patients using antidepressants had poorer HRQoL and higher levels of pain before and after THR and on average experienced less satisfaction from the treatment. Ten percent of the population (n=943) acquired antidepressant medication at least once during the preoperative period. Unfortunately over a third of these individuals had no text indicating the reason for the prescription (Figure 8) therefore this information could not be included in the regression models. The majority of antidepressant patients obtained their medication continuously throughout the observation period (Figure 9). The prevalence of antidepressant usage increased with the self-reported anxiety/depression severity in the fifth dimension of the preoperative EQ-5D survey (Figure 10).

Figure 10. Proportion of Responses to Preoperative EQ-5D Anxiety/Depression Dimension

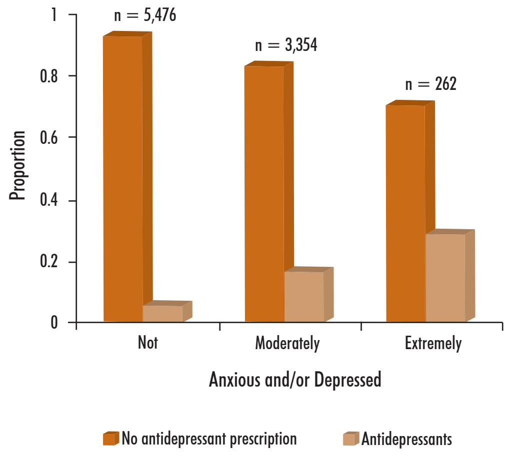


Figure 8. Diagnoses for Antidepressant Prescriptions

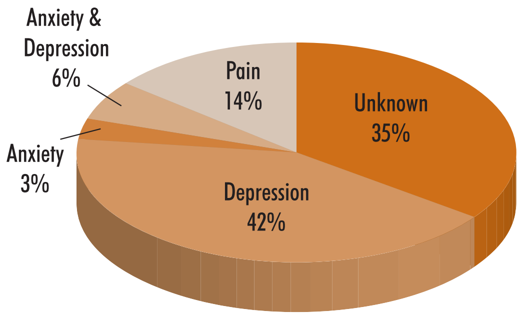
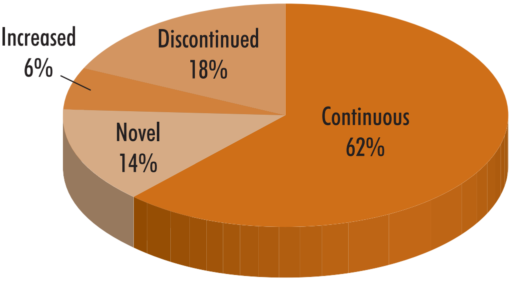


Figure 9. Antidepressant Medication Usage Patterns



After BMA, Charnley classification, preoperative EQ VAS, and the fifth dimension of the preoperative EQ-5D were the four variables consistently included in all regression models (Table 12). Hip order was not significant for any of the analyses.

Table 12. Variables Included in Each Regression Analysis as Selected by BMA

	Patient-reported Outcomes			
Predictors	**EQ-5D Index**	**EQ VAS**	**Pain VAS**	**Satisfaction VAS**
Age (years)	X	X	X	X
Gender				X
Charnley Classification	X	X	X	X
Preoperative EQ-5D Index	X	X		
Preoperative EQ VAS	X	X	X	X
Preoperative Pain VAS	X		X	
Hip Order				
5th Dimension of EQ-5D	X	X	X	X
N06A*	X	X	X	
Usage*	X	X		

** The models were tested with either the dichotomous antidepressant variable (N06A) or the categorical antidepressant usage variable (usage) at one time. All other variable posterior probabilities remained the same regardless of whether the N06A or usage variable was included.*

Linear regression results with the slope coefficients (coeff.) and the 95% confidence intervals (CI_{95}) for each variable.

Figure 11. Influence of Categorical Variables on Postoperative EQ-5D Indices

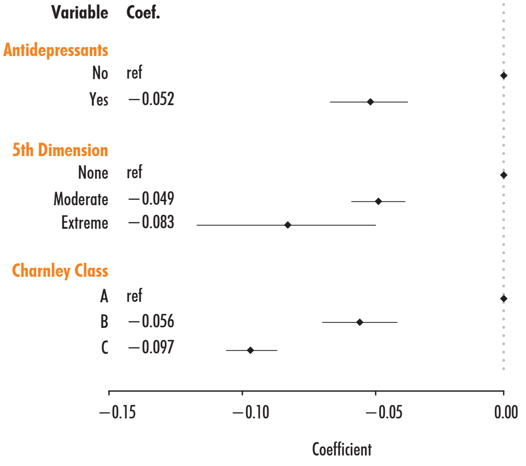


Table 13. Influence of Continuous Variables on Postoperative EQ-5D Indices

	Coeff.	CI_{95}
Age (years)	-0.002	-0.002, -0.001
Pre EQ-5D Index		
Low	0.497	0.296, 0.698
High	0.034	0.014, 0.055
Pre EQ VAS	0.001	0.0006, 0.0010
Pre Pain VAS	-0.001	-0.0009, -0.0003

Figure 11 and Table 13
The EQ-5D index was the only model that benefited from the use of piecewise linear regression splines with a change point at a preoperative EQ-5D index of 0.051. Individuals with preoperative EQ-5D index values less than 0.051 had low preoperative HRQoL while those with index values of 0.051 or greater, had high preoperative HRQoL.

Figure 12. Influence of Categorical Variables on Postoperative EQ VAS Scores

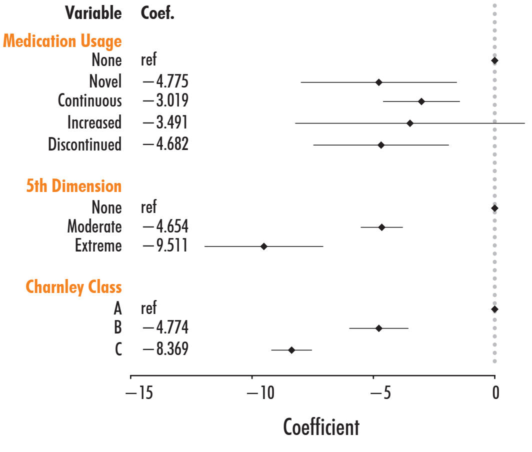


Table 14. Influence of Continuous Variables on Postoperative EQ VAS Scores

	Coeff.	CI_{95}
Age (years)	-0.233	-0.271, -0.194
Pre EQ-5D Index	2.645	1.218, 4.072
Pre EQ VAS	0.121	0.102, 0.140

Figure 12 and Table 14
Usage trends only influenced the EQ-5D index and EQ VAS. Of the various usage trends, increased antidepressant dosages had the greatest negative influence on EQ-5D indices at 1 year. Conversely, increased dosage was the only usage trend that was not associated with the patient-reported EQ VAS at 1 year.

Categorical variables with a coefficient of zero acted as the reference. Confidence intervals which include zero indicate patient demographics that were not significantly different from the reference category.

Figure 13. Influence of Categorical Variables on Postoperative Pain VAS Scores

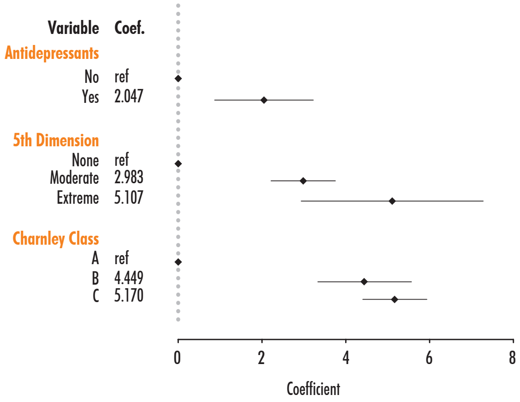


Table 15. Influence of Continuous Variables on Postoperative Pain VAS Scores

	Coeff.	CI_{95}
Age (years)	0.083	0.047, 0.119
Pre EQ VAS	-0.045	-0.062, -0.027
Pre Pain VAS	0.058	0.036, 0.079

Figure 13 and Table 15
Antidepressant medications were associated with increased patient-reported pain at one year, but those reporting severe anxiety/depression on the preoperative EQ-5D survey also reported significantly higher postoperative pain.

Figure 14. Influence of Categorical Variables on Postoperative Satisfaction VAS Scores

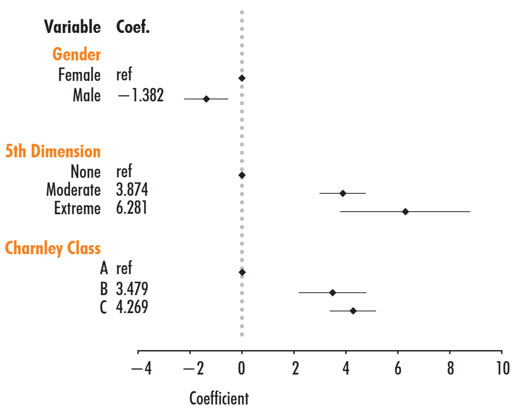


Table 16. Influence of Continuous Variables on Postoperative Satisfaction VAS Scores

	Coeff.	CI_{95}
Age (years)	0.207	0.165, 0.249
Pre EQ VAS	-0.043	-0.062, -0.024

Figure 14 and Table 16
Antidepressant prescriptions were not associated with satisfaction with outcomes 1 year after surgery. However, self-reported anxiety/depression was associated with satisfaction as much as or more than the presence of additional musculoskeletal comorbidities.

Paper IV

Improved statistical analysis of pre- and post-treatment patient-reported outcome measures (PROMs): The applicability of piecewise linear regression splines

This work explores several linear relationships to identify the best way to understand the relationship between preoperative scores and postoperative scores using EQ-5D indices from the SHAR as the example.

Results

The vast majority of the 36,625 patients studied reported improved postoperative HRQoL, however there were patients who reported no change in their EQ-5D index and a proportion who actually got worse after THR (Figure 15).

If the relationship between the preoperative and postoperative scores was normally distributed, a distribution like that of figure 16A would be expected; however, the responses were actually multimodal as shown in figure 16B where the red arrow indicates the most common pre- and postoperative responses.

Regression modeling showed that a two-line regression equation best described the relationship between patient-reported pre- and postoperative EQ-5D indices (Figure 17). The estimated change-point was at EQ-5D

Figure 15. Change in EQ-5D Indices from Preoperative to 1 Year Postoperative

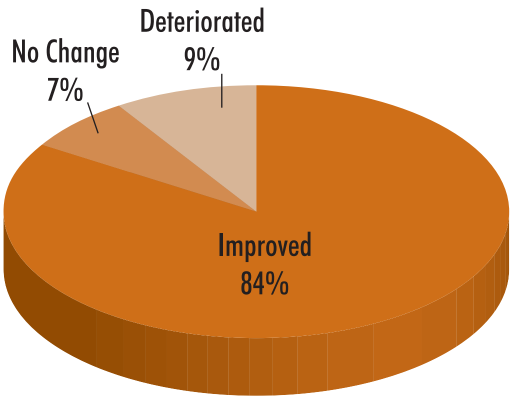
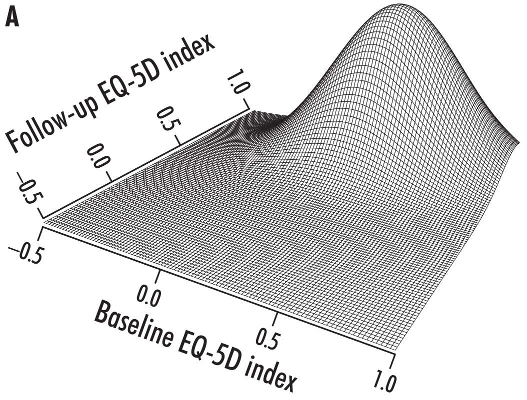


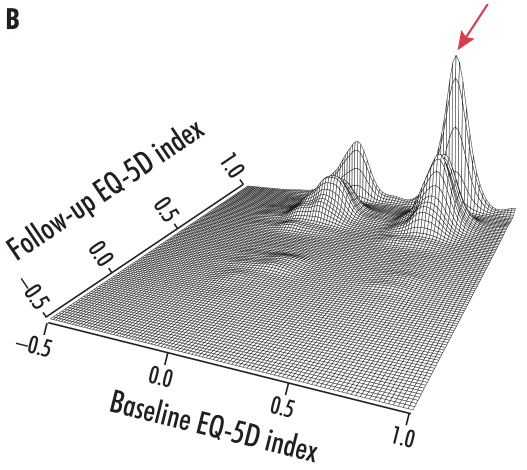
Figure 16. Distribution of Pre- and Postoperative EQ-5D Indices

A



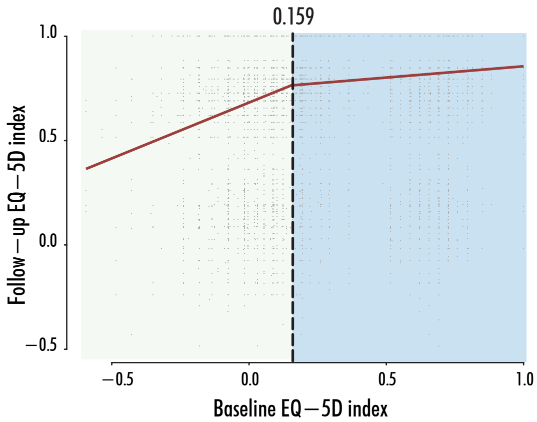
This figure indicates the distribution of pre- and postoperative indices if the relationship were normally distributed with the vertical dimension indicating the frequency of response combinations

B



This figure indicates the true multimodal distribution of pre- and postoperative EQ-5D index scores with the red arrow indicating the most common combination of repsonses

Figure 17. Distribution of Pre- and Postoperative EQ-5D Indices with Two-line Piecewise Linear Regression Splines



The splines describe the relationships between the preoperative and postoperative EQ-5D indices with a single change point at the baseline score of 0.159. Those to the left of the line are considered to have low preoperative HRQoL and those to the right high preoperative HRQoL.

index 0.159 (95% CI: 0.135, 0.182) as indicated by the dotted line. Those who had a preoperative EQ-5D index of 0.159 or less were classified as having low preoperative HRQoL (n=13,157 (35.9 %)) and those over 0.159 as having high preoperative HRQoL (n=23,468 (64.1 %)).

For those with low preoperative HRQoL, each one-unit increase in their preoperative index resulted in an improvement of 0.528 (95%CI: 0.480, 0.576) in their postoperative index. The intercept for this piece of the linear regression line was 0.682 (95%CI: 0.677, 0.686), which indicates a trend of postoperative improvement.

For individuals who had a high preoperative HRQoL, they reported a lower rate of improvement in the postoperative index: 0.106 (95% CI: 0.091, 0.121). The intercept for the second half of the linear regression line was 0.748 (95%CI: 0.742, 0.755), also indicating a trend of improvement, just less dramatically so.

Paper V

The EQ-5D-5L improves on the EQ-5D-3L for health-related quality of life assessment in patients undergoing total hip arthroplasty

This study compared patient responses on two versions of the EQ-5D survey, the original three-level and the new five-level, to identify if the five-level version might provide a better profile of preoperative and postoperative THR patient health status.

Results

The intermediate response options in the new five-level version of the EQ-5D survey were used by the majority of patients in the majority of dimensions in the preoperative group particularly in the dimensions most commonly associated with hip OA; mobility, usual activities, and pain (Figure 18). The postoperative patients were more likely to select the same answer options in four of the five dimensions in both versions of the survey, except for the pain dimension, where they were more likely to use the new responses (Figure 19).

No problems in all five dimensions on both the three-level and five-level surveys was the most common response for postoperative patients. However, for the pain/discomfort dimension the new responses were chosen 43% of the time, which was the most common for this dimension for these patients.

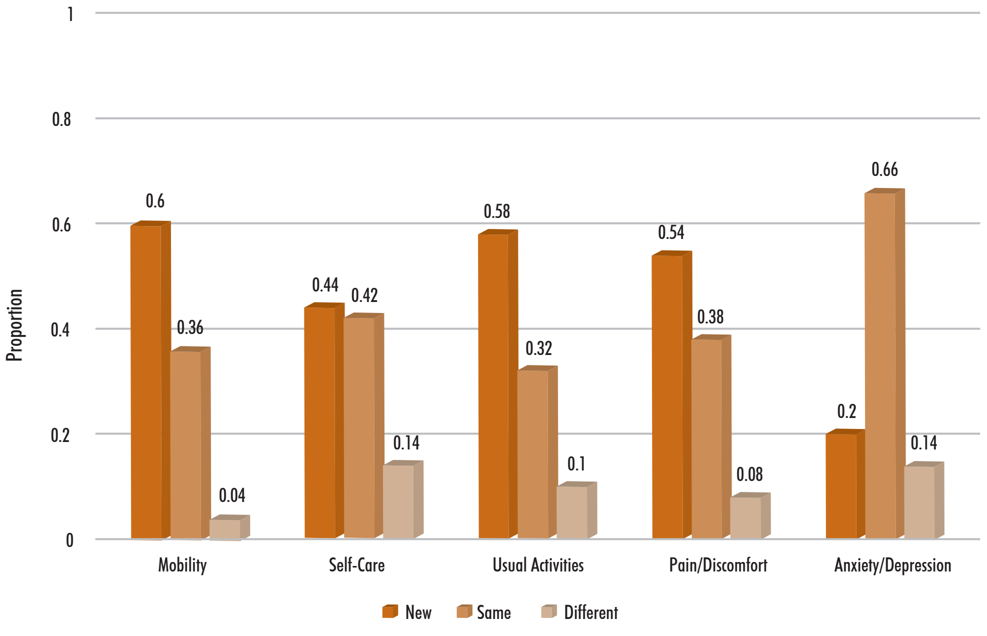
Some patients chose response options that were more than one response away from their first, these were classified as different in both pre- and postoperative patients. These different responses were as high as 14% in some dimensions in the preoperative patients and as high as 18% in one dimension (pain/discomfort) for postoperative patients.

Floor effects were not a major problem for either patient group in either survey. The only dimension with floor effects above 6% was in the pain/discomfort dimension for preoperative patients in the three-level survey, which was significantly decreased to 4% in the five-level survey (p=0.008).

Ceiling effects were diminished by the new response options in the five-level survey in all dimensions for both patient groups; however not all decreases were significant. The only dimension with a significant decrease in ceiling effects for the preoperative patients was in self-care. Conversely, all dimensions except anxiety/depression were significantly diminished for the postoperative group (Figure 20).

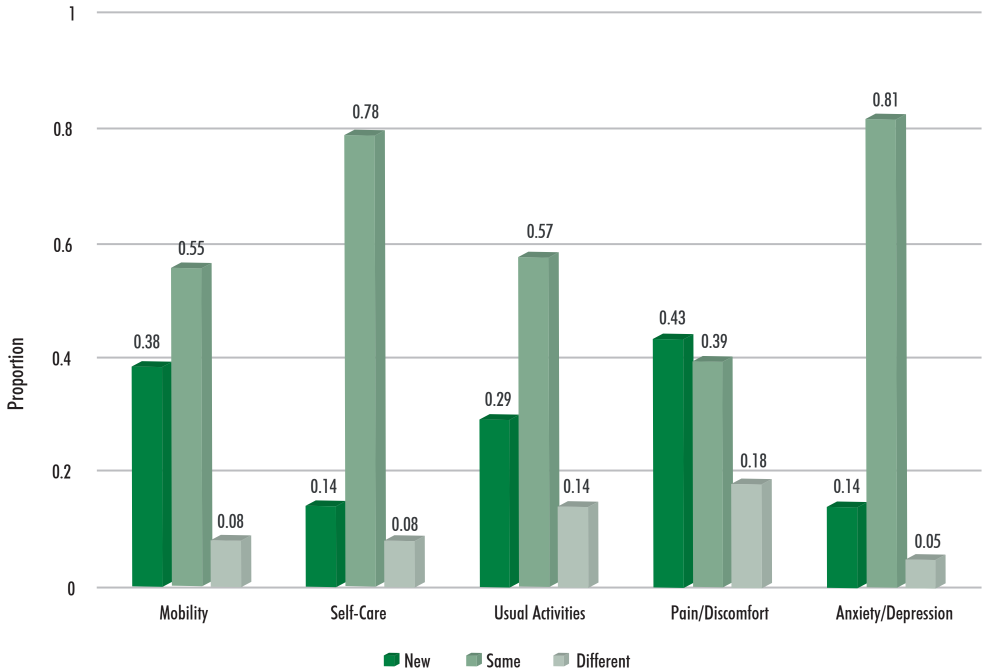
The convergent validity between the survey versions and their corresponding EQ VAS was good. Change in the VAS score from the three-level survey to the five-level survey was correlated to the response patterns in the pain/discomfort dimension. Those reporting more pain from one version to the next also reported on average 13 points lower on the VAS ($p = 0.04$). Postoperative VAS scores also decreased as worse problems were reported in each of the five dimensions, but were not found to be significant (all $p > 0.05$). With the doubling of response time between the survey versions, preoperative patients reported on average 3 units less on the EQ VAS ($p = 0.05$). Despite this difference on the VAS, there was no change in the relationship observed between the response trends to the five dimensions in either group of patients and the completion time between the different survey versions.

Figure 18. Preoperative Patient Response Trends to EQ-5D Surveys



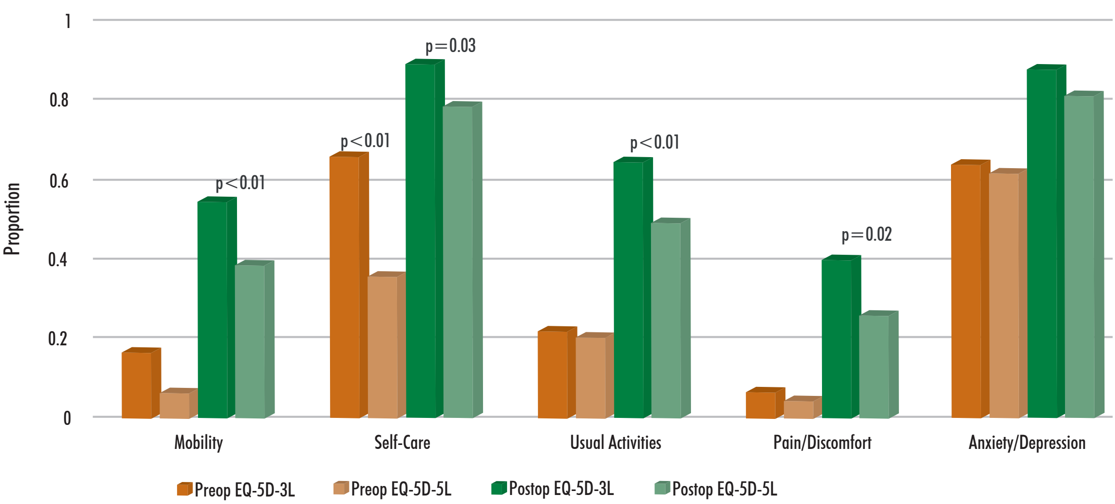
Preoperative patients took advantage of the new response options in the mobility, usual activities, and pain dimensions.

Figure 19. Postoperative Patient Response Trends to EQ-5D Surveys



Postoperative patients most commonly reported no problems in both versions of the survey except in the pain dimension where they utilized the new response options.

Figure 20. Ceiling Effects in Preoperative and Postoperative Patients



Ceiling effects were significantly diminished in only the self-care dimension for the preoperative group, but were diminished for mobility, self-care, usual activities, and pain/discomfort for the postoperative group.

Paper VI

Predicting who will be recommended for total hip replacement and those who will proceed: A tool for surgical recommendations

The final project looked at which patient factors had the greatest influence on predicting whether a patient would be recommended for THR and also whether the patient would decide to move forward with surgery.

Results

Surgical Recommendations

Of all the algorithms tested, random forest had the greatest predictive capacity for identifying the probability that a patient would be recommended for THR surgery and was therefore chosen as the selected model. If the probability threshold were set to 0.5, the random forest model had a 0.837 sensitivity and 0.569 specificity. If the probability threshold were optimized for both sensitivity (0.679) and specificity (0.751) it would be set at 0.653. Of the 17 variables included in the model, JSW, BMI, Harris hip score, OA grade, and Aberdeen participation had the greatest influence on whether a patient was recommended for surgery (Figure 21).

Those recommended for THR versus those who were not recommended now were older, had lower BMIs, narrower JWS, higher OA grades, more participation limitations according to Aberdeen, greater pain, worse HRQoL, reported less activity on the UCLA activity survey, and had worse hip symptoms according to the Harris hip score (Table 17).

Proceeding with Surgery

There were two models that stood out from those tested for predicting the probability of a patient proceeding with surgery: flexible discriminant analysis and random forest. The area under the curve of flexible discriminant analysis model was higher than that of random forest (p=0.003); however, its classification accuracy was not

Figure 21. Relative Influence of Each Variable on Surgical Recommendation

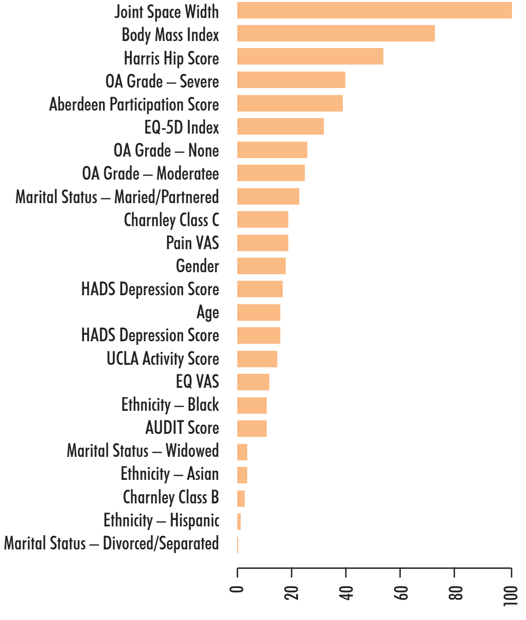


Table 17. Patient Demographics Stratified by Surgical Recommendation

	Recommendation		
	Surgery	Not Now	p
Count	192	133	
Gender – Female	103 (53.6)	82 (61.7)	0.187
Age	60.21 (10.74)	55.79 (13.59)	**0.001**
Ethnicity			0.842
White	181 (94.3)	120 (93.0)	
Black	2 (1.0)	3 (2.3)	
Asian	6 (3.1)	4 (3.1)	
Hispanic	3 (1.6)	2 (1.6)	
Marital Status			0.323
Single	24 (12.6)	26 (19.7)	
Married/Partnered	139 (73.2)	90 (68.2)	
Divorced/Separated	18 (9.5)	9 (6.8)	
Widowed	9 (4.7)	7 (5.3)	
Body Mass Index	27.79 (4.40)	29.37 (7.43)	**0.019**
Joint Space Width (mm)	0.61 (0.74)	1.47 (1.19)	**<0.001**
Osteoarthritis Grade			**<0.001**
None	0 (0.0)	5 (4.2)	
Mild	10 (5.6)	37 (31.1)	
Moderate	80 (45.2)	49 (41.2)	
Severe	87 (49.2)	28 (23.5)	
AUDIT	2.85 (3.03)	2.82 (3.46)	0.928
Aberdeen	20.90 (6.49)	18.83 (7.25)	**0.008**
HADS – Anxiety	5.01 (3.89)	5.33 (4.26)	0.480
HADS – Depression	4.88 (3.43)	4.36 (3.64)	0.197
Charnley Classification			0.773
A	78 (40.6)	57 (42.9)	
B	33 (17.2)	19 (14.3)	
C	81 (42.2)	57 (42.9)	
Pain VAS	5.82 (2.09)	5.22 (2.38)	**0.017**
EQ VAS	72.35 (18.43)	74.08 (17.86)	0.404
EQ-5D Index	0.62 (0.18)	0.66 (0.21)	**0.050**
UCLA Activity	4.76 (2.18)	5.40 (2.54)	**0.016**
Harris Hip Score	54.46 (14.45)	60.93 (18.21)	**<0.001**

uniformly better. Additionally, the specificity of random forest was better than that of flexible discriminant analysis; therefore, random forest was the selected model once again. If the probability threshold were set to 0.5, the random forest model had a 0.756 sensitivity and 0.611 specificity. If the probability threshold were optimized for both sensitivity (0.772) and specificity (0.612) it would be set at 0.588. Of the 17 variables included in the model, JSW, HHS, the EQ-5D index, gender, and preoperative pain appeared to have the greatest influence on whether a patient proceeded with surgery (Figure 22).

Those who chose to have THR versus those did not were older, had narrower JWS, had higher OA grades, had more participation limitations according to Aberdeen, had greater pain, worse HRQoL according to the EQ-5D index, reported less activity on the UCLA activity survey, and had worse hip symptoms according to the Harris hip score (Table 18).

Figure 22. Relative Influence of Each Variable on Surgery Decision

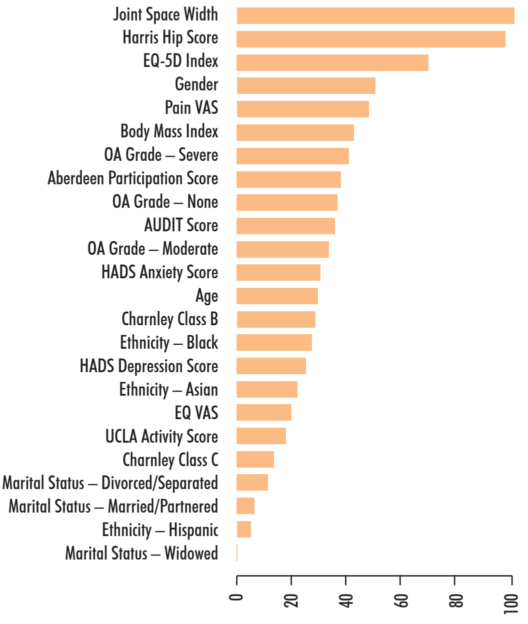


Table 18. Patient Demographics Stratified by Surgical Decision

	Proceeded with Surgery		
	Yes	No	p
Count	148	177	
Gender – Female	74 (50.0)	111 (62.7)	**0.028**
Age	60.14 (11.91)	56.95 (12.22)	**0.018**
Ethnicity			0.452
White	142 (95.9)	159 (91.9)	
Black	1 (0.7)	4 (2.3)	
Asian	3 (2.0)	7 (4.0)	
Hispanic	2 (1.4)	3 (1.7)	
Marital Status			0.961
Single	23 (15.8)	27 (15.3)	
Married/Partnered	102 (69.9)	127 (72.2)	
Divorced/Separated	13 (8.9)	14 (8.0)	
Widowed	8 (5.5)	8 (4.5)	
Body Mass Index	28.20 (4.70)	28.64 (6.72)	0.505
Joint Space Width (mm)	0.61 (0.75)	1.28 (1.16)	**<0.001**
Osteoarthritis Grade			**<0.001**
None	0 (0.0)	5 (3.2)	
Mild	7 (5.0)	40 (25.8)	
Moderate	65 (46.1)	64 (41.3)	
Severe	69 (48.9)	46 (29.7)	
AUDIT	3.11 (3.71)	2.61 (2.71)	0.163
Aberdeen	21.25 (6.50)	19.06 (7.04)	**0.004**
HADS – Anxiety	5.13 (4.23)	5.15 (3.89)	0.970
HADS – Depression	4.99 (3.50)	4.39 (3.52)	0.126
Charnley Classification			0.421
A	59 (39.9)	76 (42.9)	
B	28 (18.9)	24 (13.6)	
C	61 (41.2)	77 (43.5)	
Pain VAS	6.07 (1.93)	5.16 (2.38)	**<0.001**
EQ VAS	72.48 (17.56)	73.54 (18.73)	0.603
EQ-5D Index	0.61 (0.19)	0.67 (0.20)	**0.004**
UCLA Activity	4.57 (2.16)	5.41 (2.44)	**0.001**
Harris Hip Score	52.57 (12.89)	60.90 (17.98)	**<0.001**

Strengths and Limitations

SHAR Studies

Utilization of national register data to conduct epidemiological studies is a great strength of the first four papers. With 100% coverage and a high level of completeness, the SHAR is an invaluable tool for understanding trends in THR patients in Sweden. The only opportunity for loss to follow-up is if a patient were to emigrate, otherwise data from patients who move from one region of Sweden to another will still be captured. The use of a national personal identification number is what allowed the linkage between the SHAR and other health registers and Statistics Sweden, another great strength of papers I through III. The addition of PROMs to the SHAR has added value to the surgical and demographic data collected therein. High response rates to the PROM program allow for analyses such as those in the first four papers to identify patient factors that could either protect them from or set them at a higher risk for poor outcomes after THR. With the inclusion of PROs, the patient's voice is considered and can influence how decision-making occurs and how treatment may be improved for at-risk patients. However, observational studies like those summarized here, can only provide information about trends, but cannot speak to causation. While some consider this to be the major limitation of observational studies, it could be argued that observational studies, and particularly national observational studies, play a very important role in identifying rare risk factors that might not be identified in smaller cohort studies and randomized controlled trials. With the identification of these risk factors, targeted studies may then be executed to understand the causation of the risk.

MGH Studies

The HJR only covers a select group of surgeons at a single institution making the generalizability of studies of its patients limited. However this targeted coverage does have its benefits allowing for the collection of level IV data (radiographs). Similarly, a more extensive PROM program can be implemented in the clinic due to agreement from all participating surgeons. Although paper VI does not have the benefit of nationwide data, it is strengthened by access to both radiographs and an extensive PROM protocol. An unfortunate limitation for paper VI was the highly incomplete education data. Given that education level had an important influence on outcomes after THR as indicated in paper II, it was unfortunate that this variable could not be included in the analysis of factors influencing whether surgery would be recommended and whether the patient would proceed. A strength of paper V was the two week separation of survey administrations and the crossover design of the study when validating the EQ-5D-5L survey unlike other EQ-5D-5L validation studies.[17,50,51,75,88] This separation was important to ensure that patient responses to one version of the survey would not influence the patient's response to the second version as occurred in other validation studies with the EQ-5D-5L survey.[49]

Discussion

These studies provide a better understanding of demographics and risk factors that may predispose OA patients for poor outcomes after THR surgery as well as identify accurate methods to assess those factors and outcomes. Identification or confirmation of new and existing factors associated with PROs is an important step in improving care with THR. Patients electively choose this procedure with the hope of eliminating pain and improving their mobility. Should either of these areas not improve, it is very likely that the patient will be dissatisfied with their treatment and their HRQoL will suffer.[4,24] It is important that referring clinicians and surgeons educate patients about their risks as well as set realistic expectations for outcomes from THR. Identifying patients with the risk of poor outcomes is crucial before surgical intervention, particularly when non-surgical options are available to improve these risk factors before surgery. Identification of patients who will most benefit from THR and educating those at risk for poorer outcomes may ultimately improve patient satisfaction and the cost utility of this treatment.

The Role of Register Studies

Because register studies cannot draw conclusions about causation, these works like all observational studies, act predominantly to discover patient factors associated with outcomes. Either smaller targeted cohort studies or randomized trials can further explore the identified factors. Our register studies do not seek to define how these patients should be treated with respect to surgical versus non-surgical interventions. Once again, targeted studies can identify whether at risk patients will have greater benefits from alternative, non-surgical treatment, or delayed surgical treatment. This work provides clinicians with greater understanding of how certain OA patients may do after THR. The results are not intended to deter surgical intervention in any at risk groups, but rather identify subpopulations that may benefit from additional education, closer investigation, and greater attention before and/or after surgery. Whether specific actions regarding these patients will improve outcomes requires exploration and their efficacy will need to be determined.

Protective and Risk Factors

Comorbidities

Because THR is elective, serious and life-threatening conditions are usually a contraindication, and therefore, the population eligible for the procedure tends to be healthy. Comorbidities are defined as conditions that are not associated with the development of the condition or diagnosis of interest.[71] Medical conditions besides OA measured with the ICD-10-based comorbidity measures had a marginal influence on reported THR outcomes, but did not prove to be valuable when predicting HRQoL, pain, satisfaction, or functional abilities after THR.[60] Because these measures were designed to monitor morbidity and rehospitalization, it is not surprising that they were not closely associated with the outcomes from healthy adults treated with THR.[13,23,90] Musculoskeletal comorbidities such as bilateral hip disease (Charnley class B) or inhibited walking due to other joint conditions (Charnley class C) did have a detrimental effect on PRO after THR. These additional joint problems are unlikely to be accounted for in standard measures of comorbidity because they are likely to also be caused by OA. When measuring pain, HRQoL, and satisfaction after THR, it goes to reason that additional untreated joints affected by OA would influence recovery and reported outcomes. Therefore, when predicting outcomes after THR, it is important to account for musculoskeletal comorbidities using Charnley classification.

Socioeconomic Status

Socioeconomic status is determined by three elements occupation, income, and education. In most cases, the highest level of education attained will influence an individual's occupation and therefore their income.[28] Once an individual's education is complete, usually early in life, it remains constant unlike occupation and income making education a useful measure of socioeconomic status particularly in health studies looking at individuals who may be retired or disabled and unable to work.[16,38] Individuals in Sweden with moderate or low levels of education tended to have worse HRQoL, greater pain, and less satisfaction after THR suggesting that individuals with lower socioeconomic status were at a greater risk of poor THR outcomes. The education system in Sweden is quite different from that of the United States and therefore, the influence of education on health may vary between these nations.

Unfortunately, I was unable to explore the influence of education on health status before THR in the American population due to missing education data in the majority of patient records. However, others have also found that low levels of education have a detrimental effect on THR outcomes.[9,26,96] Despite this, patients with low levels of education do benefit from THR and should not be excluded from consideration by clinicians. They may simply benefit from additional discussions about the importance of rehabilitation and postoperative expectations.

Domestic Support

Marital status was included in the analyses in papers II and VI as a measure of domestic support. Because rehabilitation from surgery is important for favorable outcomes, we were interested to know if having a domestic partner had an influence on the decision to have THR and on how the patient would do postoperatively. Marital status was among the lowest rated variables in predicting whether a patient would choose to have THR if it were recommended. When modeling PRO at 1 year, marital status only influenced the reporting of pain. These results may mean that patients have support systems beyond domestic partnerships or that support is not an important factor in deciding to have surgery or successful recovery. A third option is that marital status is simply too crude a variable to understand the complex nature of social support, and qualitative investigation of social support might be beneficial.

Mental Health

Depression prevalence has a wide range (3.2% to 19.8%) depending on the population investigated.[10,55,62,103] Because the prevalence of depression and/or anxiety in populations with OA has been measured as high as 41% is it important for clinicians treating individuals with OA to understand how depression or anxiety might affect their patients.[5,64,70,89] Depression is linked to reduced physical function, greater experience of pain, and overall impairment in HRQoL in addition to the well-known psychological and emotional symptoms.[5,18,85,89] Physical symptoms of depression are very similar to those of OA and could confound the diagnosis of depression in patients living with OA and vice versa.[85] In fact, OA patients suffering from physical symptoms are at a higher risk of developing the psychological elements of depression.[100] Also important for clinicians to understand is that OA patients suffering from depression have a greater sensitivity to pain and decreased ability to cope with their disease, and that depressed individuals who undergo surgery tend to report less satisfaction and pain reduction with treatment.[81,85,89]

Because antidepressants provide minimal improvement in mild to moderate depression for many patients, antidepressant usage may serve as a viable stand-in for depression in health studies.[27,44,56,72] Antidepressant medications serving as a surrogate for mental health status in the Swedish OA population with THR did in fact indicate that individuals prescribed these medications had worse HRQoL and pain before and after surgery and were less satisfied with the outcomes from THR. Because the indication for the medication was unknown in over a third of patients, it unfortunately was not possible to identify if a particular condition or another had a greater influence on these outcomes. When screening for anxiety and depression preoperatively at MGH, neither condition played a pivotal role in surgical recommendations or in the patients' decision to move forward with surgery. It will be important to follow the MGH patients postoperatively to determine the influence of anxiety and depression independently on THR outcomes. Lin and collaborators showed that arthritis patients receiving enhanced depression care incorporating antidepressants and/or psychotherapy sessions had marked improvements in pain, physical function, and HRQoL even without surgical intervention.[64] Thus indicating that in the case of some patients, treatment of depression may improve their ability to cope with their OA even without surgical intervention. A better understanding of which specific mental health conditions influence outcomes after THR will allow clinicians to better target non-operative and surgical treatment of OA.

Age and Gender

Age and gender are commonly considered important risk factors necessitating control in health study models. Because the frequency of OA increases as the population ages, it is important to account for a patient's age when considering treatment with THR.[19,82] Both papers II and III found that age had a significant influence on postoperative pain and satisfaction. When considering HRQoL, age was found to have a non-linear relationship, and there was a detrimental effect on patient-reported HRQoL once individuals entered their late sixties.[35] Interestingly, the population in paper II was on average five years younger (64 years) than those in paper III (69 years) due to missing education data for the older patients. The age distribution in paper II could explain why age was not an influential variable in the postoperative EQ-5D index and EQ VAS models, but was an included variable in paper III. Age is an important consideration when comparing THR populations between nations as well. The THR population investigated at MGH was on average even younger than the two Swedish populations (60 years). Patients in the US get joint replacement at an earlier age than those in Europe and tend to have better PRO after treatment, suggesting perhaps that European clinicians may wait too long to achieve optimal outcomes.[36]

Because THR is more common in women, gender is also important to consider when studying outcomes.[41,78] Gender is also known to influence functional status, HRQoL, and satisfaction before and after THR.[9,83,92] While women reported lower HRQoL before and after THR than men, the change in their scores was greater suggesting that perhaps women may benefit from THR at a younger age than men. The exception to this appears to be women who are Charnley class C and had the least amount of improvement in HRQoL than all their peers.[34] Gender only played a very small part in the probability of clinicians at MGH recommending THR surgery, but was actually among the top five variables influencing the probability that the patient would decide to have THR.

EQ-5D

When addressing general aspects of health, the EQ-5D survey and its EQ VAS should not only to be considered outcome measures, but rather the elements of this instrument can also serve as independently reliable indicators of a patient's health which are associated with the PROs. In clinical practice, the use of these measures preoperatively is valuable for clinicians as a means to understand a patient's self-reported health status in order to more accurately encourage realistic expectations for the outcomes from surgery, despite neither having as strong a predictive power as the Charnley classification as found in paper I.

Three-level survey

The positive attributes of the original EQ-5D, that it is short and validated in 160 languages, contribute to its wide use to assess HRQoL. The EQ-5D index and EQ VAS provide users with two measures of HRQoL. The index is a weighted value based on population norms and preferences allowing comparisons between patient groups or treatment types. The index can compare HRQoL between groups of patients who received THR and those who opted for non-surgical treatment, or between THR patients and total knee replacement patients for example. We are however limited to in-country comparisons because different values sets result in different index score ranges. The British value set, used in Sweden until the recent development of the Swedish value set and utilized in papers I through IV, ranges from -0.594 to 1.000. By contrast, the American value set ranges from -0.109 to 1.000, and the new Swedish value set ranges from 0.340 to 0.969 making comparisons across published works challenging, if not impossible. In order to compare populations, the same value set may be used to calculate index scores, but the user risks inaccurate weighting of the five dimensions for one or more of the populations. The EQ VAS is a patient-centered measure of HRQoL in that the patient is asked to rate their health that day on a scale of zero to 100.

Because the scores from the EQ-5D survey provide these two different measures of HRQoL, papers II and III tested both preoperative EQ-5D indices and EQ VAS scores with Bayesian model averaging to determine if either or both should be included in the linear regression models. When looking at education, preoperative EQ-5D index and EQ VAS scores were included in each regression model looking at all four outcomes, but when investigating antidepressant prescriptions, the preoperative EQ-5D index was not included in the postoperative pain VAS or satisfaction VAS models. The differences may be accounted for in the patient populations. While there is likely to be overlap between the populations in the two studies (Figure 1), paper II had a broader time interval for inclusion and a larger observed population. However, the posterior probabilities for the preoperative EQ-5D index in paper III were zero indicating no direct influence of this variable on the postoperative pain or satisfaction VAS scores, so the population differences are unlikely to explain the exclusion of the preoperative EQ-5D index in paper III. Responses to the final question of the EQ-5D survey, whether the patient experienced anxiety or depression, did influence patient-reported pain and satisfaction at one year. In this population, inclusion of the fifth question responses seems to have taken over the role of the EQ-5D index with respect to these two outcomes. Both the EQ-5D index and EQ VAS scores remained influential for reporting of both postoperative EQ-5D indices and EQ VAS scores however.

Five-level survey

Several groups have validated the new five-level version of the EQ-5D survey for both general and disease specific patient groups supporting broad adoption of this improved instrument.[46,50,51,75,88] Our findings were similar for the preoperative and postoperative THR populations suggesting that with five response choices, patients will be able to provide a more nuanced assessment of their HRQoL than with the original version. Responses to the five health dimensions in the EQ-5D-5L had either moderate or strong correlations with the EQ VAS whereas correlations with the responses in the EQ-5D-3L were weak or absent. This finding could mean that with the development of a weighted index value set, the EQ-5D-5L index may more closely correlate with EQ VAS scores indicating a sharper representation of HRQoL with the index than with the three-level survey.

New Tools for Clinicians

The ability to anticipate and mitigate poor outcomes after THR will benefit patient HRQoL directly. Additionally, it could decrease the overall societal cost of the treatment by delaying the procedure in some, encouraging THR earlier on others, or by minimizing clinically unnecessary follow-up appointments in at risk patients. With an arsenal of tools, the clinician can hone patient care preoperatively to target those who will most benefit from THR and provide additional education or non-surgical care for patients who may not benefit in the near-term. Large register studies like those outlined here, are an important step in identifying at risk patients like those with additional joint problems, low levels of education, or active antidepressant use.

Patient Education

Many patients suffering from arthritis assume that increased pain corresponds to additional injury. Therefore, due to fear of further damage both before and after THR, some patients may reduce their activity. Contrary to this assumption, activity avoidance may contribute to physical deterioration and muscle weakness making their condition worse and making it more difficult to recover.[45] Education is important for OA patients to understand how their activity or lack thereof will influence the severity of their condition. Poorer THR outcomes in certain patient groups like those identified in papers I through III may be due to inadequate participation in rehabilitation. Several factors may be the cause; fear, low levels of motivation, heightened pain sensitivity, less capability of coping with illness, inferior ability seek medical information, and/or lower or uncertain expectations of pain relief and functional recovery.[5,64,70,85,86,92] Education about expectations, continued activity, and access to resources should be considered not only for those opting for THR, but also for those choosing non-surgical treatment for OA as education in these areas may serve as preventative measures against the development of depressive symptoms as well, which are common in this population.[100]

Indications for THR

There is a lack of consensus about indications for THR and inconsistencies across nations in the timing of treatment with this procedure. THR is provided to a younger population in the United States than in Europe for example, and Gromov et al found that American patients also had less severe OA than those in Europe.[36] The ultimate goal of works like those summarized here is to identify as many risk factors for poor PRO as possible. Identified risk factors can then be implemented into regional and/or global universal standards for ideal THR indications. Development of a decision-making tool for referring clinicians and orthopaedic surgeons will not only standardize indications, but possibly identify the ideal interval at which patients will see the greatest improvement in their outcomes from THR.

At the national level, optimization of indication and timing may minimize clinically unnecessary follow-up appointments and possibly revision, thereby improving the cost utility of treatment of OA with THR and identifying ideal treatment paths for patients with this common condition.

Conclusions

A patient's musculoskeletal comorbidities, socioeconomic status, and mental health are all important demographic elements to understand when considering treatment of hip OA with THR. These factors will influence how a patient will report their HRQoL, pain, and satisfaction one year after surgery. Gender and age also play a role in predicting these outcomes, but marital status seems to have little influence.

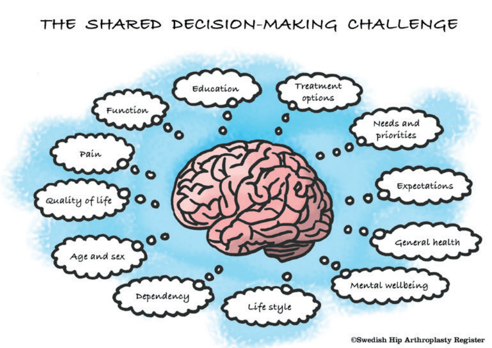
The EQ-5D survey is a brief measure of HRQoL making it appealing for large-scale studies. The newer version of this survey with five answer options appears to provide a better measure of HRQoL both before and after treatment with THR. While a weighted index for this version of the survey does not yet exist, the relationship between the preoperative score and the postoperative score will be as important to understand as in the three-level version of the survey. A two-line model with a single change point was ideal for describing the relationship between the pre- and postoperative scores in the three-level version of the EQ-5D index in the Swedish population. Prior to implementation, it will be important to establish if this holds true for the five-level version and for other populations. In Sweden, a two-line model with a single change point should be implemented when modeling the influence of additional patient factors on patient-reported HRQoL as measured by the EQ-5D-3L survey.

Because surgeons need to synthesize so many patient characteristics when deciding to recommend THR, an evidence based decision-making tool could streamline this process. Several characteristics proved valuable when estimating the probability that a patient would be recommended for THR surgery and whether they would proceed. In the arthroplasty clinic at MGH, the most critical was the space remaining in the joint as measured on an AP radiograph. Patient-reported assessment of the individual's health and hip condition were also important for both predictions, indicating that PROMs would be a valuable addition to a decision tool. Implementation of the developed probability prediction tool in the MGH arthroplasty clinic will ultimately determine if the tool might help aid the decision-making process for both clinicians and patients thereby improving treatment by streamlining patient assessment, selection, and recommendations.

Ongoing Projects

Decision-making Tool

Together these works provide information regarding patient characteristics that will indicate the need for THR, how to measure changes in the patient's condition after THR, and identify those who may be at a higher risk for poor outcomes after THR. Ideally, we will be able to make a tool with the capability of synthesizing these factors as a means to identify ideal THR candidates. The goal of this decision tool is three fold. It will identify the patients who have the greatest need for treatment with THR. It will help clinicians target patients who have the greatest capacity to improve after THR, which will in turn improve the effectiveness of this widely used treatment. Finally, more targeted recommendations for THR can ultimately improve the cost utility of the procedure by minimizing unnecessary healthcare utilization. It will be important to conduct either cohort studies or randomized trials to identify favorable alterative treatment options for those individuals who may not benefit from THR. In a collaborative effort between the SHAR and Harris Orthopaedic Laboratory, we aim to create this easy to use tool that will aid shared decision-making between the patient and either primary care or arthroplasty clinicians about whether to proceed with THR.



Outcomes at MGH

The new PROM protocol at MGH remains active, and therefore, data collection continues with postoperative collection of PROMs from the individuals included in paper VI as well as enrollment of any new patients at the clinic. Utilizing the same demographic, radiographic, and preoperative patient-reported data, a similar investigation as outlined in paper VI will be used to determine which patient factors are associated with PROs 1 year postoperatively and eventually at 3, 5, 7, and 10 years. The addition of outcome data to the prediction algorithm will not only provide an indication of who will be recommended for THR, but also will indicate the probability of improvement, no change, or degradation after treatment. After this analysis we will have a better understanding of which PROMs are helpful in predicting THR recommendations as well as outcomes after surgery and will then be able to hone the PROM protocol to minimize the burden on the patient.

Since completion of paper VI, I have been in contact with another group investigating how to improve decision-making in the arthroplasty clinic at MGH. Because education data in the charts is spotty, this group has been collecting education data directly from the patients. Through collaboration with this group, I can incorporate this important measure of socioeconomic status into our models. Inclusion of education may identify if socioeconomic status influences a patient's decision to have THR and whether it is associated with their outcomes in this US population.

Implementation of EQ-5D-5L

A similar protocol as that implemented in paper V will be utilized at centers in the western region of Sweden asking preoperative and postoperative patients to complete both versions of the EQ-5D survey. Ethical committee approval was obtained and patient recruitment will continue until Spring of 2016. The goal of the investigation is the same of that in paper V, to validate the new EQ-5D-5L survey this time in a Swedish population.

Considerations are currently being made as to when the arthroplasty clinic at MGH will change their standard PROM protocol to include the five-level version of the EQ-5D survey in place of the three-level version. The absence of a validated index for the five-level version makes it necessary for the clinicians to agree to use the crosswalk from the three-level survey index to the five-level version.

Future Visions

Streamlining Questionnaire Burden

Successful patient-reported data collection is dependent upon straightforward survey administration with minimal burden on the patient or clinical office. The National Institutes of Health (NIH) have developed the Patient-reported Outcome Measurement Information System (PROMIS) as a means to standardize patient-reported assessment in clinical studies funded by the NIH. Some of the PROMIS measures employ computer adaptive testing (CAT) as a means to streamline the burden on the patient. CAT instruments provide the user with a bank of questions of which only a few are administered. The system draws from the pool of items based on the response to the previous item. CAT systems require a substantial initial effort to identify the ideal questions for the item bank through item response theory. Once established however, a CAT system can provide a survey that requires the patient to answer very few questions by eliminating items that do not apply to the respondent. Implementation of CAT systems in national arthroplasty registers is a theoretically desirable way to expand the use of PRO within additional national registers because the burden on the patient is minimized. However, several administrative challenges remain both for registers currently using PROMs and those who are not. Funding would be needed to conduct the initial validation studies for the CAT system or to pay for the use of an established system, dedicated staff for data management and analysis would still be required, and comparisons to previously collected PROM data may become problematic. Organizations such a the International Society of Arthroplasty Registers or the American Academy of Orthopaedic Surgeons could perhaps facilitate development of a PRO CAT system, but then language and cultural differences would need to be accounted for. Despite these challenges, a system for collecting standardized THR PROs that asks for minimal effort from the patient is desirable.

Access to THR

An important problem being discussed particularly in the United States, but in other nations as well is access to healthcare. The severity of OA symptoms can be correlated with socioeconomic status, and those with low socioeconomic status have a greater unmet need for arthroplasty.[16,38] Their willingness to receive this treatment is equal to that of their peers indicating that patient inclination does not account for the discrepancy.[42] Individuals with lower socioeconomic status tend to have low-income jobs that are more labor intensive therefore increasing their chances for development of OA.[16,53] The disparity between OA patients with low socioeconomic status and access to THR exists even in countries with seemingly equal access to healthcare.[21,22,42] It is important to understand how to provide access to THR for all who may benefit from this treatment, and although these individuals may be at risk for lower PROs, they still have marked improvements. The ability for these individuals to return or continue to work may even have a greater influence on HRQoL further down the line.

PROMs for Assessing Follow-up Needs

It has been discussed to use PROMs as a measure of whether a patient needs to be seen for a follow-up visit in the clinic. Materials development for THR has reached a plateau. While component manufacturers continue to develop new implant materials, a collection of safe, very well performing components with excellent 10-year *in vivo* track records already exist.[31] If new materials are used, regular follow-up should be encouraged to ensure their safety. If well-established, vetted materials are used, follow-up in the clinic with radiographs may not be necessary for the majority of patients. Any patient having problems or those who desire time with the clinician would still have that option, but because THR tends to be highly successful, many patients do not want to return for follow-up. With the increase in THRs administered, the pool of individuals requiring follow-up is increasing. Also with the increased longevity of THR components, younger patients are receiving this treatment. For some high volume surgeons, regular in-clinic follow-up is unmanageable for all patients.

By asking the patient to complete a set of surveys either on paper forms in the mail or through a web-based link sent through email as a precursor to an in-person clinic visit, PROMs can be used to screen patient needs. If an individual was having no problems and had favorable PROs, they might not be asked to come see the doctor. This system would provide the clinician with information regarding the patient's wellbeing with minimal effort. Whereas in the current system in the United States, if a patient does not return for follow-up, the clinician may not know whether the patient was doing so well that they did not want to return for follow-up, or whether they were very unhappy or having problems and went to see a different doctor.

The growing cost of healthcare in the United States necessitates systematic changes in how care is provided. By eliminating unnecessary office visits and clinical radiographs, the cost of treatment with THR can be minimized not to mention the burden on the patient. In this scenario, ideal PROMs would need to be identified and thresholds for those surveys would need to be established to identify when a patient might benefit from a return visit to the clinic.

Shared Decision-making

In order for shared decision-making to be successful, a large burden is placed on the healthcare professional. Accurate tools to help guide the patient to the most beneficial decision with an acceptable level of risk may exist, but the process of engaging the patient in the decision process takes skill, training, and willingness on the part of the clinician. Consultation techniques for shared decision-making have been developed. However, exploration of such techniques is limited in orthopaedics.[93] Additionally, many of these programs were designed to reduce rates of elective procedures and will likely increase the required time spent with the patient in the clinic.[104] For these reasons among others, adoption of dedicated shared decision-making programs in arthroplasty has been slow. Because the historic method of paternalistic consultation in arthroplasty conflicts with shared decision-making philosophies, bridging this divide becomes important. Orthopaedic surgeons are known for early adoption of new technologies and therefore, are likely to adopt a new electronic tool for decision-making rather than learn specialized techniques for patient engagement. How best to instigate participation in shared decision-making processes by orthopaedic surgeons needs to be explored.

Summary in English

This collection of works aimed to identify patient factors that could help predict how a patient would self-report their health-related quality of life, pain, and satisfaction with the outcomes from surgery one year after total hip replacement. Among the investigated variables were age, gender, Charnley classification, education level, marital status, antidepressant use, and the presence of comorbid conditions. Through these studies, the works also aimed to identify ideal methodologies for considering patient-reported outcomes and the influence of patient factors on them.

Linear regression was utilized in most studies to identify the predictive capacity of the aforementioned patient factors on the patient-reported outcomes at one year after total hip replacement. A new version of one of the patient-reported outcome measures, the EQ-5D, was investigated and found to have an improved ability to measure health-related quality of life in total hip replacement patients both before and after treatment. Finally, several patient-reported measures, demographics, and radiographic assessments were synthesized using a complex algorithm to identify the probability that a patient would be recommended for surgery and whether they would move forward with the treatment.

Overall, patients improved after total hip replacement. Those who started with worse patient-reported health-related quality of life and pain tended to improve similar amounts as those going into surgery with better scores; however, they never achieved absolute scores that were as high. Individuals with greater musculoskeletal comorbidities, taking antidepressants, with low or medium levels of education tended to begin and end with worse patient-reported scores.

Because these patients can be easily identified through preoperative surveys and their medical record, clinicians can quickly identify patients with whom they should more thoroughly discuss the risks for lower postoperative scores after total hip replacement. While these findings indicate who may benefit the most from total hip replacement, they are not intended to discourage treatment, but rather provide the clinician and the patient with more information for the decision-making process with respect to treatment of their hip condition.

Summary in Swedish

Målsättningen med det här avhandlingsarbetet var att identifiera patientrelaterade omständigheter och förhållanden som kan hjälpa till att förutsäga hur en patient kommer att uppleva sin hälsorelaterade livskvalitet, smärta och tillfredsställelse med resultatet av kirurgin ett år efter en höftprotesoperation. De undersökta variablerna var ålder, kön, utbildningsgrad, civilstatus, medicinering med antidepressiva läkemedel, förekomsten av andra sjukdomar och huruvida patienten har besvär med gångförmågan på grund av annan sjuklighet än höftledssjukdom. Genom de här undersökningarna syftade avhandlingsarbetet också till att identifiera lämpliga metoder för att bedöma patientrapporterat utfall och hur patientfaktorer påverkar utfallet.

Linjär regression användes i de flesta studierna för att förstå sambandet mellan de ovan nämnda patientfaktorerna och det patientrapporterade utfallet ett år efter höftprotesoperation. En ny version av frågeformuläret EQ-5D undersöktes och visade sig ha bättre förmåga att mäta hälsorelaterad livskvalitet hos höftprotesopererade patienter både före och efter operationen. I det sista delarbetet användes en rad patientrapporterade mått jämte demografiska variabler och artrosgrad mätt på röntgenbilder för att skapa en algoritm som räknar ut en sannolikhet för att en patient skulle rekommenderas att genomgå höftprotesoperation och i så fall om patienten skulle välja att gå vidare med åtgärden.

På det hela taget förbättrades patienterna efter höftprotesoperation. De som hade sämre hälsorelaterad livskvalitet och högre smärta före operationen tenderade att förbättras lika mycket som de som hade bättre utgångsvärden. De med lägre utgångsvärden förbättrades emellertid inte till samma nivå som de med högre. Individer med andra sjukdomar som påverkar gångförmågan, de som använder antidepressiva läkemedel samt låg- eller medelutbildade tenderade att ha sämre värden på de patientrapporterade utfallsmåtten såväl före som efter operationen.

Eftersom dessa patienter enkelt kan urskiljas med hjälp av frågeformulär, sjukhistoria och patientens journal kan läkaren snabbt identifiera dem som kan vara i behov av mer utförlig diskussion om risken att inte förbättras eller bli missnöjd med resultatet. Även om dessa resultat kan indikera vem som kan ha störst nytta av en höftprotesoperation är inte avsikten att avskräcka någon från operation utan istället ge läkaren och patienten mer information som underlag till det gemensamma beslutet om lämplig behandling av höftledssjukdomen.

Project Collaborators

Supervisors

Göran Garellick MD, PhD – Register Director of the Swedish Hip Arthroplasty Register and professor at the Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg

Ola Rolfson MD, PhD – Orthopaedic surgeon with the Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg and affiliated with the Swedish Hip Arthroplasty Register

Henrik Malchau MD, PhD – Professor at Harvard Medical School and Co-director of the Harris Orthopaedic Laboratory at Massachusetts General Hospital, Boston, MA, USA affiliated with the Swedish Hip Arthroplasty Register and an orthopaedic surgeon with the Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg

Collaborators

Szilárd Nemes PhD – Biostatistician with the Swedish Hip Arthroplasty Register, Gothenburg, Sweden

Kevin Rader PhD – Biostatistician/preceptor in the Statistics Department at Harvard University, Cambridge, MA, USA

Max Gordon MD, PhD – Orthopaedic surgeon with the Department of Orthopaedics, Institute of Clinical Sciences, Karolinska Institute, Stockholm, Sweden

Kristina Annerbrink MD – Psychiatrist in the Department of Psychiatry at Sahlgrenska University Hospital, Gothenburg, Sweden

Kirill Gromov MD, PhD – Orthopaedic surgeon at Department of Orthopedics, Copenhagen University Hospital Hvidovre, Copenhagen, Denmark

Andrew Freiberg MD – Chief of Arthroplasty in the Department of Orthopedics, Massachusetts General Hospital and affiliated with the Harris Orthopaedic Laboratory in Boston, MA, USA

Acknowledgements

Numerous people have supported and encouraged me throughout this work. I would like to express my deepest appreciation;

Göran Garellick, for the opportunity to work with the register, guiding me through the intricacies of observational studies, and sharing the history and landscape of Sweden with me;

Ola Rolfson, for all our thoughtful discussions, caring so much about improving care for patients, teaching me a new perspective on orthopaedic research, your friendship, and for your willingness to dance;

Henrik Malchau, for believing in me, encouraging and facilitating my professional growth, and ensuring that I sample the local drink wherever we are;

Szilárd Nemes, my unofficial fourth supervisor, for your patience, wonderful teaching skills, and willingness to help in any way you could;

Kevin Rader, for making sure I understood all the tests, your willingness to help, and for all the delicious food;

Max Gordon, for all your hard work merging the databases and teaching me about splines;

Kristina Annerbrink, for joining our project halfway through and sharing your clinical expertise;

Kirill Gromov, for your willingness to read hundreds of radiographs for me even after your fellowship was complete;

Andrew Freiberg, for your study design ideas, indispensable clinical perspective, and your encouragement;

Harry Rubash, for your support of my projects at MGH and your valuable suggestions when designing them;

Sandra Pierre, for your hard work in the clinic capturing all the PROMs at MGH and for enrolling the patients for the EQ-5D study,

Linda Johansson, for helping me navigate the intricate university system,

Kajsa Erikson, for your organizational prowess, for always making sure I have somewhere to stay in Gothenburg, and for all those lovely dinners,

William H Harris, for your interest in my work, your astute suggestions, and your encouragement;

Friends in the Harris Orthopaedic Lab, past and present, for your nonstop support and laughter;

Johanna Rolfson, Karin Garellick, and *Inger Malchau*, for your warm hospitality whenever I was in Sweden. You made me feel welcome and at home on those long stays;

Nanna Sillesen, for being my professional-woman-in-orthopaedics guru and for your wonderful friendship;

Peter and Candy Cross, for all your kind supportive words and help whenever I needed it;

Arthur and Beverly Greene and Vanessa Fernandez Greene, for endless encouragement and instilling in me a sense of commitment to the task at hand, but also knowing when it is time to have fun;

Aurelia Skye, for all your peaceful naps and amazing smiles;

Eben Cross, for our adventures amidst all the work and reminding each other what is truly important. The symbiosis.

References

1. U.S. Valuation of the EuroQol EQ-5D™ Health States. Edited, Rockville, MD, Agency for Healthcare Research and Quality, 2005.

2. Allen Butler, R.; Rosenzweig, S.; Myers, L.; and Barrack, R. L.: The Frank Stinchfield Award: the impact of socioeconomic factors on outcome after THA: a prospective, randomized study. *Clin Orthop Relat Res,* 469(2): 339-47, 2011.

3. Amstutz, H. C.; Thomas, B. J.; Jinnah, R.; Kim, W.; Grogan, T.; and Yale, C.: Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am,* 66(2): 228-41, 1984.

4. Anakwe, R. E.; Jenkins, P. J.; and Moran, M.: Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. *J Arthroplasty,* 26(2): 209-13, 2011.

5. Axford, J.; Butt, A.; Heron, C.; Hammond, J.; Morgan, J.; Alavi, A.; Bolton, J.; and Bland, M.: Prevalence of anxiety and depression in osteoarthritis: use of the Hospital Anxiety and Depression Scale as a screening tool. *Clin Rheumatol,* 29(11): 1277-83, 2010.

6. Barr, C. J.; Barbalace, R. J.; Wessinger, S. J.; Bragdon, C. R.; Kwon, Y. M.; and Malchau, H.: Validation of a hospital-based joint registry: quantification of errors and maximizing utility. *J Arthroplasty,* 27(10): 1766-71, 2012.

7. Beaule, P. E.; Dorey, F. J.; Hoke, R.; Le Duff, M.; and Amstutz, H. C.: The value of patient activity level in the outcome of total hip arthroplasty. *J Arthroplasty,* 21(4): 547-52, 2006.

8. Bharmal, M., and Thomas, J., 3rd: Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. *Value Health,* 9(4): 262-71, 2006.

9. Bischoff-Ferrari, H. A.; Lingard, E. A.; Losina, E.; Baron, J. A.; Roos, E. M.; Phillips, C. B.; Mahomed, N. N.; Barrett, J.; and Katz, J. N.: Psychosocial and geriatric correlates of functional status after total hip replacement. *Arthritis Rheum,* 51(5): 829-35, 2004.

10. Blazer, D. G.; Kessler, R. C.; McGonagle, K. A.; and Swartz, M. S.: The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry,* 151(7): 979-86, 1994.

11. Brazier, J.; Roberts, J.; Tsuchiya, A.; and Busschbach, J.: A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ,* 13(9): 873-84, 2004.

12. Brozek, J. L.; Guyatt, G. H.; and Schunemann, H. J.: How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. *Health Qual Life Outcomes,* 4: 69, 2006.

13. Charlson, M. E.; Pompei, P.; Ales, K. L.; and MacKenzie, C. R.: A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis,* 40(5): 373-83, 1987.

14. Charnley, J.: The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *J Bone Joint Surg Br,* 54(1): 61-76, 1972.

15. Clement, N. D.; Muzammil, A.; Macdonald, D.; Howie, C. R.; and Biant, L. C.: Socioeconomic status affects the early outcome of total hip replacement. *J Bone Joint Surg Br,* 93(4): 464-9, 2011.

16. Cleveland, R. J.; Schwartz, T. A.; Prizer, L. P.; Randolph, R.; Schoster, B.; Renner, J. B.; Jordan, J. M.; and Callahan, L. F.: Associations of educational attainment, occupation and community poverty with hip osteoarthritis. *Arthritis Care Res (Hoboken)*, 2012.

17. Craig, B. M.; Pickard, A. S.; and Lubetkin, E. I.: Health problems are more common, but less severe when measured using newer EQ-5D versions. *J Clin Epidemiol,* 67(1): 93-9, 2014.

18. Creed, F.; Morgan, R.; Fiddler, M.; Marshall, S.; Guthrie, E.; and House, A.: Depression and anxiety impair health-related quality of life and are associated with increased costs in general medical inpatients. *Psychosomatics,* 43(4): 302-9, 2002.

19. Danielsson, L., and Lindberg, H.: Prevalence of coxarthrosis in an urban population during four decades. *Clin Orthop Relat Res,* (342): 106-10, 1997.

20. Devlin, N. J., and Krabbe, P. F.: The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ,* 14 Suppl 1: S1-3, 2013.

21. Dixon, T.; Shaw, M.; Ebrahim, S.; and Dieppe, P.: Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need. *Ann Rheum Dis,* 63(7): 825-30, 2004.

22. Dixon, T.; Urquhart, D. M.; Berry, P.; Bhatia, K.; Wang, Y.; Graves, S.; and Cicuttini, F. M.: Variation in rates of hip and knee joint replacement in Australia based on socio-economic status, geographical locality, birthplace and indigenous status. *ANZ J Surg,* 81(1-2): 26-31, 2011.

23. Elixhauser, A.; Steiner, C.; Harris, D. R.; and Coffey, R. M.: Comorbidity measures for use with administrative data. *Med Care,* 36(1): 8-27, 1998.

24. Espehaug, B.; Havelin, L. I.; Engesaeter, L. B.; Langeland, N.; and Vollset, S. E.: Patient satisfaction and function after primary and revision total hip replacement. *Clin Orthop Relat Res,* (351): 135-48, 1998.

25. EuroQolGroup: EuroQol – a new facility for the measurement of health related quality of life. *Health Policy,* 16-3, 1990.

26. Fortin, P. R. et al.: Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. *Arthritis Rheum,* 42(8): 1722-8, 1999.

27. Fournier, J. C.; DeRubeis, R. J.; Hollon, S. D.; Dimidjian, S.; Amsterdam, J. D.; Shelton, R. C.; and Fawcett, J.: Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA,* 303(1): 47-53, 2010.

28. Fowler, F. J.: Improving Survey Questions: Design and Evaluation. Edited, 190, Sage Publications, 1995.

29. Garellick, G.; Karrholm, J.; Rogmark, C.; Rolfson, O.; and Herberts, P.: Swedish Hip Arthroplasty Register Annual Report 2011. Edited, 2011.

30. Garellick, G.; Malchau, H.; and Herberts, P.: Specific or general health outcome measures in the evaluation of total hip replacement. A comparison between the Harris hip score and the Nottingham Health Profile. *J Bone Joint Surg Br,* 80(4): 600-6, 1998.

31. Garellick, G.; Rogmark, C.; Karrholm, J.; and Rolfson, O.: Swedish Hip Arthroplasty Register Annual Report 2012. Edited, 2012.

32. Geerts, W. H.; Pineo, G. F.; Heit, J. A.; Bergqvist, D.; Lassen, M. R.; Colwell, C. W.; and Ray, J. G.: Prevention of venous thromboembolism: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest,* 126(3 Suppl): 338S-400S, 2004.

33. Genell, A.; Nemes, S.; Steineck, G.; and Dickman, P. W.: Model selection in medical research: a simulation study comparing Bayesian model averaging and stepwise regression. *BMC Med Res Methodol,* 10: 108, 2010.

34. Gordon, M.; Frumento, P.; Skoldenberg, O.; Greene, M.; Garellick, G.; and Rolfson, O.: Women in Charnley class C fail to improve in mobility to a higher degree after total hip replacement. *Acta Orthop,* 85(4): 335-41, 2014.

35. Gordon, M.; Greene, M.; Frumento, P.; Rolfson, O.; Garellick, G.; and Stark, A.: Age- and health-related quality of life after total hip replacement. *Acta Orthop,* 85(3): 244-9, 2014.

36. Gromov, K.; Greene, M. E.; Sillesen, N. H.; Troelsen, A.; Malchau, H.; Huddleston, J. I.; Emerson, R.; Garcia-Cimbrelo, E.; and Gebuhr, P.: Regional differences between US and Europe in radiological osteoarthritis and self assessed quality of life in patients undergoing total hip arthroplasty surgery. *J Arthroplasty,* 29(11): 2078-83, 2014.

37. Gunther, O. H.; Roick, C.; Angermeyer, M. C.; and Konig, H. H.: The responsiveness of EQ-5D utility scores in patients with depression: A comparison with instruments measuring quality of life, psychopathology and social functioning. *J Affect Disord,* 105(1-3): 81-91, 2008.

38. Hannan, M. T.; Anderson, J. J.; Pincus, T.; and Felson, D. T.: Educational attainment and osteoarthritis: differential associations with radiographic changes and symptom reporting. *J Clin Epidemiol,* 45(2): 139-47, 1992.

39. Hansen, V. J.; Greene, M. E.; Bragdon, M. A.; Nebergall, A. K.; Barr, C. J.; Huddleston, J. I.; Bragdon, C. R.; and Malchau, H.: Registries collecting level-I through IV Data: institutional and multicenter use: AAOS exhibit selection. *J Bone Joint Surg Am,* 96(18): e160, 2014.

40. Harris, W. H.: 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,* 51(4): 737-55, 1969.

41. Hawker, G. A.; Wright, J. G.; Coyte, P. C.; Williams, J. I.; Harvey, B.; Glazier, R.; and Badley, E. M.: Differences between men and women in the rate of use of hip and knee arthroplasty. *N Engl J Med,* 342(14): 1016-22, 2000.

42. Hawker, G. A.; Wright, J. G.; Glazier, R. H.; Coyte, P. C.; Harvey, B.; Williams, J. I.; and Badley, E. M.: The effect of education and income on need and willingness to undergo total joint arthroplasty. *Arthritis Rheum,* 46(12): 3331-9, 2002.

43. Herdman, M.; Gudex, C.; Lloyd, A.; Janssen, M.; Kind, P.; Parkin, D.; Bonsel, G.; and Badia, X.: Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res,* 20(10): 1727-36, 2011.

44. Hermens, M. L.; van Hout, H. P.; Terluin, B.; Ader, H. J.; Penninx, B. W.; van Marwijk, H. W.; Bosmans, J. E.; van Dyck, R.; and de Haan, M.: Clinical effectiveness of usual care with or without antidepressant medication for primary care patients with minor or mild-major depression: a randomized equivalence trial. *BMC Med,* 5: 36, 2007.

45. Heuts, P. H.; Vlaeyen, J. W.; Roelofs, J.; de Bie, R. A.; Aretz, K.; van Weel, C.; and van Schayck, O. C.: Pain-related fear and daily functioning in patients with osteoarthritis. *Pain,* 110(1-2): 228-35, 2004.

46. Hinz, A.; Kohlmann, T.; Stobel-Richter, Y.; Zenger, M.; and Brahler, E.: The quality of life questionnaire EQ-5D-5L: psychometric properties and normative values for the general German population. *Qual Life Res,* 23(2): 443-7, 2014.

47. Hoeting, J. A.; Madigan, D.; Raftery, A. E.; and Volinsky, C. T.: Bayesian Model Averaging: A Tutorial. *Statistical Science,* 14(4): 382-417, 1999.

48. ISAR: International Society of Arthroplasty Registries Bylaws. Edited, 2013.

49. Janssen, M. F.; Birnie, E.; Haagsma, J. A.; and Bonsel, G. J.: Comparing the standard EQ-5D three-level system with a five-level version. *Value Health,* 11(2): 275-84, 2008.

50. Janssen, M. F.; Pickard, A. S.; Golicki, D.; Gudex, C.; Niewada, M.; Scalone, L.; Swinburn, P.; and Busschbach, J.: Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res,* 22(7): 1717-27, 2013.

51. Jia, Y. X. et al.: Comparison between the EQ-5D-5L and the EQ-5D-3L in patients with hepatitis B. *Qual Life Res*, 2014.

52. Johnson, J. A., and Pickard, A. S.: Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta, Canada. *Med Care,* 38(1): 115-21, 2000.

53. Juhakoski, R.; Malmivaara, A.; Lakka, T. A.; Tenhonen, S.; Hannila, M. L.; and Arokoski, J. P.: Determinants of pain and functioning in hip osteoarthritis – a two-year prospective study. *Clin Rehabil,* 27(3): 281-7, 2013.

54. Kass, R. E., and Raftery, A. E.: Bayes Factors. *J. Amer Statist Assoc,* 90: 773-795, 1995.

55. Kessler, R. C.; Berglund, P.; Demler, O.; Jin, R.; Koretz, D.; Merikangas, K. R.; Rush, A. J.; Walters, E. E.; and Wang, P. S.: The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA,* 289(23): 3095-105, 2003.

56. Khan, A.; Leventhal, R. M.; Khan, S. R.; and Brown, W. A.: Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *J Clin Psychopharmacol,* 22(1): 40-5, 2002.

57. King, M. T.: A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res,* 11(2): 171-84, 2011.

58. Kopec, J. A., and Willison, K. D.: A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol,* 56(4): 317-25, 2003.

59. Larson, C. B.: Rating scale for hip disabilities. *Clin Orthop Relat Res,* 31: 85-93, 1963.

60. Lavernia, C. J.; Laoruengthana, A.; Contreras, J. S.; and Rossi, M. D.: All-Patient Refined Diagnosis-Related Groups in Primary Arthroplasty. *Journal of Arthroplasty,* 24(6): 19-23, 2009.

61. Learmonth, I. D.; Young, C.; and Rorabeck, C.: The operation of the century: total hip replacement. *Lancet,* 370(9597): 1508-19, 2007.

62. Li, X.; Sundquist, J.; and Sundquist, K.: Age-specific familial risks of depression: a nation-wide epidemiological study from Sweden. *J Psychiatr Res,* 42(10): 808-14, 2008.

63. Liang, M. H.: Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. *Med Care,* 38(9 Suppl): II84-90, 2000.

64. Lin, E. H. et al.: Effect of improving depression care on pain and functional outcomes among older adults with arthritis: a randomized controlled trial. *JAMA,* 290(18): 2428-9, 2003.

65. Luo, N.; Johnson, J. A.; Shaw, J. W.; Feeny, D.; and Coons, S. J.: Self-reported health status of the general adult U.S. population as assessed by the EQ-5D and Health Utilities Index. *Med Care,* 43(11): 1078-86, 2005.

66. Mahomed, N. N.; Arndt, D. C.; McGrory, B. J.; and Harris, W. H.: The Harris hip score: comparison of patient self-report with surgeon assessment. *J Arthroplasty,* 16(5): 575-80, 2001.

67. Mahomed, N. N.; Liang, M. H.; Cook, E. F.; Daltroy, L. H.; Fortin, P. R.; Fossel, A. H.; and Katz, J. N.: The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol,* 29(6): 1273-9, 2002.

68. Malchau, H.: On the Importance of Stepwise Introduction of New Hip Implant Technology. In *Department of Orthopaedics*. Edited, Sweden, Gothenburg University, 1995.

69. Mancuso, C. A.; Jout, J.; Salvati, E. A.; and Sculco, T. P.: Fulfillment of patients' expectations for total hip arthroplasty. *J Bone Joint Surg Am,* 91(9): 2073-8, 2009.

70. Marks, R.: Comorbid depression and anxiety impact hip osteoarthritis disability. *Disabil Health J,* 2(1): 27-35, 2009.

71. Mnatzaganian, G.; Ryan, P.; and Hiller, J. E.: Does Co-morbidity Provide Significant Improvement on Age Adjustment when Predicting Medical Outcomes? *Methods of Information in Medicine,* 53(2): 115-120, 2014.

72. Moncrieff, J., and Kirsch, I.: Efficacy of antidepressants in adults. *BMJ,* 331(7509): 155-7, 2005.

73. Naal, F. D.; Impellizzeri, F. M.; and Leunig, M.: Which is the best activity rating scale for patients undergoing total joint arthroplasty? *Clin Orthop Relat Res,* 467(4): 958-65, 2009.

74. Patrick, D. L., and Deyo, R. A.: Generic and disease-specific measures in assessing health status and quality of life. *Med Care,* 27(3 Suppl): S217-32, 1989.

75. Pickard, A. S.; De Leon, M. C.; Kohlmann, T.; Cella, D.; and Rosenbloom, S.: Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. *Med Care,* 45(3): 259-63, 2007.

76. Pollard, B.; Dixon, D.; Dieppe, P.; and Johnston, M.: Measuring the ICF components of impairment, activity limitation and participation restriction: an item analysis using classical test theory and item response theory. *Health Qual Life Outcomes,* 7: 41, 2009.

77. Pollard, B.; Johnston, M.; and Dieppe, P.: Exploring the relationships between International Classification of Functioning, Disability and Health (ICF) constructs of Impairment, Activity Limitation and Participation Restriction in people with osteoarthritis prior to joint replacement. *BMC Musculoskelet Disord,* 12: 97, 2011.

78. Quintana, J. M.; Arostegui, I.; Escobar, A.; Azkarate, J.; Goenaga, J. I.; and Lafuente, I.: Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. *Arch Intern Med,* 168(14): 1576-84, 2008.

79. Raftery, A. E.: Bayesian model selection in social research. *Sociological Methodology 1995, Vol 25,* 25: 111-163, 1995.

80. Research, US Department of Health and Human Services FDA Centers for Drug Evaluation and Research, Biologics Evaluation and Research, and Devices and Radiological Health: Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health and Quality of Life Outcomes,* 4: 79, 2006.

81. Riediger, W.; Doering, S.; and Krismer, M.: Depression and somatisation influence the outcome of total hip replacement. *Int Orthop,* 34(1): 13-8, 2010.

82. Rolfson, O.; Dahlberg, L. E.; Nilsson, J. A.; Malchau, H.; and Garellick, G.: Variables determining outcome in total hip replacement surgery. *J Bone Joint Surg Br,* 91(2): 157-61, 2009.

83. Rolfson, O.; Karrholm, J.; Dahlberg, L. E.; and Garellick, G.: Patient-reported outcomes in the Swedish Hip Arthroplasty Register: results of a nationwide prospective observational study. *J Bone Joint Surg Br,* 93(7): 867-75, 2011.

84. Rolfson, O.; Salomonsson, R.; Dahlberg, L. E.; and Garellick, G.: Internet-based follow-up questionnaire for measuring patient-reported outcome after total hip replacement surgery-reliability and response rate. *Value Health,* 14(2): 316-21, 2011.

85. Rosemann, T.; Backenstrass, M.; Joest, K.; Rosemann, A.; Szecsenyi, J.; and Laux, G.: Predictors of depression in a sample of 1,021 primary care patients with osteoarthritis. *Arthritis Rheum,* 57(3): 415-22, 2007.

86. Ross, C. E., and Wu, C.-L.: The Links Between Education and Health. *Am Sociol Rev,* 60(5): 719-45, 1995.

87. Saunders, J. B.; Aasland, O. G.; Babor, T. F.; de la Fuente, J. R.; and Grant, M.: Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction,* 88(6): 791-804, 1993.

88. Scalone, L.; Ciampichini, R.; Fagiuoli, S.; Gardini, I.; Fusco, F.; Gaeta, L.; Del Prete, A.; Cesana, G.; and Mantovani, L. G.: Comparing the performance of the standard EQ-5D 3L with the new version EQ-5D 5L in patients with chronic hepatic diseases. *Qual Life Res,* 22(7): 1707-16, 2013.

89. Scopaz, K. A.; Piva, S. R.; Wisniewski, S.; and Fitzgerald, G. K.: Relationships of fear, anxiety, and depression with physical function in patients with knee osteoarthritis. *Arch Phys Med Rehabil,* 90(11): 1866-73, 2009.

90. Sharabiani, M. T. A.; Aylin, P.; and Bottle, A.: Systematic Review of Comorbidity Indices for Administrative Data. *Medical Care,* 50(12): 1109-1118, 2012.

91. Shaw, J. W.; Johnson, J. A.; and Coons, S. J.: US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care,* 43(3): 203-20, 2005.

92. Singh, J. A., and Lewallen, D.: Age, gender, obesity, and depression are associated with patient-related pain and function outcome after revision total hip arthroplasty. *Clin Rheumatol,* 28(12): 1419-30, 2009.

93. Slover, J.; Shue, J.; and Koenig, K.: Shared decision-making in orthopaedic surgery. *Clin Orthop Relat Res*, 470(4): 1046-53, 2012.

94. Soderman, P., and Malchau, H.: Is the Harris hip score system useful to study the outcome of total hip replacement? *Clin Orthop Relat Res,* (384): 189-97, 2001.

95. Soderman, P.; Malchau, H.; and Herberts, P.: Outcome of total hip replacement: a comparison of different measurement methods. *Clin Orthop Relat Res,* (390): 163-72, 2001.

96. Stevens, M.; Wagenmakers, R.; Groothoff, J. W.; Bulstra, S. K.; van den Akker-Scheek, I.; and Zijlstra, W.: Physical activity behavior after total hip arthroplasty (THA): a prediction based on patient characteristics. *Patient Educ Couns,* 69(1-3): 196-9, 2007.

97. Streiner, D. L., and Norman, G. R.: Health Measurement Scales: a practical guide to their development and use. Edited, Oxford University Press, 2008.

98. Sullivan, P. W.; Lawrence, W. F.; and Ghushchyan, V.: A national catalog of preference-based scores for chronic conditions in the United States. *Med Care,* 43(7): 736-49, 2005.

99. Tönnis, D.: Congenital Dysplasia and Dislocation of the Hip in Children and Adults. Edited, Berlin, Germany, Spinger-Verlag, 1987.

100. Tylee, A., and Gandhi, P.: The importance of somatic symptoms in depression in primary care. *Prim Care Companion J Clin Psychiatry,* 7(4): 167-76, 2005.

101. van Hout, B. et al.: Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health,* 15(5): 708-15, 2012.

102. Wamper, K. E.; Sierevelt, I. N.; Poolman, R. W.; Bhandari, M.; and Haverkamp, D.: The Harris hip score: Do ceiling effects limit its usefulness in orthopedics? *Acta Orthop,* 81(6): 703-7, 2010.

103. Wilhelm, K.; Mitchell, P.; Slade, T.; Brownhill, S.; and Andrews, G.: Prevalence and correlates of DSM-IV major depression in an Australian national survey. *J Affect Disord,* 75(2): 155-62, 2003.

104. Youm, J.; Chan, V.; Belkora, J.; and Bozic, K. J.: Impact of Socioeconomic Factors on Informed Decision Making and Treatment Choice in Patients With Hip and Knee OA. *J Arthroplasty*, 2014.

105. Zahiri, C. A.; Schmalzried, T. P.; Szuszczewicz, E. S.; and Amstutz, H. C.: Assessing activity in joint replacement patients. *J Arthroplasty,* 13(8): 890-5, 1998.

106. Zigmond, A. S., and Snaith, R. P.: The hospital anxiety and depression scale. *Acta Psychiatr Scand,* 67(6): 361-70, 1983.

SWEDISH HIP ARTHROPLASTY REGISTER

Harris Orthopaedic Laboratory · Established 1969

HOL 1969