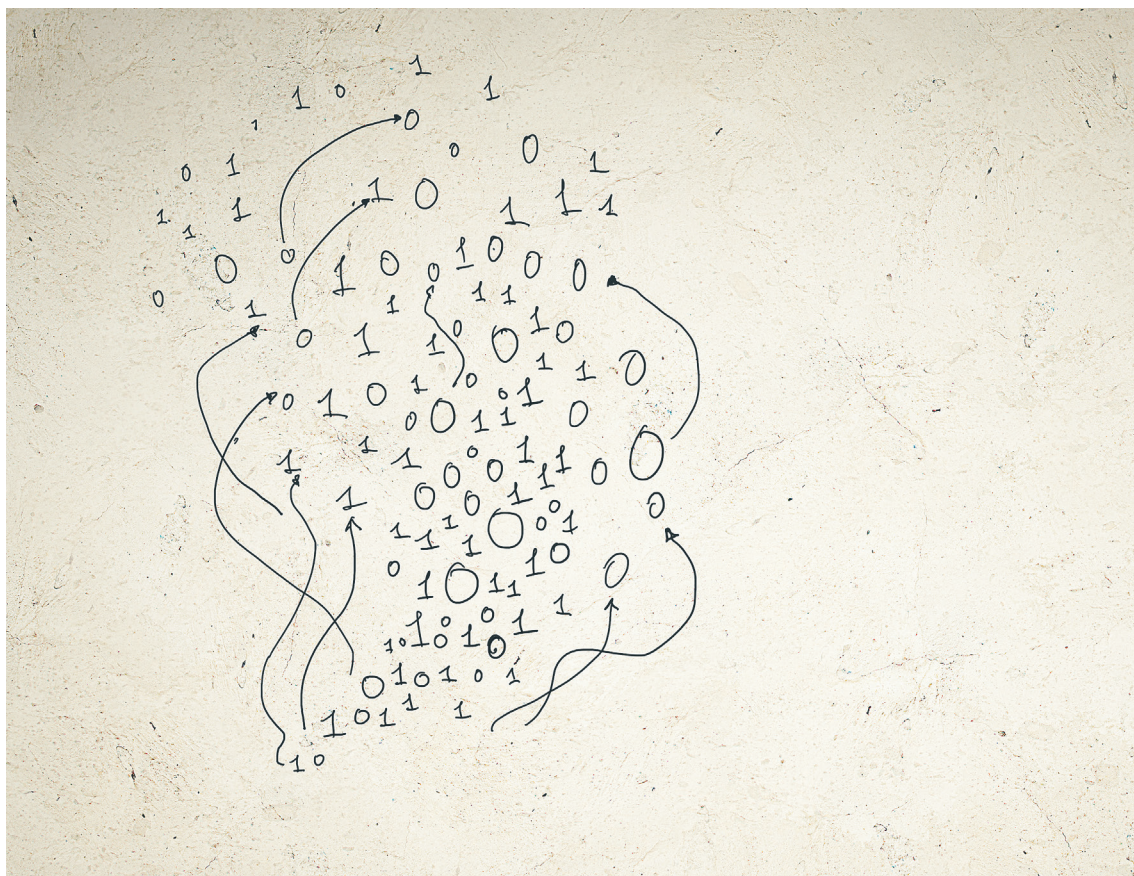


Evaluation of patient-related factors influencing outcomes after total hip replacement



Max Gordon



**Karolinska
Institutet**

From DEPARTMENT OF CLINICAL SCIENCES,
DANDERYD HOSPITAL
Karolinska Institutet, Stockholm, Sweden

EVALUATION OF PATIENT-RELATED FACTORS INFLUENCING OUTCOMES AFTER TOTAL HIP REPLACEMENT

Max Gordon



**Karolinska
Institutet**

In association with
the Swedish Hip Arthroplasty Register

Stockholm 2014

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.

Printed by Elanders AB

© Max Gordon, 2014

ISBN 978-91-7549-526-2

To my family

Evaluation of patient-related factors influencing outcomes after
total hip replacement
THESIS FOR DOCTORAL DEGREE (Ph.D.)

by

Max Gordon

Principal Supervisor:

Olof Sköldenberg, MD, PhD
Karolinska Institutet
Department of Clinical Sciences,
Danderyd Hospital
Division of orthopaedics

Opponent:

Prof. Karl Michaëlsson
Uppsala University
Uppsala Clinical Research Center
Division of orthopaedics

Co-supervisors:

Prof. André Stark
Karolinska Institutet
Department of Clinical Sciences,
Danderyd Hospital
Division of orthopaedics

Prof. Göran Garellick
Sahlgrenska Academy at
the University of Gothenburg
Institute of Clinical Sciences
Division of orthopaedics

Prof. Johan Kärrholm
Sahlgrenska Academy at
the University of Gothenburg
Institute of Clinical Sciences
Division of orthopaedics

Examination Board:

Prof. Li Felländer-Tsai
Karolinska Institutet
Department of Clinical Science Intervention
and Technology
Division of orthopaedics

Assoc. prof. Fredrik Granath
Karolinska Institutet
Department of Clinical Epidemiology Unit

Prof. Olle Svensson
Umeå Universitet
Department of Surgical and
Perioperative Sciences
Division of orthopaedics

ABSTRACT

Background

Basic patient factors such as age, sex, and comorbidities are poorly understood in the context of re-operation rates and patient-reported outcome measures after total hip replacement (THR) surgery.

Aims

The aims of this thesis were to investigate:

1. If comorbidity measures developed for mortality outcomes also are applicable for re-operations.
2. How age influences health-related quality of life (HRQoL).
3. If sex/age interacts with self-administered Charnley classification in regards to HRQoL.
4. The generalizability of HRQoL estimates.
5. If antidepressant use impacts patient-reported outcomes.

Methods

All studies were based upon the Swedish Hip Arthroplasty Register (SHAR) database. For study I, data from SHAR was cross-matched with the National Patient Register through which 3 different comorbidity measures were calculated: the Charlson score, the Elixhauser score, and the Royal College of Surgeons (RCS) Charlson's score. The three scores were then compared using survival analysis with implant re-operation performed between 0 to 2 years and 2 to 12 years.

In study II and III we used the SHAR's PROM database with HRQoL outcomes as measured by EQ-5D and EQ-VAS. In study II we modeled age using linear regression in combination with restricted cubic splines in order to study the relationship between age and HRQoL. In study III we used linear regression with interaction terms evaluated by ANOVA-tests, subset, and EQ-5D dimension specific analyses.

In study IV we linked the SHAR's PROM database to the National Patient Register and a Danish cross-sectional sample. The Charlson comorbidity measure was calculated as in study I, and effect modification by country was investigated through terms of interaction, evaluated as in study III.

In study V, we cross-matched the SHAR's PROM database with the Prescribed Drug Register. We calculated the usage of antidepressants using regular expressions. Measures for compliance, treatment change, and indication were retrieved from the prescription text.

Results

Study I

0-2 years, only the Elixhauser score showed significant risk increase with increased score for both 1-2 and ≥ 3 comorbidities. The predictive C-statistic in this period for the Elixhauser score was poor, 0.52. None of the measures proved to be of any value between 2-12 years.

Study II

Both the EQ-5D index and EQ VAS exhibited a non-linear relationship with age, they were fairly unaffected by age until the patient's late sixties, after which it had a negative impact.

Study III

We found that women in category C had a poorer EQ-5D outcome compared to men. This effect was mostly due to the fact that women failed to improve in the mobility dimension, only 40% improved, while 50% of men improved. Age did not interact with Charnley class. We also found that the classification performed best without splitting or aggregating classes.

Study IV

Danish patients had an overall higher EQ-5D index and EQ VAS than Swedish patients (p-value < 0.001). After regression analysis, the estimated coefficients for sex, age, or the Charlson score did not differ between countries for either the EQ-5D index (p-value = 0.83) or EQ VAS (p-value = 0.41) one year after THR.

Study V

Antidepressants were used by 9% of the cases (n = 954). Patients using antidepressants had poorer HRQoL, more pain, and experienced less satisfaction. Preoperative antidepressant use was independent of patient-reported anxiety/depression in predicting PROs one year after THR. Discontinuation of treatment was negatively associated with pain and satisfaction at one year.

Conclusions

Study I

We failed to validate any of the scores for re-operations after total hip arthroplasties, although the Elixhauser score may be useful for estimating the comorbidities relevant to the risk of re-operation within 2 years. The comorbidity associated risk increase was small, and is undoubtedly best suited to the study of large samples and not individual patients.

Study II

There is a non-linear relationship for age and HRQoL in patients receiving THR; resulting in residual confounding if treated as a simple linear term or categorically in the regression. The implication of this is important, as age is a common confounder. The same applies to the preoperative EQ-5D index and EQ VAS.

Study III

The self-administered Charnley classification is a reliable instrument with interesting properties easy to utilize in everyday clinical practice. There is also strong evidence that women in Charnley class C fail to improve their mobility as much as men.

Study IV

There are clear similarities in how basic predictors influence patient-reported outcomes in patients with THR in Sweden and Denmark. Apparent cultural, social, and other such differences among these countries are not reflected in these predictors.

Study V

Antidepressants have a negative influence on patient-reported outcomes 1 year after THR, independent of the pre-operative EQ-5D anxiety/depression dimension. We also found that discontinuation of treatment prior to surgery is associated with poorer outcomes in the dimensions of pain and satisfaction.

LIST OF SCIENTIFIC PAPERS

- I. M. Gordon, A. Stark, O. G. Sköldenberg, J. Kärrholm, and G. Garellick, “The influence of comorbidity scores on re-operations following primary total hip replacement Comparison and validation of three comorbidity measures,” *Bone Joint J*, vol. 95-B, no. 9, pp. 1184–1191, Sep. 2013.
- II. M. Gordon, M. Greene, P. Frumento, O. Rolfson, G. Garellick, and A. Stark, “Age and health related quality of life after total hip replacement: Decreasing gains in patients above 70 years of age,” *Acta Orthop*, vol. 85, p. ??-??, 2014.
- III. M. Gordon, P. Frumento, O. Sköldenberg, M. Greene, G. Garellick, and O. Rolfson, “Women in Charnley class C fail to a higher degree to improve mobility after total hip replacement: a nationwide registry study on Charnley class and health related quality of life” *Acta Orthop*, vol. 85, p.??-??, 2014.
- IV. M. Gordon, A. Paulsen, S. Overgaard, G. Garellick, A. B. Pedersen, and O. Rolfson, “Factors influencing health-related quality of life after total hip replacement - a comparison of data from the Swedish and Danish hip arthroplasty registers,” *BMC Musculoskeletal Disorders*, vol. 14, no. 1, p. 316, Nov. 2013.
- V. M. E. Greene, O. Rolfson, M. Gordon, K. Annerbrink, H. Malchau, and G. Garellick, “Does the Use of Antidepressants Predict Patient-Reported Outcomes Following Total Hip Replacement Surgery?,” *Manuscript*, 2014.

TABLE OF CONTENTS

1	Introduction.....	1
1.1	Outcomes.....	1
1.1.1	Patient survival.....	1
1.1.2	Implant Survival.....	1
1.1.3	Adverse events.....	3
1.1.4	Patient-reported outcome.....	3
1.2	Patient factors.....	5
1.2.1	Age.....	5
1.2.2	Sex.....	6
1.2.3	Comorbidities.....	8
1.2.3.1	Indices and scores.....	9
1.2.3.2	Charnley class.....	10
1.2.3.3	Mental health.....	11
1.2.3.4	BMI and obesity.....	11
1.2.3.5	Diabetes.....	12
1.2.3.6	Other diseases.....	13
1.3	Summary.....	14
2	Aims.....	15
3	Methods.....	16
3.1	Study design.....	16
3.2	Data sources.....	16
3.2.1	The Swedish Hip Arthroplasty Register.....	16
3.2.1.1	Patient-reported outcome measures.....	16
3.2.2	The Swedish National Patient Register.....	17
3.2.3	The Prescribed Drug Register.....	17
3.2.4	The Danish study sample.....	17
3.3	Patient selection.....	17
3.3.1	Study I.....	17
3.3.2	Study II.....	17
3.3.3	Study III.....	18
3.3.4	Study IV.....	18
3.3.5	Study V.....	18
3.4	Outcomes.....	18
3.4.1	Study I.....	18
3.4.2	Study II to IV.....	19
3.4.3	Study V.....	19
3.5	Exposure variables.....	19
3.5.1	Study I.....	19
3.5.2	Study II.....	20
3.5.3	Study III.....	20
3.5.4	Study IV.....	20
3.5.5	Study V.....	20
3.6	Statistics.....	22
3.6.1	Survival analysis.....	22
3.6.1.1	Cox proportional hazards background.....	22
3.6.1.2	Competing risk regression.....	22
3.6.1.3	Proportional hazards assumption.....	23
3.6.1.4	C-statistic.....	23

3.6.2	Linear regression.....	23
3.6.2.1	Interactions.....	24
3.6.2.2	Heteroskedasticity	24
3.6.2.3	Residuals	25
3.6.3	Quantile regression	25
3.6.4	Non-linearity.....	25
3.6.5	Software.....	26
4	Results.....	27
4.1	Study I.....	27
4.1.1	Early re-operations.....	27
4.1.2	Late re-operations	27
4.1.3	Other analyses.....	30
4.1.3.1	Sex	30
4.1.3.2	Age.....	30
4.2	Study II.....	31
4.2.1	Non-linearity of age	32
4.2.2	Pre-operative HRQoL	33
4.3	Study III	34
4.3.1	EuroQoL dimensions	36
4.3.2	Grouping or subdividing the classification	36
4.3.3	Self-reporting.....	36
4.4	Study IV	38
4.5	Study V.....	39
4.5.1	Antidepressant effect.....	39
4.5.2	EQ-5D anxiety/depression and antidepressants.....	39
4.5.3	Sensitivity analyses	41
5	Discussion.....	43
5.1	Study I.....	43
5.1.1	Discussion of results	43
5.1.2	Strengths	43
5.1.3	Limitations.....	44
5.2	Study II.....	45
5.2.1	Discussion of results	45
5.2.2	Strengths	46
5.2.3	Limitations.....	46
5.3	Study III	47
5.3.1	Discussion of results	47
5.3.2	Strengths	48
5.3.3	Limitations.....	49
5.4	Study IV	49
5.4.1	Discussion of results	49
5.4.2	Strengths	49
5.4.3	Limitations.....	50
5.5	Study V.....	50
5.5.1	Discussion of results	50
5.5.2	Strengths	51
5.5.3	Limitations.....	51
5.6	General Discussion	52
5.6.1	Data quality	52
5.6.2	Code quality.....	53

5.6.3	Patient selection.....	54
5.6.4	Minimal clinically important difference	54
5.6.4.1	Distribution based methods	55
5.6.4.2	The anchor based method	56
5.6.4.3	Risk and MCID	57
5.6.4.4	Changing management or estimating base level.....	57
5.7	Balancing complexity	58
6	Conclusions	59
6.1	Study I	59
6.2	Study II.....	59
6.3	Study III.....	59
6.4	Study IV	59
6.5	Study V.....	59
7	Future perspectives	60
8	Acknowledgements.....	61
9	Sammanfattning på svenska	63
10	References	66

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
ANOVA	Analysis of variance
ASA	The American Society of Anesthesiologists (ASA) Physical Status classification system
BIC	Bayesian information criterion
BMI	Body Mass Index (kg/m^2)
CI	Confidence interval
DHR	The Danish Hip Arthroplasty Register
EQ-5D	Euroqol 5 dimensions
EQ VAS	Euroqol VAS (mm)
GRoC	Global rating of change
HRQoL	Health-related Quality of Life
ICD	International Statistical Classification of Diseases
MCID	Minimal clinically important difference
MPR	Medication possession ratio
NPR	The national patient register
PDR	Prescribed Drug Register
PROM	Patient-Reported Outcome Measure
RCS Charlson's score	Royal College of Surgeons Charlson score/index
SD	Standard deviation
SF-36/12/8	The Short Form (36/12/8) Health Survey
SHAR	The Swedish Hip Arthroplasty Register
THR	Total Hip Replacement
TTO	Time-trade-off
VAS	Visual Analogue Scale (mm)
WOMAC	Western Ontario and McMaster Universities Arthritis Index

LIST OF RISK MEASURES

RISK MEASURES

Incidence proportion or cumulative incidence The proportion of a population that *develops* a condition/disease during a given time period

$$\text{Incidence prop.} = \frac{\text{No. new cases}}{\text{Total no. of subjects}}$$

Incidence rate The number of subjects that develop a condition/disease in relation to the *total exposed time*, e.g. if two subjects are studied for 2 years, the total exposed time accumulates to 4 years

$$\text{Incidence rate} = \frac{\text{No. new cases}}{\text{Total no. of observed days/months/years}}$$

Odds The relation between those with a condition (diseased) and those without a condition (healthy)

$$\text{Odds} = \frac{\text{No. diseased}}{\text{No. healthy}} = \frac{\text{No. diseased}}{\text{Total no.} - \text{No. diseased}}$$

Prevalence proportion The proportion of a population *with* a condition/disease at a given time point

$$\text{Prevalence prop.} = \frac{\text{No. individuals with a condition}}{\text{Total no. of subjects}}$$

RATIOS

Hazard Ratio The relation between two hazards of experiencing an event (see section 3.6.1 for details)

Odds ratio The ratio between two odds is similar to the concept of relative risk, although this is restricted to rare diseases. It is commonly used due to the popularity of logistic regression. If manually estimated the formula is:

$$\text{Odds ratio} = \frac{\text{Odds for exposed}}{\text{Odds for controls}}$$

Rate ratio The ratio between two incidence rates

$$\text{Rate ratio} = \frac{\text{Incidence rate for exposed}}{\text{Incidence rate for controls}}$$

Relative risk or risk ratio The ratio between two incidence proportions

$$\text{Relative risk} = \frac{\text{Incidence proportion for exposed}}{\text{Incidence proportion for controls}}$$

1 INTRODUCTION

Prior to total hip replacement (THR) surgery, patients with hip osteoarthritis were offered treatments such as resection arthroplasty, weight distribution, arthrodesis and even bed rest with hip traction [1], [2]. When THR emerged as a viable treatment alternative it revolutionized orthopaedics, referred to by some as “*the operation of the century*” [3]. The first successes of THRs appeared in the late 1930s and 30 years later, with Charnley’s ground-breaking work on low-friction arthroplasties [4], it truly caught on. The indications have since then widened; today even 100 year old patients are being operated [5].

1.1 OUTCOMES

The success of a THR is measured by 4 main outcome categories: (1) patient survival, (2) implant survival, (3) other adverse events, and (4) patient-reported outcome measures (PROM). Although technical outcomes such as component positioning, soft tissue preservation or leg length restoration can be viewed as a type of outcome, they all aim at addressing issues within the 4 main outcomes.

1.1.1 Patient survival

The patient surviving the surgery is the most fundamental premise for any kind of surgery. The mortality associated with the THR procedures occurs soon after the surgery; effects related to bleeding, pulmonary embolism, and other per-operative causes quickly decline during the first months [6]. The 90-day mortality rate is the most common mortality measure [7], although some studies report in-hospital mortality [8] or 30-day mortality [9].

Prior to the aseptic technique introduced in late 19th century, mortality rates of 50% were common [10]. Even today there seems to be improvement in the mortality rates [11]–[13], and the current 90-day mortality in Sweden is less than 1% [14], similar to rates in other countries [6], [8], [11], [12], [15]–[19].

1.1.2 Implant Survival

Implant survival relates to the survival of the implanted material. There are two terms frequently used for estimating implant survival, re-operations and revisions:

- **Re-operation** is a wider definition whereby any further surgery affecting the operated hip is included.

- **Revision** is a subgroup of re-operations including only the surgeries where the implant is exchanged.

The Swedish Hip Arthroplasty Register (SHAR) has from its start in 1979 used re-operations as a measurement [20]. Re-operations and revisions can be further subdivided into the underlying cause:

- **Infection:** Often referred to as periprosthetic joint infection, it occurs most frequently within the first year of surgery [21], [22] with an estimated 2 year cumulative incidence of 1% [23].
- **Fracture:** Periprosthetic fractures occur both early and late after THR. Early re-operations within 6 months are frequently technical complications and are associated with cementless stems [24]. The average time to fracture is 7 years, and the estimated cumulative incidence about 1% for 5 years, and 3-4% at 10 years [25], [26].
- **Instability:** A THR dislocates more easily than the innate hip joint. The incidence of dislocation is difficult to measure and ranges in international studies between 1 and 5% [27]–[30]. Between 2009 and 2012, 0.3% in Sweden were re-operated within 2 years due to dislocation [31, p. 52]. Most dislocations occur early on, 50%-70% occur within the first 3 postoperative months [32], [33], although late dislocations (> 5 years) more often require re-operation [34].
- **Loosening/lysis:** Loosening occurs when the bone retracts from the implant with a radiolucent line visible on x-rays. If the process is localized and the implant is clearly not loose, it is referred to as osteolysis [35]. Common inflammatory pathophysiology for the processes has been hypothesized [36], [37]. The cumulative incidence for loosening/osteolysis is 0.2% at 2 years, 0.8% at 5 years, and 2% at 10 years (study I).
- **Other:** There are multiple other causes such as technical failure or implant fracture for re-operations that amount to a cumulative 10 year incidence of less than 3% (study I).

The 10 year re-operation rate is currently at an all-time low, less than 5% for patients operated due to primary osteoarthritis [14]. This coincides with the new

having increasing difficulty in improving on the existing designs [38], currently 3 stems constitute 99% of all implanted cemented stems in Sweden [31, p. 13].

1.1.3 Adverse events

Death and re-operations are only a subset of adverse events after THR. Other events such as cardiovascular, thromboembolic events can be equally important, while minor events such as urinary tract infection may be quality of care indicators. Adverse events in medical practice are fairly common, expected around 3-15% [39]–[43], and perhaps even more common in orthopaedics [44]–[46]. There is currently no validated registry-based tool for the Swedish population, although readmissions combined with ICD-codes are reported on a yearly basis in Sweden [14].

1.1.4 Patient-reported outcome

The THR indication is primarily to improve quality of life, and is very successful at this [3]. Charnley reported in his landmark paper on long-term results in 1972 that the majority of patients had excellent results regarding pain and walking ability [47]. As the indications have widened and patient demographics changed, the importance of patient-reported outcome measures (PROMs) have reentered center stage [48].

Furthermore, with the improvement in implant survival and low re-operation, PROMs are becoming increasingly important for evaluating THR [49]. Although hip resurfacing has fallen out of favor, many of its main proponents argued for it because of better PROMs [50]–[52].

In orthopaedics there are two main types of PROMs: generic measures and disease specific measures. Generic measures usually span different health domains, and can be used irrespective of the particular disease being studied; this has the advantage of allowing comparison between different diseases [53]. Generic measures are frequently also referred to as health related quality of life (HRQoL). There are 2 main HRQoL instruments used in orthopaedics:

- **Euroqol 5 dimensions (EQ-5D):** The EQ-5D™ was developed in the late 80s by the EuroQol group as a standardised, non-disease-specific instrument for describing and evaluating health states [54]. The tool consists of 5 different dimensions and a visual analogue scale (VAS). The dimensions are weighted according to a value set in an index where values of 0 correspond to states equal to death and 1 to perfect health. The 5 dimensions are:
 1. Mobility
 2. Self care

3. Usual activities
 4. Pain
 5. Depression
- **Short Form 36 (SF-36):** The SF-36® has its roots in the 70s but was formalized in the late 80s [55], and is now managed by the Medical Outcomes Trust. The tool consists of 36 items that are aggregated into 8 different domains that in turn can be divided into physical and mental health summaries. The domains and summaries are usually expressed on a scale ranging from 0 to 100 where higher scores indicate better health for that domain. Simpler alternatives to the 36 questions have been developed by QualityMetric, SF-12® [56] and SF-8™ [57]. The 8 domains¹ are:
 1. Bodily pain
 2. Physical functioning
 3. Role limitations due to physical health
 4. General health
 5. Mental health
 6. Vitality
 7. Social functioning
 8. Role limitations due to emotional health.

Disease specific quality of life measures and hip specific outcome measures are separated by some into two entities [58]. As they are both limited in their extent, they are grouped together under disease specific measures in this thesis. There are currently 3 commonly used disease specific measures for hip osteoarthritis and THR:

- **Harris hip score:** The score was originally introduced for evaluating acetabular fractures [59, p. 69], but has since been widened to other hip related diseases [60], [61]. Harris hip score was originally not intended for self-reporting, but studies have shown that this is feasible [62]. The score consists of 10 items that are merged into a score ranging from 0 to 100 where higher scores indicate better hip function.
- **Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC):** The WOMAC® score was developed in the 80s and consists of 24 items that are summarised into 3 subscales [63]–[65]. The subscores range from 0 to 100 where higher scores indicate poorer hip function. The 3 subscales are:
 1. Pain
 2. Stiffness

¹ 1-4 are part of the physical health summary, 5-8 are part of the mental health summary

3. Physical function

- **Oxford hip score:** The score consists of 12 items add up to a score from 12 to 60 where higher scores indicate poorer hip function [66]. The scoring was updated by the original authors in 2007; the new score ranges from 0 to 48 and has flipped the scale, i.e. higher number indicates better hip function [67].

This is just a handful of the available disease specific hip scores [58], [68] and there is an ongoing effort in validating and translating new measures [69], [70]. PROMs are generally believed to be less susceptible to interviewer's bias as the majority of them are self-reported, since patients answer the questionnaire without any aid from study personnel. The lack of supervision can increase the risk of misunderstood questions and skipped answers [71, p. 500]. Recall bias can still occur and it is therefore important that the health states are prospectively recorded [72]. There is also a risk of response bias, i.e. that the patient responds to the survey differently from those non-responding. This stresses the importance of maximizing response rates [73]–[75].

PROMs are also believed to be more susceptible than other outcomes to cultural influence. Dieppe et al. reported large variations between and within countries in disease severity when deciding to operate [76]. How these variations affect the PROMs is sparsely explored.

1.2 PATIENT FACTORS

This thesis is limited to the results regarding patients operated with THR due to primary osteoarthritis. In 2012 almost 16,000 patients were operated with a primary THR and primary osteoarthritis amounted to more than 80% of all surgeries [14]. Other important indications are fractures, idiopathic femoral head necrosis, childhood disease and rheumatoid arthritis. Fractures have become increasingly common [77], [78] while rheumatoid arthritis has rapidly decreased due to efficient disease-modifying anti-rheumatic agents [79].

1.2.1 Age

Age is an important confounder in studying most health related outcomes [80, p. 5], and perhaps the most commonly adjusted variable. Increased re-operation rates for younger patients due to aseptic loosening is today widely recognized [31, p. 79], [81]–[86]. On the other hand, the risk of fractures and dislocations is higher among elderly. Cook et al. showed that patients age 80 and above have an odds ratio of 4.4 (95% confidence interval (CI), 2.9–6.4) for fractures compared to patients below 80 [26]. Malkani et al. reported that patients 85 and above were

associated with an almost 50% higher relative risk of early dislocation compared to patients between 65 and 69 [87].

As expected, age is also related to mortality within 90 days. SooHoo et al. reported an odds ratio of 2.6 (95% CI, 2.2 to 3.0) for patients over the age of 75 compared to patients between 65 and 75 [88]. Hunt et al. studied over 400,000 patients and similarly reported an increase in the elderly, although they noted that this increase was primarily among men. The Kaplan-Meier 90-day mortality estimates for patients age 80 and above were 1.9% for men (95% CI, 1.7 to 2.11) vs 1.1% for women (95% CI, 1.0 to 1.2); the corresponding rates for patients between 65 and 69 were 0.3% for men (95% CI, 0.3 to 0.4) vs 0.2% for women (95% CI, 0.1 to 0.2) [12].

Regarding the impact of age on HRQoL after THR, there is currently no consensus [89]. Some studies report decreasing improvement in HRQoL after THR with age [90]–[94], while others indicate little effect if any [95]–[99]. Furthermore, it has not been properly investigated if the effect of age is similar throughout the range of ages, or if there is an accelerated decline for the elderly.

1.2.2 Sex

Men compared to women have a marginally higher risk for revisions [81], [86], [100], [101, pp. 40, 49] where the risk increases with age [31, p. 79]. The Norwegian Arthroplasty Register noted a hazard ratio of 1.3 (95% CI, 1.1 to 1.5) [102], while the Australian Orthopaedic Association National Joint Replacement Registry reported in 2013 an even smaller difference, hazard ratio 1.06 (95% CI, 1.02 to 1.1) [85], and Inacio et al. reported on an American cohort of 34,140 women having an increased risk for revision, hazard ratio 1.3 (95% CI, 1.1 to 1.5) [103]. Part of these differences could be explained by sex interactions with implant type [104].

Part of the increased risk may be due to infections [105]. Men are associated with increased hazard ratios in Nordic countries ranging between 1.5 and 2.5 for periprosthetic joint infection [22], [106], [107], although two large American studies fail to confirm these findings [103], [108]. Conversely, some report a higher risk for aseptic loosening for women with hazard ratios of 1.3 to 1.4 [103], [109], although this has not been unanimously shown [82]. Women also seem to dislocate 2 times more often than men [27], [29] though Conroy et al. could not find any support for this in an Australian cohort of 65,992 patients [110].

Men have marginally elevated mortality rates [12], [17]. SooHoo et al. reported a hazard ratio of 1.2 (95% CI, 1.1 to 1.4) for mortality within 90 days [88]. Memtsoudis et al. reported an odds ratio of 1.7 (95% CI, 1.6 to 1.9) for in-

hospital mortality [8]. McMinn et al. studied the long term mortality and found similarly higher mortality rates for men, hazard ratio 1.5 (95% CI, 1.5 to 1.6) [111].

Sex differences have also been noted for PROMs, although the differences are inconclusive. Slover et al. found for 1,357 THR patients that men reported a 4 points higher improvement for Harris hip score [112], less than 10% of the average improvement after a THR (45-50 points) [113], [114]. Smith et al. similarly found when studying 1,318 THR patients that men perform better, although they only investigated whether a patient reached the maximum Harris hip score of 100 [115].

Hajet et al. could not find any sex differences for the Oxford hip score of 5,038 patients but the study was compromised by the fact that more than half of the patients failed to respond to the survey [116]. Quintana et al. also failed to find any significant differences in 590 patients 6 months after THR for the WOMAC score [98]:

- 14 points for men vs 16 points for women¹ regarding pain
- 20 points for men vs 23 points for women¹ regarding stiffness
- 22 points for men vs 25 points for women¹ regarding physical function

Somewhat surprisingly they found significant differences favoring men when using the SF-36 tool:

- 51 points for men vs 40 points for women² regarding physical functioning
- 32 points for men vs 16 points for women² regarding role limitations due to physical health
- 56 points for men vs 47 points for women² regarding bodily pain
- 34 points for men vs 30 points for women² regarding general health
- 40 points for men vs 36 points for women² regarding mental component summary scale¹.

¹ The value is the improvement according to linear regression for healthy patients below 70 years of age without the mental component summary scale estimate, i.e. $\text{intercept} + \beta_{\text{sex}}$

² The value is the improvement according to linear regression for healthy patients below 70 years of age without the mental component summary scale estimate, i.e. $\text{intercept} + \beta_{\text{sex}}$

Judge et al. also looked at the SF-36 physical functioning dimension of 282 patients and found that women have a lower probability of achieving clinically meaningful improvements, odds ratio 0.4 (95% CI, 0.2 to 0.7) [117]. McGuigan et al. noted that men have slightly better outcomes for the SF-36 dimensions: bodily pain, role limitations due to physical health, mental health, vitality, and social functioning [90]. Two more recent studies with mixed hip and knee arthroplasty cohorts support these findings [96], [118].

Contrasting with this is the large study by Stevens et al. on 653 patients that also included BMI. They found that men reported poorer WOMAC outcomes, while no relation was seen for SF-36 [119]. Lavernia et al. investigated sex differences for 532 patients at a minimum of 2 years after their THR using 5 different scores; the WOMAC subscales were the only outcomes that correlated to a higher improvement among women but unfortunately they did not attempt to analyze their data further using any of the regression methods [120].

It is possible that the differences arise due to different control for confounding, although if sex had a stronger effect on patient-reported outcomes after THRs, the results would most likely have been more consistent.

1.2.3 Comorbidities

Comorbidities' influence on outcome after arthroplasties has lately gained a lot of interest. The number of articles in PubMed mentioning both arthroplasty and comorbidity has more than quadrupled between 2001 and 2011 (Figure 1). A comorbidity can be defined as [121]:

“...diseases or medical conditions unrelated in etiology or causality to the principal diagnosis that coexist with the disease of interest”.

In the THR context this interprets as any disease other than the one behind the reason for the THR.

¹ A summary of the four different mental health components (vitality, social function, role emotional, and mental health)

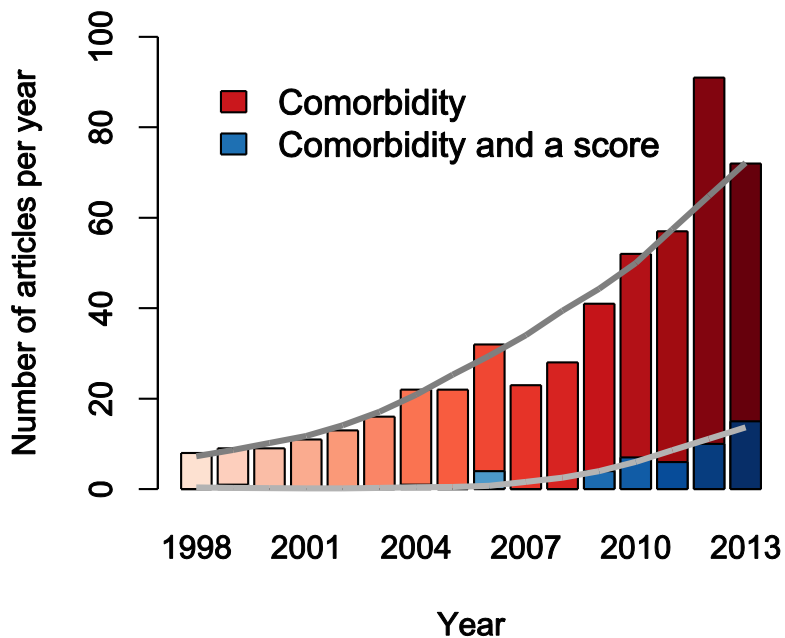


Figure 1. Number of articles according to Pubmed that mention comorbidities.

1.2.3.1 Indices and scores

Due to the multitude of different comorbidities they are frequently summarized in scores/indices. The ASA score is probably the most commonly known comorbidity measure [122]. While the ASA score impacts THR outcomes [123], [123]–[125], it is generally not available in administrative databases.

Administrative databases that log institutional and non-institutional care usually contain ICD codes indicating both the main disease and comorbidities of importance. The availability of the codes has spurred the development of ICD-based scores as illustrated by Figure 1. Unfortunately there are different versions of ICD-codes requiring translation and validation between different versions. For instance, the US uses ICD-9 while we in Sweden use version 10 since 1997, and only 15% of studies comparing and evaluating comorbidities use both systems [126].

There are currently two commonly used ICD-based scores that are available for both ICD -9 and ICD-10: the Charlson comorbidity score or index [127] and the Elixhauser comorbidity measure [128]. The Charlson score has also been simplified and adapted to a surgical setting by the Royal College of Surgeons (RCS) [129].

According to the Charlson score comorbidities have been found to increase overall revision rates. Johnsen et al. reported for scores above 2, hazard ratios of 2.3 (95% CI, 1.6 to 3.5) for the first month, 3.0 (95% CI, 2.1 to 4.5) between one month and six months, and 2.8 (95% CI, 2.3 to 3.3) between six months and 8.6 years [130]. Paterson et al. reported on 20,290 patients that Charlson scores of 1 correlated to increased risk of revision within 1 year, a hazard ratio of 2.1 (95% CI, 1.4 to 3.1). They could not find any risk increase for scores of 2 or above [131]. According to the Charlson score comorbidities have also been shown to increase the risk for infections [105], [107], [132], dislocations [87], early mortality, and adverse events within 90 days [13], [19], [88], [131].

Scores in relation to PROMs have been sparsely investigated. The Charlson score has been associated with poorer SF-36 results in the physical domain for 590 patients [98]. Estimates using the questionnaire based on the Functional Comorbidity Index [133], with similar comorbidities to the Charlson score, indicate similar results for SF-36 outcomes 3, 10, and 16 years after surgery [134], [135].

1.2.3.2 Charnley class

The Charnley classification was introduced in 1972 [47] as a comorbidity grouping for walking disabilities. It is today a widely recognized predictor for patient-reported outcomes after both total hip and knee replacements [93], [136]–[141], even though this has not been unanimously shown [90], [99], [142]. The classification has a simple design with 3 classes:

- A. One involved hip
- B. Two involved hips but no other joints
- C. Some other factor contributing to failure to achieve normal locomotion, such as rheumatoid arthritis, senility, hemiplegia, and cardiovascular or respiratory disability.

Despite the simplicity there has been concern that the classification suffers from inter-observer variability [121], e.g. the class C reported by McGuigan et al. was around 4% [90] while Lavernia et al. reported 30% [142]. To complicate further, some split the B-class into 2 groups, those that have not operated the contralateral hip joint (B1) and those that have (B2) [138], [143]. These inconsistencies can cause difficulties when clinicians try to use the classification system to inform their patients.

Self-administered classification as used for the nationwide follow-up program run by the SHAR may limit the interobserver variability. In the self-administered classification, Charnley class C is both common, 45% among patients eligible for THR, and a strong predictor for poor patient-reported outcomes [93].

The Charnley classification has been sparsely investigated for mortality and re-operation rates. Münger et al. reported on a large European case-control study on 5,035 patients where patients with Charnley class B, i.e. a comorbidity suggesting less mobility and constituting one third of the sample, had fewer cases of aseptic loosening with an odds ratio of 0.8 (95% CI, 0.6 to 0.96). Class C did not show any risk reduction but these patients represented only 2% of the study sample [82].

Charnley classification is more frequently used together with PROMs. Rolfson et al. reported on a large nationwide sample that Charnley class A and B improved on average 0.38 while class C improved 0.35 on the HRQoL EQ-5D index, the corresponding improvement for pain VAS was 49 vs 46 mm [136]. For a disease specific instrument such as the Harris hip score the corresponding averages after 5 years (90 patients) are 94 vs 78 points for [141], and after 13 years, 92 vs 72 points [137].

1.2.3.3 Mental health

Vissers et al. report in their systematic review that mental health negatively correlates to knee arthroplasty outcomes, but it is unclear for hip arthroplasty outcomes. For hip arthroplasties they summarized the evidence as limited, conflicting, or absent [144].

Since then a few interesting studies have appeared. Stundner et al. studied 1 million patients and found that depression and anxiety negatively impact hospital stay, higher cost, and non-routine discharge. Somewhat surprisingly they report lower odds for perioperative in-hospital mortality, odds ratio of 0.5 (95% CI, 0.3 to 0.9) for patients with depression and 0.6 (95% CI, 0.4 to 0.9) for patients with anxiety [145]. Depression has also been associated with increased risk of periprosthetic joint infections, with hazard ratio 1.4 (95% CI, 1.1 to 1.7) [146].

1.2.3.4 BMI and obesity

Obesity is a known risk factor for re-operations. The main concern for obese patients are periprosthetic joint infections. This has been shown to be especially true for the very obese¹ with reported odds ratios ranging between 5 and 18 [147]–[149]. It is possible that there is also a difference between the sexes and the impact of obesity; Lübbecke et al. found in 2,495 THRs that obese women, BMI ≥ 30 , compared to obese men have been have a higher risk both for revision

¹ Super obese > 50 in the Malinzak et al. and Everhart et al study, morbidly obese > 40 for the Jämsen et al. study

due to infection, relative risk 16 (95% CI, 3.4 to 76) vs 1.0 (95% CI, 0.2 to 5), and for revision due to dislocation, relative risk 3.0 (95% CI, 1.3, 7.1) vs 1.8 (95% CI, 0.9, 3.6) [150]. A meta-analysis by Haverkamp et al. further added that obese patients had an increased risk of aseptic loosening and venous thromboembolism [151]. The loosening effect was also found by Münger et al. with an increase in odds ratio of 1.03 for every BMI unit (95% CI, 1.00 to 1.05) [82].

The data regarding PROMs and BMI are currently conflicting. Stevens et al. found in a larger retrospective study on 653 patients that BMI had a modestly negative influence on the WOMAC score and the SF-36 self-perceived general health subscale [119]. Similarly Davis et al found when studying 1,617 THRs that for every 10 points of BMI the Harris hip score decreased by 3.0% (95% CI, 1.6 to 4.4) and all 8 SF-36 components except for mental health with effects ranging between 3 and 10% [152].

On the other hand, Aranda Villalobos et al. looked at BMI > 28 kg/m² and found in 70 prospectively recruited patients that the improvement for the WOMAC functional component and the SF-12 physical component was greater for obese patients [153]. Dowsey et al. in turn found no differences in 471 patients regarding improvement measured by Harris hip score or SF-12 when using WHO' BMI categories [154]. Jones et al. found that there was increased pain at 6 months but that the differences were negligible after 3 years for a mixed cohort of 520 hip and knee arthroplasties [155]. Studies with a 5-year follow-up support these Jones et al.'s results [150], [156].

1.2.3.5 Diabetes

Diabetes is a known risk factor for infectious diseases [157] and especially uncontrolled hyperglycemia has been associated with surgical site infections [148], [158]. While some suggest that with modern techniques there is no increased risk for either deep or superficial wound infections [159], this is partly supported by Namba et al.'s finding that after adjusting for BMI the risk increase for diabetes disappeared [108], but few are ready to dismiss diabetes as a risk factor. Malinzak et al. found in a mixed hip and knee arthroplasty cohort of 6,108 patients that patients with diabetes were 3 times as likely to suffer a periprosthetic joint infection compared to non-diabetic patients even after adjusting for BMI [147]. Pedersen et al. found an increased relative risk of 1.5 (95% CI, 1.0 to 2.2) for revision due to periprosthetic joint infection within 2 years of surgery although they did not adjust for BMI [160].

Marchant et al. investigated a mixed hip and knee arthroplasty cohort of 1 million patients comparing patients with uncontrolled diabetes (0.4%) and

controlled diabetes (10.2%) to those without diabetes, although data regarding BMI was not available in this analysis. They found that controlled diabetes was associated with urinary tract infections, odds ratio of 1.2 (95% CI, 1.2 to 1.3), while patients with *uncontrolled diabetes* had [161]:

- Increased mortality: odds ratio 2.7 (95% CI, 1.7 to 4.4)
- Cerebrovascular accident: odds ratio 4.1 (95% CI, 2.4 to 7.0)
- Urinary tract infection: odds ratio 2.5 (95% CI, 2.0 to 3.0)
- Ileus: odds ratio 2.5 (95% CI, 1.7 to 3.4)
- Infection: odds ratio 2.3 (95% CI, 1.4 to 3.8)
- Postoperative hemorrhage or shock: odds ratio 1.8 (95% CI, 1.3 to 2.6)

Diabetes and PROMs have been sparsely investigated. Judge et al. report a small sample of 282 patients where they found that diabetes increased the risk of failure to improve after surgery according to a physical functioning score of the SF-36 instrument. However, this effect did not remain significant after adjusting for covariates [117].

1.2.3.6 Other diseases

Queally et al. investigated the effect of neurological diseases and concluded that THR can have a good outcome among these patients [162]. Patients with Parkinson's disease are commonly believed to have a higher dislocation rate due to their poor muscular control, although evidence is inconclusive. Meek et al. found in a large registry study on 14,314 THRs that the dislocation rates were not more frequent among patients with stroke or Parkinson's disease [163]. Jämsen et al. found in a case-control study on 857 patients with Parkinson's disease and 2,571 controls that patients with Parkinson's have an increased risk for hip dislocations during the 1st year, hazard ratio 2.3 (95% CI, 1.02 to 5.3) [164].

Bozic et al. recently reported on comorbidities and risk of revisions within 12 months for 56,030 patients 65 years and above¹. They found that out of 29 comorbidity categories, 6 were associated with a higher risk of revision [165]:

- Depression, hazard ratio 1.6 (95% CI, 1.4 to 1.9)
- Rheumatologic disease, hazard ratio 1.3; 95% CI, 1.11–1.57)
- Psychoses, hazard ratio 1.3 (95% CI, 1.1 to 1.7)
- Renal disease, hazard ratio 1.3 (95% CI, 1.1 to 1.6)
- Chronic urinary tract infection, hazard ratio 1.2 (95% CI, 1.01 to 1.3)

¹ The article does not state the surgery indication or the proportion of patients with primary osteoarthritis

- Congestive heart failure, hazard ratio 1.2 (95% CI, 1.01 to 1.4)

The same group also studied 90-day mortality in 40,919 THR patients and found that neurological diseases such as dementia, hemiplegia or paraplegia, and cerebrovascular disease increase the risk estimates. They also found that congestive heart failure, metastatic cancer, psychosis, renal disease, and chronic pulmonary disease correlate with increased mortality rates [146]. Memtsoudis et al's study on in-hospital mortality of almost 7 million procedures, in a mixed hip and knee arthroplasty cohort, confirms the findings that dementia, renal disease, and cerebrovascular disease are associated with increased risk of dying [8].

Singh et al. investigated if heart disease, peripheral vascular disease, renal disease, chronic obstructive pulmonary disease, diabetes, or connective tissue disease increase the risk for persistent moderate/severe pain after 2 and 5 years. Out of the six comorbidities they found that peripheral vascular disease had an almost significant risk of higher odds ratio for moderate/severe pain after 2 years, odds ratio 1.5 (95% CI, 1.0 to 2.4), and somewhat surprisingly renal disease was associated with a lower odds ratio, 0.6 (95% CI, 0.3 to 1.0) [166].

1.3 SUMMARY

Patient characteristics impact the outcomes after THR. The impact is not universal between the different outcomes, i.e. some may improve implant survival while at the same time correlate with poorer PROMs. It is also highly likely that many comorbidities coexist in complex fashion, for instance there are well-established reciprocal links between depression and obesity [167], and obesity and diabetes [168]. This may partially explain why different studies vary regarding the estimates. Furthermore, there are more inconsistencies regarding PROMs than other outcomes and little seems to be known of the underlying mechanisms. Worth noting is that important patient characteristics such as smoking [169], educational status [100], [170], and alcohol consumption [33], [171] are beyond the scope of this thesis.

2 AIMS

The overall aim of this thesis is to estimate the impact of common patient factors on outcomes after THR by using the Swedish Hip Arthroplasty Register, and cross-matching with other population-based registers.

The specific aims of this thesis were to investigate:

1. If comorbidity measures developed for mortality outcomes also are applicable for re-operations.
2. How age influences health-related quality of life (HRQoL).
3. If sex/age interacts with self-administered Charnley classification in regards to HRQoL.
4. The generalizability of HRQOL estimates.
5. If antidepressant use impacts patient-reported outcomes.

3 METHODS

3.1 STUDY DESIGN

All studies except study III are retrospective nation-wide cohort studies. Study III is a cross-sectional study as there is no pre-operative data.

3.2 DATA SOURCES

The five manuscripts included in this thesis are combined from five different data sources (Figure 1).

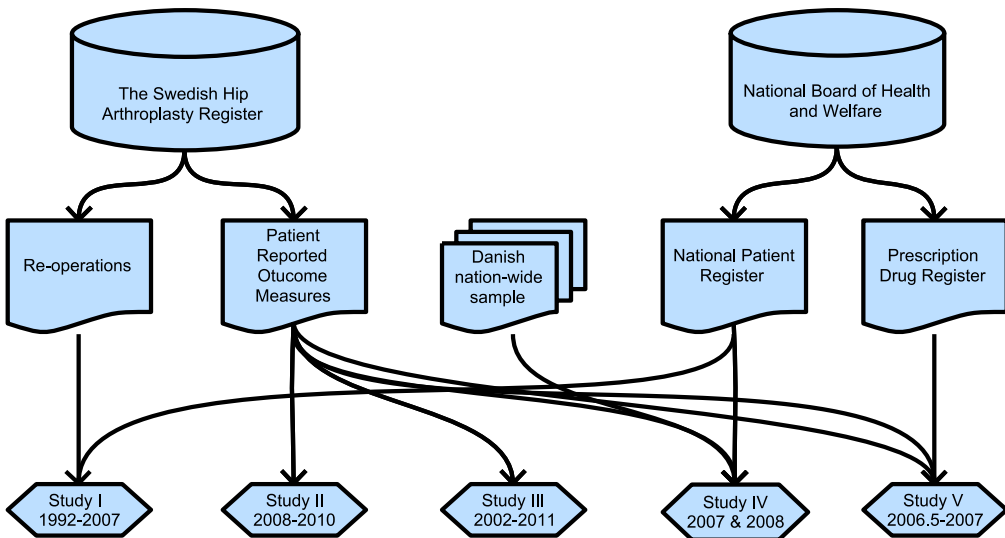


Figure 2: The basis for the study datasets and the surgery years included

3.2.1 The Swedish Hip Arthroplasty Register

Since its inception in 1979 the Swedish Hip Arthroplasty Register has collected data on all primary THRs and their re-operations associated to the primary THR performed in Sweden [20]. The registry is the second oldest quality register in the world and includes 98% of all patients operated with a total hip arthroplasty from all Swedish hospitals [14]. Since 1992 the patient's personal registration number has been collected, allowing follow-up that is patient and hip-specific.

3.2.1.1 Patient-reported outcome measures

In 2002 a program for gathering patient-reported outcome measures (PROMs) was adopted. The program reached full nation-wide coverage in 2008 [93]. A

PROM questionnaire is presented to all elective THR patients pre-operatively and 1, 6, and 10 years after surgery. Each hospital is responsible for data collection and registration within an online database.

3.2.2 The Swedish National Patient Register

The Swedish National Patient Register was started in 1964 and includes all in-patient care in Sweden since 1987 with discharge codes according to ICD-9 and ICD-10 and admission/discharge dates. It was previously known as the in-patient register but as of 2001 it also includes outpatients visits [172]. The proportion of patients with a given diagnoses where the registry code is deemed correct (the positive predictive value) is around 85-90% [173].

3.2.3 The Prescribed Drug Register

The Prescribed Drug Register (PDR) was initiated in July 2005 and includes any withdrawn prescriptions. Prescriptions that are never withdrawn by patients and drugs bought over the counter without prescriptions are not included. The data fields used were the drug ATC-code, number of pills, and prescription text [174].

3.2.4 The Danish study sample

The Danish Hip Arthroplasty Register (DHR) has collected nation-wide samples from the Danish THR population on different PROMs. These include among other items the same HRQoL measure as used by the Swedish Hip Arthroplasty Register, the EQ-5D instrument [175].

3.3 PATIENT SELECTION

3.3.1 Study I

We included all patients operated between 1992 and 2007 that had received a THR due to primary osteoarthritis. Patients that had received a resurfacing arthroplasty or an implant used less than 500 times were excluded.

3.3.2 Study II

We included all patients operated between 2008 and 2010 with THR due to primary osteoarthritis. We only included patients above 40 years of age due to rarity of osteoarthritis prior to 40 years and the risk for indication misclassification.

Patients with re-operations or death within 1.5 years from index operation were excluded. If a patient had two records due to bilateral surgery, only the first hip with complete data was selected.

3.3.3 Study III

Similar to study II, we included all patients operated between 2008 and 2010 with THR due to primary osteoarthritis above 40 years of age. We included only those that had filled out the pre-operative questionnaire. Patients with re-operations or death within 1.5 years from index operation were excluded. If a patient had two records due to bilateral surgery, only the first hip with complete data was selected. We also excluded patients where the second surgery occurred prior to filling out the one-year PROMs questionnaire.

3.3.4 Study IV

We included from the SHAR all patients operated in 2006 and 2007, and from the DHR sample we included all patients operated in 2008 with complete one year EQ-5D. Patients re-operated within a year or with missing values in any of the outcome scores were excluded from the analysis. For bilateral cases, only the first operation with complete data was selected.

3.3.5 Study V

We included patients operated from July 2006 through December 2007. We included only those that had filled out the pre-operative form. If a patient had bilateral hip surgery within the observation period, only the first hip with complete data was included. Patients with re-operations within 1.5 years from index operation were excluded.

3.4 OUTCOMES

3.4.1 Study I

In study I the outcome was re-operations between 0-2 years and 2-12 years. Re-operation was defined as an open surgical procedure localized to the hip joint that could in some way be related to a previous THR, regardless of whether any of the parts or the entire implant was exchanged, removed, or left in situ.

The time to event was defined as the period between the initial surgery and the day of re-operation, death or end of study period, whichever came first.

3.4.2 Study II to IV

In studies II, III and IV, the outcome was HRQoL measured by the EQ-5D index and the EQ VAS 1 year after surgery. The EQ-5D form consists of 6 items, 5 Likert questions and the EQ VAS. The questions span 5 dimensions of health:

1. Mobility
2. Self-care
3. Usual activities
4. Pain/discomfort
5. Anxiety/depression.

Each dimension has 3 levels of severity, generating a total of 243 combinations representing different health states. There are different value sets that may be used to translate these health states into a utility index.

In studies II and III, we used the Swedish experience-based time-trade-off (TTO) value set to translate the answers into a score between 0.34 and 0.97; on a scale 0 represents death and 0.97 maximum attainable HRQoL by the EQ-5D measurement [176]. As the Swedish tariff was published after study IV was published we used the Danish TTO tariff [177] for all patients, ranging from -0.624 to 1.

The EQ VAS, in turn, consists of a vertical VAS ranging from 0 to 100 where the patient is asked to mark their HRQoL and 100 corresponds to full health.

3.4.3 Study V

In study V we used HRQoL, pain VAS, and satisfaction VAS as outcome measures. HRQoL was measured as above using the Swedish tariff. The VASs for pain and satisfaction are similar to the EQ VAS with possible scores from 0 to 100 on a horizontal scale. A score of 0 on the VAS pain scale corresponds to no pain. Similarly, complete satisfaction with the outcome of the operation corresponds to 100 on the satisfaction VAS.

3.5 EXPOSURE VARIABLES

3.5.1 Study I

The exposure variables in study I were three different comorbidity measures calculated using ICD-codes from the NPR. All codes occurring during admissions 1 year prior to the surgery were included.

Of the three measures, the Elixhauser score [128] is the most detailed with 30 different categories; Charlson's [127] consists of 19 categories, while Royal College of Surgeons (RCS) Charlson's [129] has only 14 categories. There is, as

expected, a large overlap between categories in the three measures. The Elixhauser and the RCS Charlson score were calculated by counting comorbidities without any pre-assigned weights, i.e. a score of 2 indicates that the patient has 2 comorbidities as defined by that particular score. The Charlson score was calculated by applying the original weights and ranged between 1 and 6, for instance myocardial infarction generates 1 point while metastatic cancer generates 6 points. All scores were grouped into three categories: 0, 1-2 and ≥ 3 points.

3.5.2 Study II

The exposure in study II was age at time of surgery. Instead of modeling age as traditionally done by using categories or a linear variable, we used splines. The spline allows the line to bend according to predefined rules. This allows for a smooth relationship throughout the variable's span, with minimal residual confounding [178].

There are several different spline flavors, we used restricted cubic splines that use cubic terms in the center of the data and restrict the ends to a straight line; this helps to avoid the center distorting the ends. The flexibility of a spline is chosen by the number of knots, more knots allow a more detailed description of the relationship. To avoid overfitting the regression model by choosing too many knots, i.e. making the line too wiggly, the number of spline knots was chosen using the Bayesian information criterion (BIC) [80].

3.5.3 Study III

In study III the exposure was the Charnley classification interaction with sex and age at surgery. In the PROM program, Charnley classification is assessed by the patient pre-operatively and at 1 year by using two questions:

- 1) Do you have any symptoms from the other hip?
- 2) Do you have problems walking because of other reasons (e.g., pain from other joints, back pain, angina, or any other medical condition impairing walking capacity)?

3.5.4 Study IV

The exposure was the patient's country interaction with sex, age, and Charlson score. Charlson score was calculated as described in study I.

3.5.5 Study V

The exposure was withdrawal of antidepressants 1 year prior to surgery. Antidepressants were identified through ATC codes matching N06A in the PDR.

To avoid studying groups that received antidepressants for other conditions such as addiction, cancer, dementia, or Parkinson's disease, patients taking any medication with ATC code N05A, N04, N06D, N03, N07B, N06AX12, or L02BA01 were excluded.

Patients with suspected chronic pain syndrome were treated as a separate group and were not included in the antidepressant group, even if they received antidepressants. This group was identified by prescription text specifying chronic pain, pain syndrome, or prescriptions of Gabapentin (N03AX12), Pregabalin (N03AX16), or Tryptizol (N06AA09), where at least one of the prescription texts mentioned pain. Patients with Gabapentin or Pregabalin without the mention of pain were not included into the antidepressant group even if they also had antidepressant treatment.

The antidepressant group was further subdivided into 4 groups according to prescription texts:

1. **Depressed:** texts indicating depression and no mention of anxiety.
2. **Anxiety:** texts indicating anxiety and no mention of depression.
3. **Mixed:** texts with both depression and anxiety mentioned were categorized as a separate mixed group.
4. **Other/unknown:** when there was no text indicating either depression or anxiety.

In order to differentiate severity and stages of disease, 3 treatment types were identified:

- **Novel treatment**, indicating non-chronic disease.
Definition: no purchases during the first 6 months of observation. Patients with daily dispensed medications were considered novel if they had not received antidepressants the first 14 days.
- **Discontinued**, possibly indicating less severe disease or poor adherence.
Definition: no further antidepressant purchases occurred since an expected purchase. To limit the risk of drug hoarding the minimum time for discontinuation was set to 3 months. Patients with daily dispensed medications were considered discontinued if they had not received medication for 14 days.
- **Dose increase**, possibly indicating a more severe disease.
Definition: the daily milligram intake increased between prescriptions. Prescription text with escalating dosages were not included.

3.6 STATISTICS

3.6.1 Survival analysis

Re-operation rates were modeled using Cox proportional hazards [179], [180] in study I. The Cox proportional hazard model is a popular survival method that allows estimating individual parameters without knowing the baseline hazard. The method enables estimating the hazard ratio between different variables as a form of relative risk.

3.6.1.1 Cox proportional hazards background

The hazard $h(t_j)$ is defined as the j :th patient's risk of experiencing an event, $f(t_j) = P(T = t_j)$, if he/she has survived that far, $S(t_j) = P(T > t_{j-1})$ [181]. The $P(\dots)$ denotes a probability ranging from 0 to 1.

$$h(t_j) = P(T = t_j | T > t_{j-1}) = \frac{f(t_j)}{S(t_{j-1})} \quad (1)$$

The Cox proportional hazard model states that an individual's hazard function is:

$$h_i(t) = e^{\beta_1 x_{i1} + \dots + \beta_m x_{mi}} h_0(t) \quad (2)$$

Note that the $h_0(t)$ is outside the exponential function and independent of the individual's variables, the x -parameters. If this assumption holds true, i.e. that the parameters do not affect the hazard over time, the estimates, the β -parameters, can be interpreted as hazard ratios. For instance if we have two individuals in a randomized controlled trial with two treatment arms, their functions would be:

$$\begin{aligned} h_{treated}(t) &= e^{\beta_{treatment} * 1} h_0(t) = e^{\beta_{treatment}} h_0(t) \\ h_{not\ treated}(t) &= e^{\beta_{treatment} * 0} h_0(t) = e^0 h_0(t) = h_0(t) \end{aligned} \quad (3)$$

If we want to compare their hazards, we see that the hazard ratio (HR) is independent of the $h_0(t)$:

$$\begin{aligned} HR &= \frac{h_{treated}(t)}{h_{not\ treated}(t)} = \frac{e^{\beta_{treatment}} h_0(t)}{h_0(t)} \\ &= e^{\beta_{treatment}} \end{aligned} \quad (4)$$

3.6.1.2 Competing risk regression

Since a dead patient can no longer be re-operated, death can be viewed as a possible competing risk to re-operation. The results were therefore validated using competing risk regression [181]–[184]. The competing risk regression is

similar to Cox proportional hazards regression with the important exception that it investigates the subhazard (denoted as $\tilde{h}(t)$) instead of the general hazard function. If all subhazard functions are summed they combine into the Cox regression model hazard function, $h(t)$, i.e. for p number outcomes:

$$h(t) = \sum_{i=1}^p \tilde{h}_i(t) = \frac{\sum_{i=1}^p f_i(t)}{S(t)} \quad (5)$$

3.6.1.3 Proportional hazards assumption

The proportional hazards assumption was investigated using Schoenfeld residuals [185], [186]. These should align along a straight line for each estimate. If hazard ratios increase over time, the overall hazard ratio for the risk factor will be overestimated. Conversely, decreasing hazard ratios will lead to underestimation of overall hazard ratio [183].

3.6.1.4 C-statistic

In order to provide an estimate for the scores' ability to identify patients at risk of re-operations and to compare with the scores' evaluation on mortality, we used the area under the ROC-curve for a logistic regression model. A value of 0.5 indicates that a model does no better at identifying patients than pure chance, while a value of 1 would mean that the model manages to correctly identify all patients that are later re-operated.

3.6.2 Linear regression

For studies II to IV we used ordinary least square linear regression models. In the linear regression studies the mean and each estimate should be interpreted as that variable's effect on the mean. The general form is written as:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (6)$$

This translates into matrix format where each patient is one row in the Y and X matrix and the β are the unknown estimates of interest:

$$y = X\beta \quad (7)$$

Estimating the β through the above formula gives:

$$\beta = (X'X)^{-1}X'y = Hy \quad (8)$$

Explanation: The X' indicates the X matrix transposed, i.e. flipped horizontally. By multiplying the transposed matrix with the original, a square matrix is generated. The benefit of the square matrix form is that it allows for calculating the inverse, the $(X'X)^{-1}$ part, and can thus be moved to the y side of the equal

sign, thus solving the equation for β . The complex $(X'X)^{-1}X'$ is often referred to as the 'hat matrix' (H) matrix since Hy gives the predicted values, the \hat{y} where the $\hat{}$ ('hat') indicates that it is estimated and not the true value. The diagonal elements in this matrix relate to the impact that a particular observation has on the model, also referred to as the leverage.

3.6.2.1 Interactions

Interactions were investigated with multiplicative terms. For instance if a sex and Charnley class with three classes (A, B, C) interact, the model changes from:

$$y = \beta_0 + \beta_{Sex}Sex + \beta_B B + \beta_C C \quad (9)$$

To:

$$y = \beta_0 + \beta_{Sex}Sex + \beta_B B + \beta_C C + \beta_{B*Sex} B * Sex + \beta_{C*Sex} C * Sex \quad (10)$$

Where the multiplicative term estimates were investigated through ANOVA-testing the null hypothesis:

$$H_0: \beta_{B*Sex} = \beta_{C*Sex} = 0 \quad (11)$$

3.6.2.2 Heteroskedasticity

In order to construct valid p-values and confidence interval, the linear regression needs to estimate the error for each estimate (β). As the studied outcomes were bounded, i.e. had a ceiling and a floor, we used robust covariance matrices in order to limit the impact of the heteroskedasticity on the model [187].

Ordinary least square regression calculates the variance of the β through the residual errors. The residual error is the part that is not explained by the model, illustrated below by the ε :

$$y = X\beta + \varepsilon \quad (12)$$

By extracting the ε we get:

$$\varepsilon = y - \hat{y} = y - X\beta = y - (X'X)^{-1}X'y \quad (13)$$

By further separating the y we can simplify the formula:

$$\varepsilon = (I_n - (X'X)^{-1}X')y = (I_n - H)y \quad (14)$$

The I_n is the identity matrix, a square matrix with n rows/columns containing only 0 except for the diagonal that contains 1. The identity matrix serves a similar purpose to the number 1 in regular mathematics. This equation illustrates the relationship between the H -matrix and the errors.

The variance for $\hat{\beta}$ is:

$$\text{var}(\hat{\beta}) = \text{var}(Hy) = H\text{var}(y)H' = (X'X)^{-1}X'\widehat{\Omega}X(X'X)^{-1} \quad (15)$$

Under Gauss-Markov assumptions the variance estimates for the $\widehat{\Omega}$ is $\hat{\sigma}^2 I_n$ whereby the above simplifies to:

$$\text{var}(\hat{\beta}) = (X'X)^{-1}\hat{\sigma}^2 = (X'X)^{-1}\frac{\sum_{i=1}^n \widehat{\varepsilon}_i^2}{(n-p)} \quad (16)$$

White [188] noted that under heteroskedasticity the errors will not simplify into a single $\hat{\sigma}^2$, and the more complex variance estimator formula (15) should be retained. He suggested that instead of the $\hat{\sigma}^2 I_n$ for the $\widehat{\Omega}$, the squared errors for each row, $\widehat{\Omega} = \text{diag}\{\widehat{\varepsilon}_1^2, \dots, \widehat{\varepsilon}_n^2\}$, give a better estimate. Together with MacKinnon he further expanded this into the HC3 method [189], that was later popularized by Long and Ervin's method comparison [190]:

$$\widehat{\Omega} = \text{diag}\left\{\frac{\widehat{\varepsilon}_1^2}{(1-h_1)^2}, \dots, \frac{\widehat{\varepsilon}_n^2}{(1-h_n)^2}\right\} \quad (17)$$

The h in the formula originates from the diagonal elements on the H-matrix, see formula (14). By adding this to the denominator it elegantly inflates each element's error to its actual impact on the model. This since the h carries information about each observation's leverage on the estimates (see above).

3.6.2.3 Residuals

If the residuals indicate a non-normal distribution, the model mean may be difficult to interpret. Non-normal distributed errors may also affect inference but with the large sample size this impact is limited due to the central limit theorem. The normality of the residuals was investigated through histograms and QQ-plots.

3.6.3 Quantile regression

For study III we also used quantile regression [191] in order to study effects of Charnley classification on poorly performing patients. The quantile regression is similar to the linear regression with the exception that it studies different quantiles. Quantiles lack distribution assumptions and allow a more detailed study of a variable's impact on the population.

3.6.4 Non-linearity

Throughout all studies we tried to avoid categorizing variables or refraining to simple linear relationships. By linear we mean that a continuous variable such as

age has the same impact between ages 40 and 41 as 80 and 81. We used non-linear relationship as described in section 3.5.2. The non-linearity was evaluated by investigating the coefficients through ANOVA as was done for interactions (see section 3.6.2.1).

In order to limit the risk of overfitting we controlled the flexibility of the splines using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) [80, pp. 23–24, 202–203]. Smaller AIC and BIC indicate a better model. The AIC is defined as:

$$AIC = 2k - \ln(\hat{L}) \quad (18)$$

Where k is the number of parameters in the model and L is the maximum likelihood for the given model. BIC is similar and can be approximated by the formula:

$$BIC = k \cdot \ln(n) - 2\ln(\hat{L}) \quad (19)$$

The n indicates the number of study subjects. Due to the large n in our studies the BIC will select simpler models than the AIC.

3.6.5 Software

All statistical computations were performed in R [192]. The statistical computations were performed using the rms-package [193], while the ggplot2 [194] and Gmisc-package were used for graphical output. For reproducible research, dynamic documents were used together with the knitr-package [195]. SPSS combined with the Python-plugin-in was used for extracting variables and cleaning data. Regular expressions [196] were used for text data mining.

4 RESULTS

4.1 STUDY I

4.1.1 Early re-operations

We identified 134,423 total hip arthroplasties (114,072 patients) with primary osteoarthritis (Table 1). 1,826 (1.4%) arthroplasties had been re-operated within 2 years. The two main reasons for early re-operations within 2 years were dislocations and infections.

The risk for re-operation increased continuously in the RCS Charlson and Elixhauser scores (Figure 3). The Elixhauser score was the only measure where both comorbidity groups were significantly different from the healthy group, and it was least affected by the confounders. None of the score estimates showed significant changes when using only ICD-10 codes. The C-statistic for the best performing measure, the Elixhauser score, was 0.52 (95% CI, 0.51 to 0.53).

We also compared the results with an unweighted Charlson score, and the difference in the estimates was negligible. The weighted van Walraven version of the Elixhauser score exhibited similar estimates to the unweighted score, although it lacked the continuous increase.

4.1.2 Late re-operations

From the original cohort 118,065 total hip arthroplasties had been observed for more than 2 years, average follow-up was 7.1 (SD: 3.2) years. In this group 4,244 (3.6%) arthroplasties had been re-operated. The main reason for re-operation in the later period was aseptic loosening, and the second most common reason was dislocation.

In the period 2-12 years, none of the scores had any significant impact on the re-operation rates. When we applied a non-weighted Charlson score the category ≥ 3 had a close to significant estimate $p = 0.053$, hazard ratio of 1.9 (95% CI, 1.0 to 3.7). This estimate also increased when looking at only patients with ICD-10 codes, 2.6 (95% CI, 1.1 to 6.2). We also investigated the van Walraven version of the Elixhauser score, but it did not perform better than the unweighted score.

Variable	0 - 2 years n = 134,423		2 - 12 years n = 118,065	
	No. of patients	No. (%) of re-operations	No. of patients	No. (%) of re-operations
Age				
< 50	3,351	55 (1.6%)	3,006	273 (9.1%)
50 to 59	16,755	245 (1.5%)	15,148	979 (6.5%)
60 to 69	52,676	728 (1.4%)	46,317	1,158 (2.5%)
70 to 79	41,777	503 (1.2%)	37,121	1,659 (4.5%)
≥ 80	19,864	295 (1.5%)	16,473	175 (1.1%)
Sex				
Female	77,134	897 (1.2%)	68,203	2,028 (3.0%)
Male	57,289	929 (1.6%)	49,862	2,216 (4.4%)
No. surgery				
First	112,032	1,487 (1.3%)	98,993	3,762 (3.8%)
Second	22,391	339 (1.5%)	19,072	482 (2.5%)
Charlson score				
0	119,427	1,572 (1.3%)	106,070	3,957 (3.7%)
1-2	13,914	235 (1.7%)	11,279	273 (2.4%)
≥ 3	1,082	19 (1.8%)	716	14 (2.0%)
RCS Charlson score				
0	123,299	1,657 (1.3%)	109,334	4,035 (3.7%)
1-2	10,897	164 (1.5%)	8,583	206 (2.4%)
≥ 3	227	5 (2.2%)	148	3 (2.0%)
Elixhauser score				
0	106,908	1,386 (1.3%)	95,896	3,665 (3.8%)
1-2	25,071	389 (1.6%)	20,399	546 (2.7%)
≥ 3	2,444	51 (2.1%)	1,770	33 (1.9%)
Type of prosthesis				
Cemented	122,833	1,621 (1.3%)	108,642	3,626 (3.3%)
Uncemented	4,160	74 (1.8%)	3,167	194 (6.1%)
Hybrid	4,122	62 (1.5%)	3,904	397 (10.2%)
Rev. hybrid	3,308	69 (2.1%)	2,352	27 (1.1%)
Hospital type				
University hospital	15,598	197 (1.3%)	14,184	823 (5.8%)
County hospital	54,645	898 (1.6%)	48,096	1,821 (3.8%)
Rural hospital	55,590	618 (1.1%)	48,485	1,425 (2.9%)
Private hospital	8,590	113 (1.3%)	7,300	175 (2.4%)

Table 1: Study population characteristics for study I

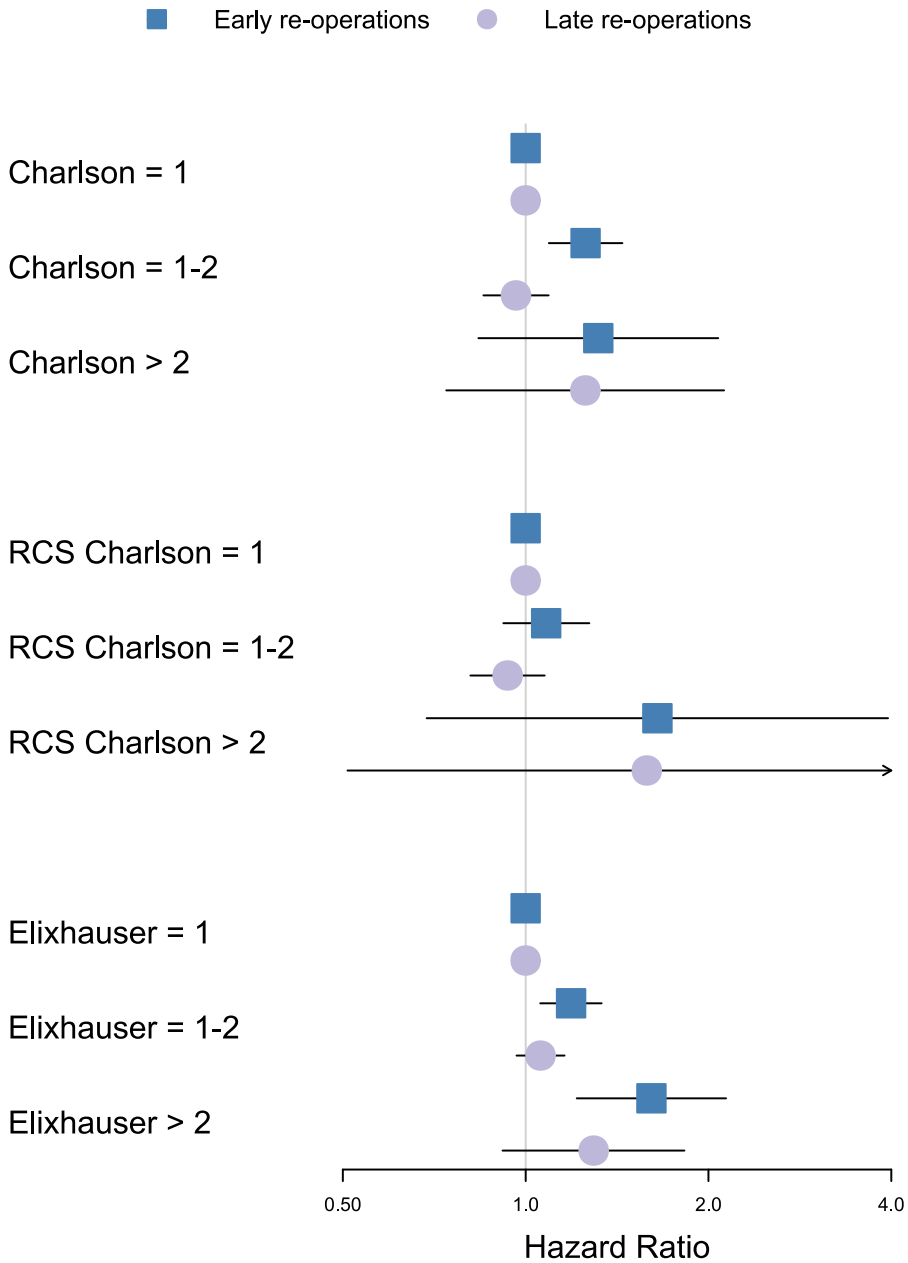


Figure 3: The different scores and their hazard ratios. All scores are fully adjusted for confounders. None of the scores' hazard ratio estimates go above 2, and only the Elixhauser score indicates a continuous increase with increasing score.

4.1.3 Other analyses

4.1.3.1 Sex

Women had a lower re-operation rate for both the early and late period, hazard ratio of 0.7 (95% CI, 0.6 to 0.8) during the first period and 0.7(95% CI, 0.6 to 0.7) in the later period. During both periods we observed a difference in the effect of implant fixation between the sexes. Women had a higher hazard ratio with all implants other than fully cemented, with the only exception for reverse hybrids in the later period. In men the choice of fixation had no certain influence during any of the periods.

4.1.3.2 Age

The age estimate changed considerably between the two time periods. In the early period, 0-2 years, the effect was negligible, while in the later period, 2-12 years, the risk for re-operation decreased with increasing age (Figure 4).

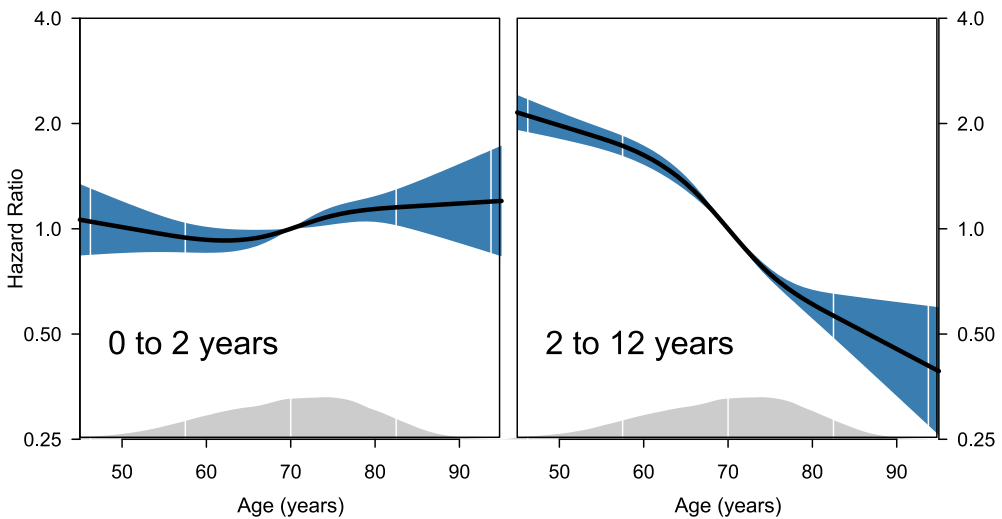


Figure 4: Adjusted estimate for age and risk for re-operation where the median age is the reference. The grey area at the bottom indicates the age distribution in the study population.

4.2 STUDY II

We identified 27,245 patients during this period that had been operated due to primary osteoarthritis with complete data. The majority of patients improved according to both measures. For EQ VAS twice as many failed to improve compared to the EQ-5D index (16% vs 9%). This failure to improve increased in patients above 80 years of age; 20% of these elderly patients failed to improve their EQ VAS score and 13% failed to improve their EQ-5D index. However, only a minority (14%) of patients were above 80 years of age (Figure 5).

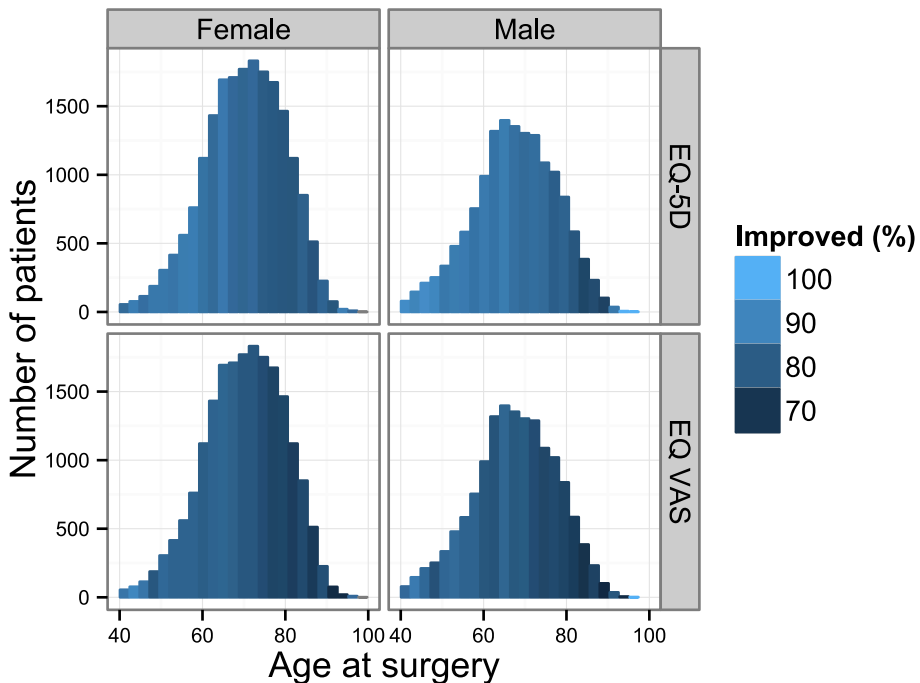


Figure 5: Age distribution between the sexes and the proportion that fail to improve their HRQoL according to EQ-5D and EQ VAS. Darker colors indicate a larger proportion that fail to improve 1 year after surgery compared to their pre-operative value.

4.2.1 Non-linearity of age

The regression model confirmed that age was associated with a decrease in the 2 HRQoL outcomes; a 40-year-old patient compared to an 80-year-old patient had on average a 0.030 higher gain in EQ-5D index (95% CI, 0.023 to 0.037) and a 5.4 mm higher gain in EQ VAS (95% CI, 4.2 to 6.6). Furthermore the age-related decrease was non-linear for both measurements, with a decrease that started in the patient's late sixties (Figure 6). The differences between patients below 70 years were minimal compared to patients above 70 years, the estimates for the EQ-5D index and the EQ VAS increased twofold for each decade after 70 (Table 2).

The dimensions that corresponded closest to the overall spline were mobility and usual activities. We found no support for interaction between age and sex for either outcome.

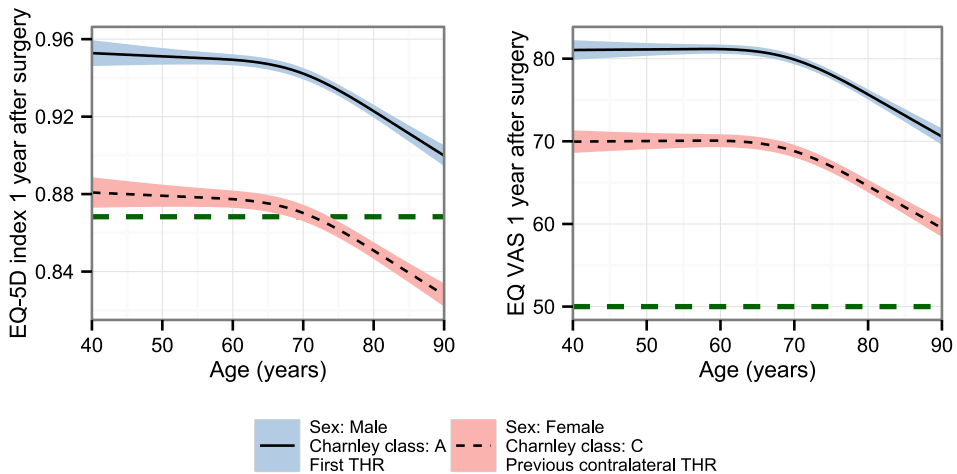


Figure 6: The relation between the EQ-5D index and the EQ VAS 1 year postoperatively and the patient's age at surgery. Pre-operative EQ-5D index and EQ VAS were set to the most frequently occurring values (Index = 0.87; VAS = 50) and are indicated by the horizontal dashed lines. The change before and after surgery is the height above this line, i.e. anything above is an improvement. The 2 lines differ only in height; the solid line with blue confidence interval indicates the optimal combination of covariates (male sex, first hip and Charnley class A) while the dotted line with pink confidence interval indicates the least favorable combination (female sex, previous contralateral hip surgery and Charnley class C). The pain VAS was set to the median, 65mm.

Age (years)	Estimate	95% CI	P-value
EQ-5D index			
40 vs. 70 years	0.011	0.003 - 0.018	0.004
50 vs. 70 years	0.009	0.004 - 0.013	< 0.001
60 vs. 70 years	0.007	0.005 - 0.009	< 0.001
70 vs. 70 years	Ref.	-	-
80 vs. 70 years	-0.019	-0.022 - -0.017	< 0.001
90 vs. 70 years	-0.042	-0.048 - -0.037	< 0.001
EQ VAS (mm)			
40 vs. 70 years	1.2	-0.1 - 2.4	0.078
50 vs. 70 years	1.2	0.4 - 2.0	0.002
60 vs. 70 years	1.3	1.0 - 1.6	< 0.001
70 vs. 70 years	Ref.	-	-
80 vs. 70 years	-4.2	-4.7 - -3.8	< 0.001
90 vs. 70 years	-9.3	-10.3 - -8.3	< 0.001

Table 2: Age estimates with 70 years as reference

4.2.2 Pre-operative HRQoL

The pre-operative EQ-5D index and the EQ VAS explained substantially more than age; 3.9 times for the EQ-5D index and 1.9 times for the EQ VAS when comparing the variables' impact on the model R-square value. Both pre-operative HRQoL values exhibited a non-linear relationship with the postoperative value. Patients with low pre-operative values had the biggest gain, although they did not reach the same absolute levels. With increasing pre-operative values the improvement lessens, and patients with pre-operative values close to the ceiling actually declined on average 1 year later.

4.3 STUDY III

Complete data was available for analysis on 26,249 patients. Women were slightly older, experienced more pain, and categorized themselves more frequently in Charnley class C (Table 3).

Women in Charnley class C had a poorer outcome than men (Figure 7). This interaction was strongest for the EQ-5D index, p-value for the EQ-5D index < 0.001 and for EQ VAS 0.0075. Age did not interact with Charnley class, p-value for the EQ-5D index 0.57 and EQ VAS 0.30.

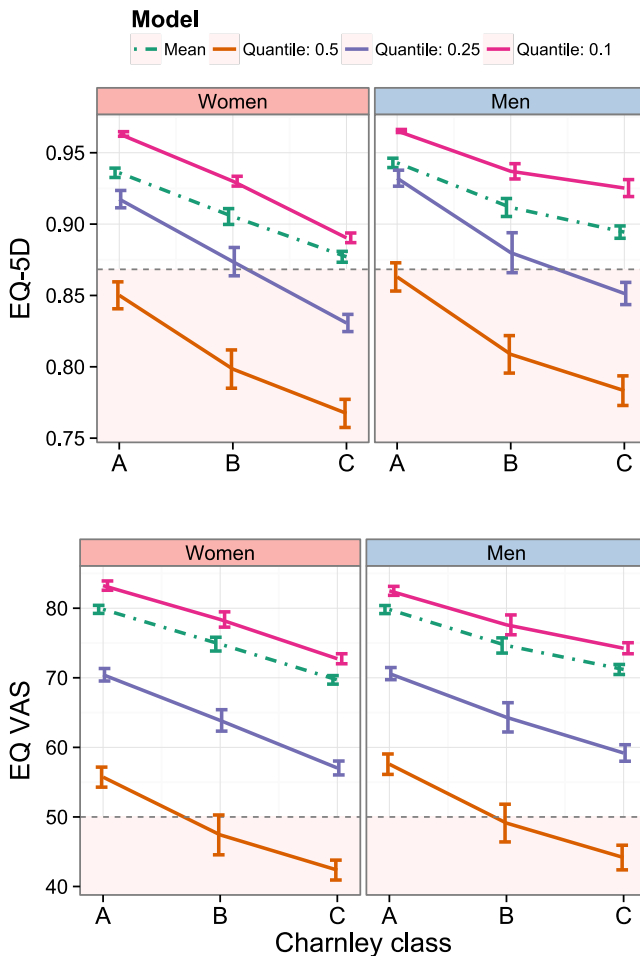


Figure 7: An illustration of the different impact of Charnley class depending on sex. The red area indicates values below the pre-operative index value. The dashed line indicates the predictions from a linear regression model while the solid lines depict quantile regression predictions at the different quantiles.

Variable	Women (n = 16,322)	Men (n = 12,263)
Age (SD)	70 (\pm 10)	67 (\pm 10)
Pain VAS (mm)		
Mean (SD)	64 (\pm 16)	59 (\pm 16)
Missing	2 (0%)	5 (0%)
Charnley class		
A	7,145 (44%)	6,604 (54%)
B	1,788 (11%)	1,262 (10%)
C	7,387 (45%)	4,396 (36%)
Missing	2 (0%)	1 (0%)
Previous contralateral THR		
No	13,010 (80%)	10,102 (82%)
Yes	3,312 (20%)	2,161 (18%)
HRQoL Estimates		
EQ-5D index (Pre-operative)		
Mean (SD)	0.72 (\pm 0.12)	0.75 (\pm 0.11)
Missing	7 (0%)	8 (0%)
EQ-5D index (1 yr)		
Mean (SD)	0.87 (\pm 0.11)	0.89 (\pm 0.10)
Missing	1,292 (8%)	1,015 (8%)
EQ VAS (Pre-operative)		
Mean (SD)	52 (\pm 22)	58 (\pm 22)
Missing	14 (0%)	11 (0%)
EQ VAS (1 yr)		
Mean (SD)	75 (\pm 21)	78 (\pm 19)
Missing	1,294 (8%)	1,014 (8%)

Table 3: Basic characteristics of study population, displaying the absolute numbers (%) for proportions and mean (\pm SD) of continuous variables. Missing is always presented as percentages for that particular group.

4.3.1 EuroQoL dimensions

The sex interaction was mostly mediated by the mobility dimension where men in class C improved more frequently than women. 50% of men vs 40% of women improved at least 1 grade. The dimensions typically affected by the Charnley class were pain/discomfort and mobility. The pain/discomfort did not differ between class B and C, while both mobility and usual activities decreased with higher Charnley class.

4.3.2 Grouping or subdividing the classification

Each class of the Charnley classification had a unique impact on the HRQoL one year after surgery. The model with Charnley class A and B grouped had poorer BIC-estimates, indicating an inferior model. The Charnley classification's contribution to the model improved by 12% for the EQ-5D index, and by 11% for the EQ VAS when we separated the A and B class.

When analyzing the subdivision of class B, 3,156 patients were re-assigned to subclass B2 as they had according to SHAR previously operated the contralateral hip. Out of these, 2,749 patients were from class A and 407 patients from class B. While there was a considerable change between classes, the regression models did not improve by applying this knowledge; both the EQ-5D index and EQ VAS had poorer BIC with the split B class.

4.3.3 Self-reporting

Women reported Charnley class C more frequently than men both before (45% vs. 36%) and 1 year after surgery (47% vs. 37%). We also found a considerable change in reported class before and after surgery (Figure 8), 39% changed class. Crossovers that should be rare according to the original Charnley classification were fairly common, 18% changed either to class A from class B or C, or to class B from class C. The difference between men and women regarding crossovers was small, although women tended to a higher degree deteriorate to class C 1 year later.

In the 1,284 patients that we excluded from the analysis due to contralateral surgery prior to reporting the 1 year follow-up form, class A was the least common class (6%). Class B was larger, approximately 4 times more frequent than in the studied population (46%), while the C-class proportion was similar to studied population (47%).

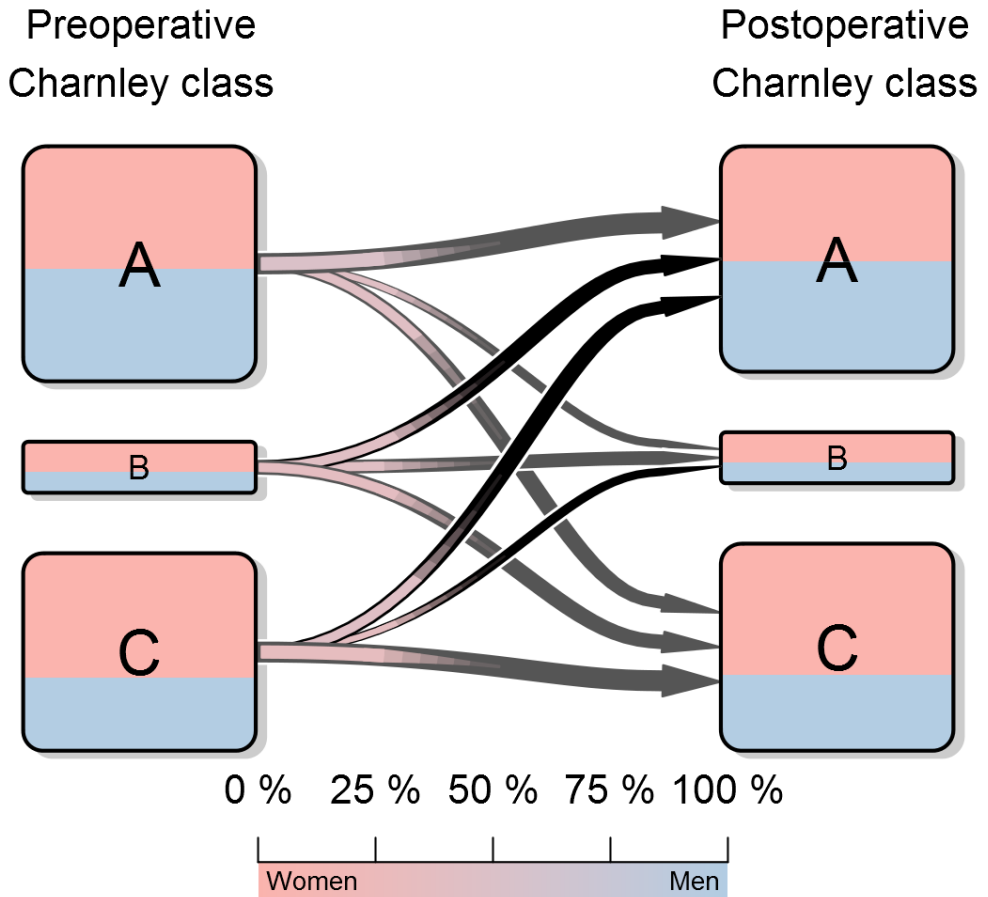


Figure 8: The crossover between Charnley classes from before to after surgery, the patients before surgery are on the left while those to the right are after surgery. Pink indicates the proportion of women in each group while blue corresponds to the proportion of men. The size of the arrow is proportional to the percentage of patients leaving that specific class, and the color of the gradient corresponds to the sex proportion for each transition according to the color bar.

4.4 STUDY IV

We identified 14,560 patients from the SHPR registry and 632 patients from the Danish nation-wide sample. The Danish patients had an overall higher EQ-5D index and EQ VAS than Swedish patients (p-value < 0.001). Apart from this difference the two cohorts were similar.

We found no difference between the countries in how sex, age or the Charlson comorbidity index affected the EQ-5D index (p-value = 0.83) or EQ VAS (p-value = 0.41) 1 year after hip replacement (Figure 9).

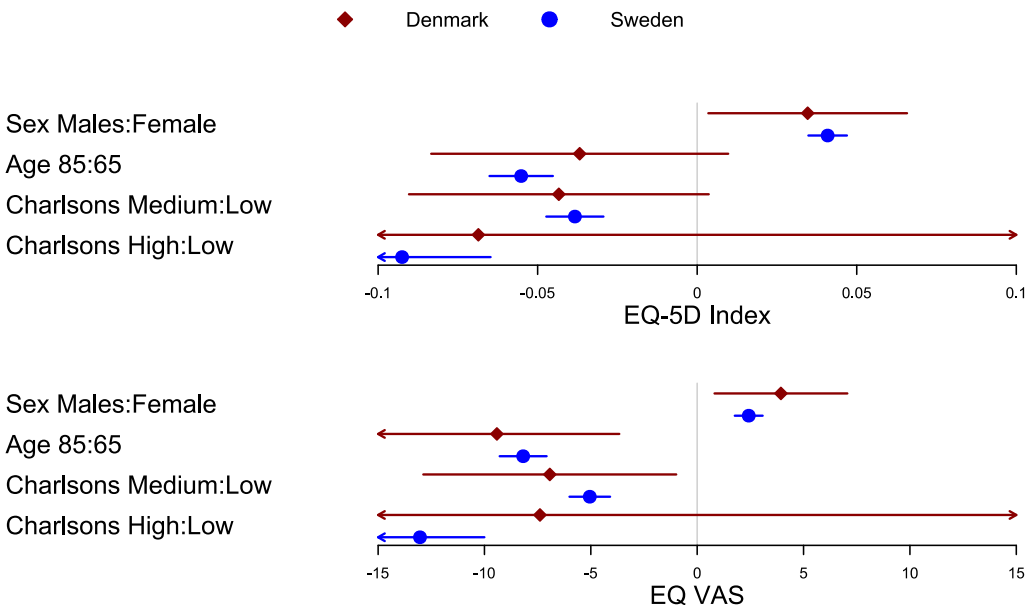


Figure 9: Comparison of factors influencing EQ-5D index between Swedish and Danish patients. Forest plot with 95% confidence intervals for the estimates of EQ-5D index one year after THR for gender (reference=female), age 85 years (reference=65 years), and medium or high Charlson (reference=low Charlson) for Swedish (blue) and Danish (red) patients.

4.5 STUDY V

We identified 10,700 patients that fulfilled our selection criteria. Up to 1 year before surgery, 9% of these patients acquired antidepressants at least once (n = 954). Of the patients treated with antidepressants, 51% had a prescription text indicating depression (n = 487), 3% indicated anxiety (n = 31), and 8% had text indicating both depression and anxiety (n = 75). Of the entire study population, 2% of patients (n = 177) were identified as being treated for a suspected pain syndrome with antidepressant medications.

4.5.1 Antidepressant effect

The use of antidepressants significantly decreased improvement in the EQ-5D index, EQ VAS, and pain 1 year after surgery after adjusting for gender, age, Charnley classification, pre-operative pain, pre-operative HRQoL, self-reported pre-operative anxiety/depression, pain treatment, and antidepressant (Table 4). Antidepressants did not affect satisfaction after surgery.

Among the treatment types we identified 146 patients in the novel treatment group (15%), 150 patients in the discontinued group (16%), and 93 that had increased their dosage (10%). New and increased dosage of antidepressants 1 year before surgery did not appear to influence any of the outcomes at one year. However, discontinuation of antidepressant usage was negatively associated with the patients' pain and mildly associated with their satisfaction at 1 year (Table 5). The effect of treatment discontinuation on the pain outcome was also similar to those with a suspected pain syndrome (Table 4).

4.5.2 EQ-5D anxiety/depression and antidepressants

The prevalence of antidepressant usage increased with increasing severity in the EQ-5D anxiety/depression dimension: 5% for those patients that reported no anxiety/depression, 14% for those patients that moderate anxiety/depression, and 26% for those patients that extreme anxiety/depression used antidepressant. There was no interaction between antidepressant use and the pre-operative anxiety/depression dimension, the smallest p-value identified was 0.1 for satisfaction (as measured by ANOVA).

The majority (66%) of the patients reporting some form of anxiety/depression reported less anxiety/depression 1 year after surgery. This number was slightly lower among those treated with antidepressants (52%). While a few patients deteriorated in this dimension, those without any problems increased from 60% to 79% (Figure 10).

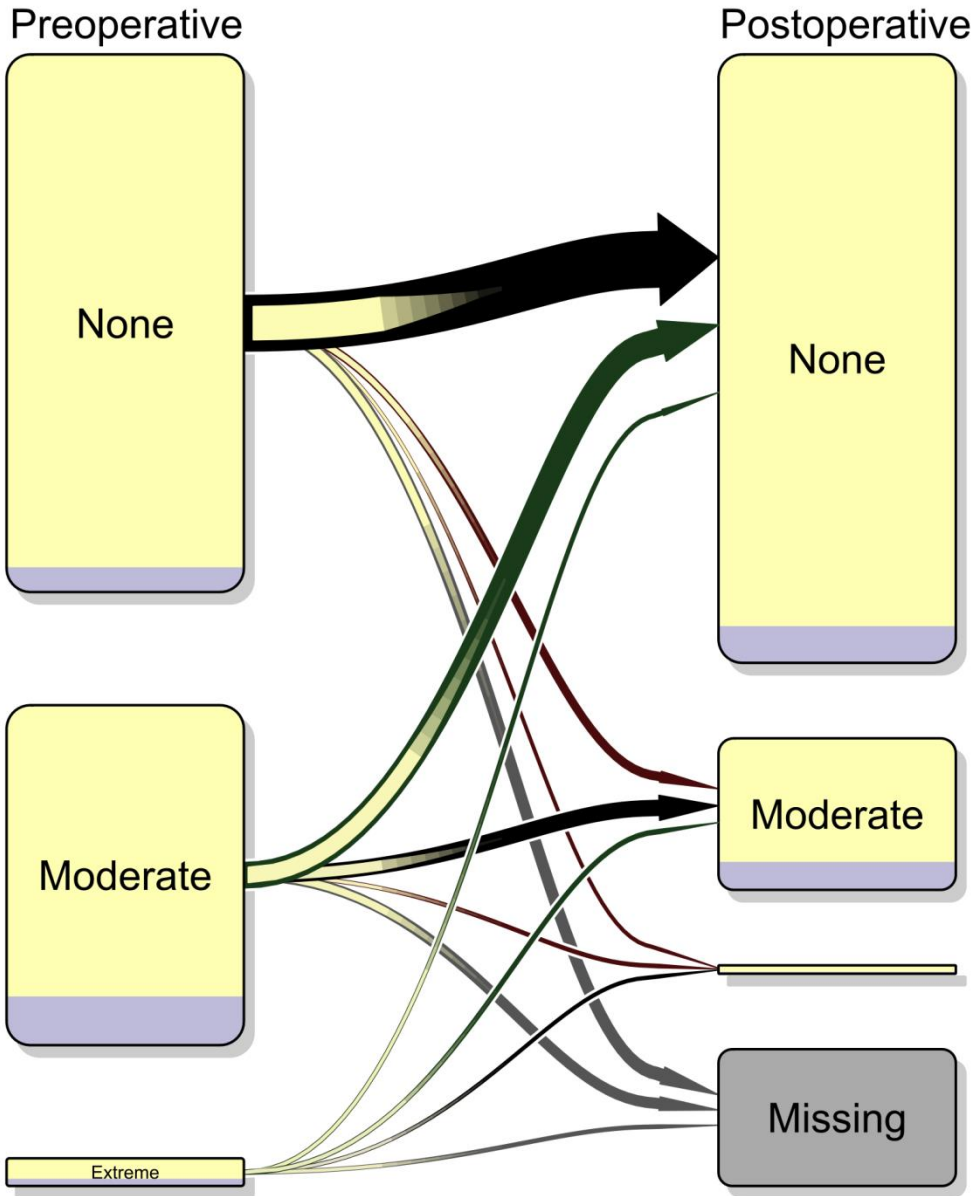


Figure 10: The crossover between anxiety/depression states from before to after surgery, the patients before surgery are on the left while those to the right are after surgery. The blue area indicates the proportion of patients with antidepressant treatment in each group and the missing are those that failed to return the follow-up form. Improvements are highlighted by green arrows while deteriorations by red arrows. The size of the arrow is proportional to the percentage of patients leaving that specific class.

4.5.3 Sensitivity analyses

75% of the patients had at least 2 medication acquisitions up to 1 year before surgery (n = 711). The results were mostly unchanged when those with poor MPR were excluded (MPR < 0.8). Patients treated for depression, the most common indication, had also similar estimates to the overall group, although pain and satisfaction with THR 1 year after surgery were not significant. There was a slightly higher proportion of missing values among those receiving antidepressant treatments (p-value from Fisher's exact test = 0.068). The imputation resulted in the confidence intervals shrinking and discontinuation became significantly associated with satisfaction.

	The EQ-5D index		The EQ VAS	
	Estimate	95% CI	Estimate	95% CI
EQ-5D Anxiety/depression dimension				
None	ref	-	ref	-
Moderate	-0.038	-0.049 to -0.025	-4.5	-5.4 to -3.7
Extreme	-0.075	-0.122 to -0.032	-9.8	-13.2 to -6.7
Treatment with antidepressants				
No	ref	-	ref	-
Yes	-0.042	-0.059 to -0.023	-3.0	-4.6 to -1.5
Indication subgroups^a				
No treatment	ref	-	ref	-
Depressed	-0.039	-0.062 to -0.015	-4.7	-7.0 to -2.6
Anxious	-0.005	-0.086 to 0.071	-2.6	-9.4 to 4.4
Both anxious and depressed	-0.014	-0.083 to 0.046	0.1	-4.7 to 4.9
Unknown indication	-0.050	-0.080 to -0.021	-1.5	-3.6 to 0.7
Treatment change^b				
No change	ref	-	ref	-
Novel treatment	-0.003	-0.056 to 0.049	-1.0	-5.1 to 2.7
Dose increase	-0.036	-0.112 to 0.034	0.2	-5.0 to 4.9
Discontinued treatment	-0.004	-0.051 to 0.042	-2.9	-6.9 to 1.0
Pain syndrome				
No	ref	-	ref	-
Yes	-0.125	-0.169 to -0.081	-7.2	-10.4 to -3.9

Table 4: Regression coefficients and 95% confidence intervals (CI) for health related quality of life outcomes. Negative estimates indicate poorer outcomes.

^a The indication subgroups use the same regression variables as the general model but has the diagnosis groups instead of the antidepressant variable.

^b The differences in treatment changes build on the general model with the addition of the treatment changes, i.e. the values are adjusted for the basic antidepressant treatment effect.

	Pain VAS		Satisfaction VAS	
	Estimate	95% CI	Estimate	95% CI
EQ-5D Anxiety/depression dimension				
None	ref	-	ref	-
Moderate	2.8	2.0 to 3.6	3.7	2.7 to 4.6
Extreme	4.7	1.8 to 7.7	6.1	2.9 to 9.6
Treatment with antidepressants				
No	ref	-	ref	-
Yes	1.5	0.1 to 3.0	0.6	-0.9 to 2.3
Indication subgroups^a				
No treatment	ref	-	ref	-
Depressed	1.6	-0.5 to 3.6	1.5	-0.8 to 3.5
Anxious	2.5	-5.2 to 10.5	0.4	-7.3 to 10.0
Both anxious and depressed	0.1	-4.4 to 4.8	-4.8	-8.5 to -0.8
Unknown indication	1.5	-0.8 to 3.8	0.5	-1.9 to 3.1
Treatment change^b				
No change	ref	-	ref	-
Novel treatment	-1.4	-4.6 to 2.0	-3.2	-6.9 to 0.9
Dose increase	0.3	-4.0 to 4.8	1.8	-3.3 to 7.2
Discontinued treatment	4.5	0.5 to 9	4.5	0.0 to 9.2
Pain syndrome				
No	ref	-	ref	-
Yes	3.5	0.6 to 6.6	1.6	-1.7 to 5.0

Table 5: Regression coefficients and 95% confidence intervals (CI) for pain and satisfaction. Positive estimates indicate poorer outcomes.

^a The indication subgroups use the same regression variables as the general model but has the diagnosis groups instead of the antidepressant variable.

^b The differences in treatment changes build on the general model with the addition of the treatment changes, i.e. the values are adjusted for the basic antidepressant treatment effect.

5 DISCUSSION

5.1 STUDY I

5.1.1 Discussion of results

Of the three measures, the unweighted Elixhauser comorbidity score was closest to fulfilling the qualities that we sought after in the early period. In the late period none of the measures proved satisfactory.

We believe that predicting arthroplasty outcome from comorbidities scores is more challenging than mortality. The C-statistic of 0.52 for the best score is far from the scores' mortality models that range between 0.7 and 0.9 [197]–[199]. These results are cause for caution when interpreting comorbidity adjusted results, especially when the scores have not been validated for the studied outcome.

The difficulty in predicting re-operations in the 2-12 year period is not surprising. One possible explanation is that comorbidities' proximity in time may correlate to their impact, e.g. a cancer 10 years ago will have less impact than a cancer 1 year ago on a patient's risk of dying, and will probably also have less impact on re-operation rates. Beside this attenuation effect, the early and late periods exhibited different reasons for re-operations that may explain the lack of co-morbidity impact. It is also possible that severe comorbidities cause patients to use their hip less or die prior to any re-operation, thus having a protective effect.

Bozic et al [146] looked at specific comorbidities in the Elixhauser score and infections after total hip arthroplasties. While they found rheumatologic disease, obesity, coagulopathy, and pre-operative anemia to be significant, none of these had a hazard ratio of more than 2. Surprisingly they failed to confirm previous findings of such common risk factors as diabetes [160], [161], [200]. Identifying universal risk factors seems to be a surprisingly difficult task.

5.1.2 Strengths

Our study's main strengths are: a homogenous population, high quality of the data, and a large sample size. Choice of implant might be influenced by patient's comorbidities [111], we believe therefore that it is vital to exclude rare, unproven implants, and to adjust for fixation type. By investigating only patients with primary osteoarthritis, and excluding rare implants, the population was ideal for evaluating the comorbidity measures' impact. When studying comorbidities it is important that there is no residual confounding for the age variable. Age is

closely related to comorbidities and often regarded as the most important comorbidity measure. In this study we used a spline for age instead of age groups as commonly used [84], [88], [130], this minimizes residual confounding for age, and reduces the risk for measurement error or cut point bias[80].

5.1.3 Limitations

We had no information concerning the technical outcome, patient compliance, or the surgeon's experience. Although we stratified for type of hospital and volume, we lack detailed knowledge about the surgeon's experience, a factor that is undoubtedly important [201]–[203].

Another important limitation is that we probably failed to identify all comorbidities due to the crude nature of registry data. Although data from tertiary care or drug registry was not available during the study period, we believe that this detection bias is geared towards the less severe diseases. There is therefore reason to believe that the already weak associations would attenuate with inclusion of these undetected diseases. It is also noteworthy that the estimates were unchanged when we looked at the ICD-10 subgroup, although these patients were studied during a later period with higher prevalence of comorbidities. We interpret the increased prevalence of comorbidities in the later study period as due to improved diagnosis registration. This is due to the fact that a partly ICD-code-based funding was introduced in the 1990s, putting more emphasis on the patient's comorbidities. This is supported by the Elixhauser score, with less severe comorbidity groups, that increased more than the other scores, while the age distribution remained unchanged during the study period.

Patients with comorbidities might also be subjected to a surveillance bias. These patients may be more familiar with hospital settings and seek medical attention sooner, which affects the estimates. Although this effect is hard to rule out, we believe that the clinical problems subsequently leading to a re-operation bring the majority of patients to seek medical attention.

In the study we looked at re-operation rates, and it is important to recognize that this is just one out of many possible outcome measures after total hip arthroplasties. Other outcomes such as quality of life, non-surgically treated infections or dislocations, or early re-admissions (e.g., due to pulmonary embolism and cardiovascular disease), are also important and require separate validation. It also important to note that re-operation is a wider definition of failure than revision. Although revision does not include hip-related surgery that leaves the implant untouched, previous studies with this outcome [88], [130], [204] have found similar estimates for the Charlson score. This strengthens the external validity of our findings.

We constructed the Elixhauser comorbidity score by counting disease categories. Elixhauser et al questioned this approach [128] since the impact of disease categories may vary between different outcomes. We believe that using a score instead of enumerating all the 30 comorbidities decreases the risk of overfitting the model. We also investigated an alternative weighted Elixhauser score, as suggested by van Walraven et al [198], but it did not outperform the unweighted score. Combined with the similarity of the different Charlson scores there may be a lack of difference between weighted and unweighted scores. It is possible that this could be due to a surgeon coding bias; a more difficult patient will probably have a more rigorous work-up, and thus also a more complete coding of the comorbidities.

5.2 STUDY II

5.2.1 Discussion of results

We found a non-linear correlation between age and HRQoL 1 year after surgery. Both scores remained largely unaffected by age until a downturn in the patient's late sixties. It is worth noting that increased age was correlated with lower pre-operative HRQoL, as lower pre-operative HRQoL corresponded to a greater improvement, the majority of the elderly increasing in HRQoL 1 year after surgery.

Since the mean age for surgery in this cohort was 69 years, one could argue that surgery is performed too late in order to maximize HRQoL outcome. On the other hand, the decline in improvement may be due to natural age-related deterioration of HRQoL that is not hip related, therefore it is expected that the intervention may have a smaller impact on the gained HRQoL in the elderly. HRQoL is also just one out of many possible PROM outcomes after THR, and while it is an interesting metric it should be combined with other measures during patient consultation [93].

Our results are in line with other large cohort studies. Pennington et al. reported that within 30,203 patients age correlated with a decrease in the EQ-5D index [205]. Judge et al. showed that for 1,375 THR patients, whose age was categorized into < 50, 50 to 60 and > 60 years, both the younger and elder categories performed worse according to the Oxford Hip Score. They concluded that age had a non-linear effect, but failed to explore the relationship further [206]. Cushnaghan et al showed in a small sample of 278 patients with THR that the SF-36 score was poorer in subjects above 67 when surveyed eight years after surgery [207]. It is not surprising that many studies have failed to detect the correlation for HRQoL outcomes [97], [98], [208], as the age variable only explained a minor part of the variation. Our data suggest that adjusting for age

above 70 years will mitigate confounding. While not recommended, ignoring age prior to 70 years can be a viable alternative in those situations where splines are difficult to implement.

In addition to non-linearity for age, both the pre-operative EQ-5D index and the EQ-VAS exhibited non-linearity. Using the difference between the preoperative and the postoperative values, as many previous studies have done [93], [209], will lead to residual confounding as it assumes both linearity and that the preoperative HRQoL estimate equals 1. In these instances, the pre-operative HRQoL state carries some unadjusted information, and therefore, residual information in this variable. Furthermore, the figures demonstrate the instrument's ceiling effect. When patients begin with a high HRQoL close to the ceiling, they will on average deteriorate as any improvement cannot be measured beyond the instrument's ceiling. This can also be viewed as a regression to the mean where extreme observations tend to move toward the population mean.

5.2.2 Strengths

The main strength of our study is the nation-wide population and should therefore be representative of the common patient. The small difference between respondents and non-respondents at the 1 year follow-up also supports our belief that the results are generalizable.

Another important strength is that we used splines instead of categorization. Categorizing continuous variables introduces a cut point bias [210], [211]. Identifying the best divisor between groups by optimizing cut points can increase the risk for type I errors. For instance, a study that reports a p-value of 0.05 and that has applied optimal cut points should in reality be reporting a p-value > 0.25 [212]–[215]. If standard cut points or quantiles are used instead of optimal cut points, this usually results in a loss of power. For instance, dichotomizing a normally distributed continuous variable can be equivalent to reducing the study population by a third [216]. While these properties are well known among statisticians, we have found no clinical articles that implemented splines to model age for HRQoL outcomes.

5.2.3 Limitations

Our study is limited by the small amount of data per patient. Expanding the routine follow-up program with more questions to provide more detailed background information may enhance the regression model but it will jeopardize response rates, thus hurting the generalizability. The follow-up is also only 1 year after surgery; however previous studies have shown surprisingly little change beyond 1 year [206], and the mid-term HRQoL seem consistent [217].

It is important to remember that we looked at the chronological age, not the biological age. Using a biological marker for age may improve the accuracy of the model as the life span of humans is heterogeneous [218]. This is outside the scope of a registry study but may be plausible in a smaller study. Similarly, if a more detailed co-morbidity adjustment than the Charnley classification was applied, the effect of age may be further attenuated and even more difficult to detect.

5.3 STUDY III

5.3.1 Discussion of results

We show that Charnley class B is an independent risk factor and should not be grouped with class A. Furthermore the patient-reported Charnley class B does not benefit from subdivision when assessing HRQoL [138], [143].

There is also a sex difference in Charnley class C's impact for on the EQ-5D index between men and women, where women gained less than men. The mobility dimension seemed to be the main cause behind this interaction, where 25% more men than women improved 1 year after surgery. A possible explanation may be that Charnley class is closely correlated to postoperative ambulation [219], and that women have a different rehabilitation pattern [220]. It is also conceivable that women interpret class C differently, i.e. that the men in class C and women in class C are not comparable. As almost half of all women belonged to Charnley class C we believe that this is an interesting sub-population to further study.

Interestingly, sex differences were not seen for the EQ VAS outcome, possibly due to a different impact of mobility on HRQoL between the sexes. Furthermore, this difference between the EQ-5D and the EQ VAS illustrates how the 5 dimensions can capture different qualities than the EQ VAS, adding value to the overall patient understanding.

The Charnley classification intends to categorize patients according to their walking ability and contralateral hip disease. Consistent with this intention, the mobility and pain/discomfort dimensions in the EQ-5D index were the dimensions most affected by the self-categorized Charnley classification. Perhaps less consistent with the intention was the sizeable change between the pre-operative Charnley class and the class 1 year later. Worth noting is that a considerable group crossed over into class A from both the B and C classes, suggesting that classification based on a patient-administered questionnaire is less rigid than the original intention [47]. Conversely, patients did seem to classify themselves correctly since those who had their contralateral hip operated

soon after rarely reported class A. This is further supported by the lack of improvement when subdividing the B-class according to prior hip arthroplasties. Interestingly, the few patients with a contralateral THR who reported class B, performed worse even after adjusting for pain and pre-operative HRQoL. This could both be due to fear of surgery or genetic predisposal to postoperative pain [221]. We conclude that even though there is a difference from the original intention, patients seem to have a good understanding of their physical health, thus possibly explaining why this self-categorized Charnley classification is such a strong predictor.

To our knowledge, no other large cohort studies on Charnley classification go into the details of the classification's impact on HRQoL. In order to keep our message simple we did not investigate other important metrics such as pain and satisfaction that also are included by SHAR's PROM program. Generic measures such as the EQ-5D index are interesting metrics as they combine many dimensions of health. While these measures are not as sensitive as disease specific measures [139], they do contain common osteoarthritis characteristics such as pain, activity of daily life, and mobility, while at the same time retaining the option of detecting unexpected correlations in other dimensions such as anxiety and depression. In large cohort studies, the study group size compensates for the lack of sensitivity, while preserving the ability to detect unexpected correlations.

The estimated outcomes 1 year after surgery for different Charnley classes suggests that some patients will end up with values lower than their pre-operative level. This estimation applies to the average-aged patient with the most common pre-operative health state according to EQ-5D/EQ VAS. Shifting the references to a lower pre-operative EQ-5D index/EQ VAS or to younger patients would decrease the risk of not improving. We chose this presentation of combined estimates instead of single estimates to better illustrate different quantiles.

5.3.2 Strengths

The main strength of our study is the large nationwide cohort with a good response rate to the survey and small differences between respondents and non-respondents. We believe that the selected cohort is representative, while minimizing confounders such as early reoperations, recall bias due to contralateral surgery, and misclassification of surgery indication.

We chose in this study to exclude those that died within 1.5 years. We believe that HRQoL is interesting only if a patient survives and can benefit from the gained HRQoL. Furthermore if a patient in Charnley class C is advised not to go

through with surgery both due to risk of dying and poor HRQoL outcome, there will in a sense be a double penalty.

5.3.3 Limitations

Our main limitation is the scarcity of information per patient. The registry has during this period started collecting interesting confounders such as BMI and ASA class. In future studies these will most likely be included in the models. The Charnley classification can also be strongly influenced by simultaneous knee osteoarthritis and other lower extremity disabilities that we have not been able to adjust for.

5.4 STUDY IV

5.4.1 Discussion of results

We found no significant differences between Denmark and Sweden in how age, sex and comorbidity status influence the level of HRQoL 1 year after surgery. The consistency of these common and basic predictors facilitates pooling and comparisons between the two countries.

There may be multiple explanations for Danish patients who report higher HRQoL 1 year postoperatively than Swedish. First, population studies have shown that Danish people in general perceive their HRQoL to be higher than Swedes do [222], [223]. Assuming the EQ-5D values for Danish patients are generally higher than for Swedish patients, this would partly explain the observed difference among the THR population. Second, differences in fixation method, implants used, and surgical technique may play a role for the PROMs. For instance, the posterior approach is more common in Denmark, and previous work indicates that a posterior approach is associated with better PROMs compared to the direct lateral approach [224], [225]. Third, the difference in mean age at surgery indicates that Danish patients have surgery at an earlier stage of hip disease. Although age is included in the regression model, there may be residual confounding due to age being a proxy for stage of hip disease. Surgery at an earlier stage of disease implies a preservation of HRQoL after hip replacement. Conversely, surgery at a late stage of disease may reduce the possibility of reaching the expected level of HRQoL for the particular age group

5.4.2 Strengths

The large study population that combines a nationwide Swedish THR population and a randomly selected nationwide Danish THR population contributes to the

strength of this study. Reoperations have been excluded, which limits the effect of potential confounders related to implant survival.

5.4.3 Limitations

The cross-sectional nature of the study limits comparisons of outcomes between the two countries, and only a limited set of variables have been included in the models. Pre-operative PROMs were not available for the Danish patients and thus comparisons taking baseline levels into account were not possible.

Furthermore, the Danish population was much smaller than the Swedish and chance may affect Danish data to a greater extent than Swedish. Including a larger Danish population may result in significant differences between countries in how the investigated predictors affect the outcomes. However, the clinical relevance of such small differences can be questioned.

It is also plausible that some of the differences come from the minor difference in recruited years. The consistency in the EQ-5D index over time in the Swedish THR population suggests though that this effect should be negligible.

5.5 STUDY V

5.5.1 Discussion of results

Our analyses found that the use of antidepressants for depression significantly diminished pain reduction, inhibited improvement in HRQoL, and negatively influenced THR. In this study population, the antidepressant users were worse off before and 1 year after THR. Patients receiving pharmacological treatment for depression or pain syndrome with antidepressant medications perceived their overall HRQoL and pain to improve less than the other group.

Somewhat unexpected, the influence of antidepressants on outcomes was independent of the response to the anxiety/depression dimension of the pre-operative EQ-5D questionnaire. The lack of interaction between the anxiety/depression dimension and treatment is noteworthy, i.e. a patient that reported no anxiety/depression (responding to the treatment) does not differ from one who reports anxiety depression. This could possibly relate to the concerns regarding loss of antidepressant efficacy over the last decades, where the file drawer effect, i.e. negative studies go unpublished, and widening of indications may be part of the explanation [226]. The widening of indications is partly supported by the six-fold rise in sale of antidepressants in Sweden 1990 to 2002 [227]. Despite this increase, studies have found that only one in four of with depression requiring treatment actually receive antidepressants [228]. This

proportion is similar to the treatment prevalence among the patients reporting extreme problems in the EQ-5D anxiety/depression dimension.

The large improvement in the anxiety/depression dimension suggests that many patients were experiencing pre-operative anxiety. The smaller improvement among the patients treated with antidepressants supports this notion.

Duivenvoorden et al. have reported similar improvement rates using the Hospital Anxiety and Depression Scale [229]. As pre-operative anxiety may respond to patient education [230] there may be room for improvement regarding this parameter, although Hossain et al. could not find any decrease in satisfaction related to pre-operative distress in 448 THR patients [231].

We also found that discontinuation of antidepressant treatment had a negative influence on pain and satisfaction. The effect size was similar to the group with pain syndrome. Interestingly, discontinuation of treatment was the only treatment change which influenced any of the outcomes. To our knowledge, this has never been shown and can have direct clinical implications; simultaneous discontinuation of antidepressant treatment in conjunction with major surgery may be unwise.

A possible explanation for the depression effect could be the pain-depression link. Depression in patients with primary osteoarthritis has been correlated with sensitivity to pain and decreased ability to cope with the disease [232]. Depression is also known to be a catalyzer of pain, and addressing it early seems to improve outcomes [233], [234]. Because depressive disorders are common in patients with chronic pain, education about continued activity should be considered not only as a non-surgical treatment for primary osteoarthritis, but also as a preventative measure against the development of depression [235].

5.5.2 Strengths

The strength of this study stems from the use of the nation-wide register of primary osteoarthritis patients receiving THR. In conjunction with the nation-wide register of prescription drug purchases, we were able to create a homogenous study population that excluded important confounding groups. Specialized code allowed the identification of various stages and severity of disease for patients taking antidepressant medications. Finally, rigorous validation minimized the risk of invalid data.

5.5.3 Limitations

A limitation of this investigation was that the indication for the antidepressant medications in 37% of the patients could not be identified. These patients were simply classified as users of antidepressants, but the indication was unknown. A

review of a Swedish subpopulation found that antidepressants were primarily prescribed for depression (66%), followed by anxiety (14%), and treatment of pain (11%) [236].

Another limitation is that we cannot know with certainty whether patients took their prescribed medication. However, there is strong indication that the majority of patients were following the treatment set by their physician because they obtained their prescribed medication at least twice during the observation period. The variety of clinical indications for antidepressant prescription and the unknown compliance with treatment may limit our understanding of how psychological distress truly influenced PROMs in this population. Nevertheless, the majority of the antidepressant patients were clinically diagnosed with specific psychological conditions, and these conditions influenced PROMs after surgery.

The aim of this study was to explore pre-operative factors associated with poor patient outcomes, for this reason we did not explore the novel, continued, or discontinued use of antidepressants after surgery. Future studies that focus on these postoperative factors will inform how antidepressant treatment continues to influence PROMs after surgery.

5.6 GENERAL DISCUSSION

5.6.1 Data quality

All studies in this thesis are based upon the data in the Swedish Hip Arthroplasty Register. The register is renowned for its high data quality with a 100% of hospital coverage and a completeness of individuals about 98% for THR [14]. Primary operations are registered using a decentralized structure where each department enters the data. Data capture for reoperations is centralized, all medical records regarding the reoperations are sent to trained coordinators at the SHPR who then enter these into the database. Every year each department is also sent a list of primary and secondary interventions in order to review and report back any inconsistencies.

Since the start improvements in quality have been continuously put in place. Söderman et al identified 10% of missing revisions in which 2 hospitals constituted almost half of the missing data [237], and since 2007 an annual quality control is performed by cross-matching data with the NPR. As of 2013 there is also a monitoring system in which departments are visited by coordinators that compare the register data against the medical records.

Even with this control we recently found that one third of all re-operations due to infection were missing [23]. There is also an ongoing investigation regarding

periprosthetic fractures due to concerns that fractures of Vancouver type A and C are underreported [25]. This lack of reporting could primarily affect the results in study I with falsely low estimates. The effect is probably most relevant for early, acute re-operations, as they are frequently managed by multiple surgeons during non-office hours.

The PROM program has a pre- and postoperative response rate of around 80-90% [93]. It is currently based upon pen and paper forms that are later filled in by secretaries at different clinics. It is possible that this causes some faulty data input causing a non-differential misclassification bias, most likely diminishing the estimates [238]. Efforts into digitalizing the program using web-based questionnaires were abandoned due to low response rates [239]. While the completeness is less than for re-operations, we have found little difference in the PROM results when applying multiple imputations.

Similarly, the NPR has a high sensitivity and specificity for primary diagnoses [173], and combined with the Swedish personal identity number [240] the matching error between the registers is negligible. The misclassification bias is generally believed to be below 10% [241]. While primary ICD-codes are commonly correct, there may be less specificity among the secondary ICD-codes. This is supported by the increase in comorbidities throughout the study period seen in study I. We believe that the increase in registration is mostly due to monetary incentives for registering as the age distribution for the population has not changed over the period. This possible detection bias may weaken the estimates from study I and IV, although the ICD-10 subgroup analysis in study I showed only small differences.

The data capture for the PDR is fully automated by the pharmacies, and missing data should therefore be negligible. While there are no antidepressants currently sold over-the-counter, missing data in study IV can still theoretically occur from herbal products [242] or medications used in hospitals or from nursing home drug storerooms [243].

5.6.2 Code quality

Equally important to data quality is the code quality. The Python scripts that extracted many of the variables for this thesis amount to 5,199 lines of code (289,955 bytes). Coding errors are inevitable, depend on code complexity, and are estimated by some to be around 10 per 1,000 lines of code [244]. As the effects of coding errors can be devastating, all code used in the thesis has been subjected to both white and black box testing [245].

White box testing uses knowledge of the software's design in order to test its different components. This was deployed for the comorbidity counts where unit

tests used the previously published SAS-code as a template. For the R-packages unit tests were deployed through the standard testthat-package [246].

Black box testing is the opposite of white box testing in which no knowledge of the software design is used. The black box tests were performed using small random subsets of the original data. The results were then manually validated in order to find any unexpected bugs. The black box validation in study IV, the study with the most intense coding, was performed by Kristina Annerbrink in order to avoid creator bias.

While few studies actively report their coding quality strategies, these are a source of study errors. In macro-economics the raw data is often available and the recent debacle regarding Reinhart and Rogoff's famous optimal debt to gross domestic product (GDP) ratio [247] verifies this notion. The lack of code transparency has been noted in the statistical community [248] and this is also why we have openly published the central portion of the code used in this thesis on [gforge.se](https://github.com/gforge/se).

5.6.3 Patient selection

Due to lack of detailed patient knowledge we have tried to select patient cohorts as clean as possible in each study. In study I we excluded rare implants in order to limit the risk of patients participating in implant studies. The Australian registry showed that none out of 167 new hip implants performed better than the old, and almost one third had a higher revision rate [38]. Thus it is highly likely that small series of implants, often restricted to studies, will perform worse. As studies frequently exclude patients with multiple comorbidities, this can be viewed as a form of confounding by indication [249].

In study II-V we excluded patients that had an early re-operation or death. There is a high probability that motivation for reporting differs in these groups, patients may be too sick to report, experience a general resignation towards health-care, frustration, and other feelings not intended to be captured by the HRQoL-form. There is also a risk that including these would result in a sense of double penalty for the patients at increased risk for re-operations or death.

5.6.4 Minimal clinically important difference

As register studies contain thousands of patients, small differences with little clinical relevance may reach statistical significance. For re-operations, risk differences of 3% are often considered clinically relevant [104] but for PROMs the interpretation is less obvious.

Jaeshke et al [250] defined minimal clinically important difference (MCID) as:

“... the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management”

The MCID is a tool for clinicians to discern what treatment effects should encourage them to change their practice. Some differentiate between the minimally important distance (MID) and MCID [251] but for this discussion they will be used interchangeably. There are currently 2 main methodologies for estimating the MCID: distribution based methods and anchor based methods.

5.6.4.1 Distribution based methods

The general idea for distribution based methods is that the variance in the measurement relates to the MCID. Norman et al looked at 6 different MCID-studies and noted that the MCID frequently coincides with one half of the standard deviation (SD). They therefore suggested that this may be an alternative metric when no established MCID exists [252].

A more theoretical approach is to use the standard error of measurement (SEM); this is the error that an individual makes when re-estimating his/her status according to the current outcome. Wyrwich et al found that 1 SEM frequently corresponds to the MCID [253].

Both tools relate to the stability of the measurements, and suggest that a poor, unreliable instrument would demand a higher MCID than an accurate one. This can be viewed as if an observer were to identify a difference; the improvement needs to be larger than the noise. This further suggests that the MCID is the same no matter the intervention. If the intervention improves a patient's well-being in a way similar to the measurement's design, this may be reasonable. Conversely, if the measurement's design is not tailored to the intervention, this assumption is flawed (Figure 11). HRQoL is arguably not tailored to measuring THR outcomes, and using the distribution of the measurement would therefore not be a sensible option.

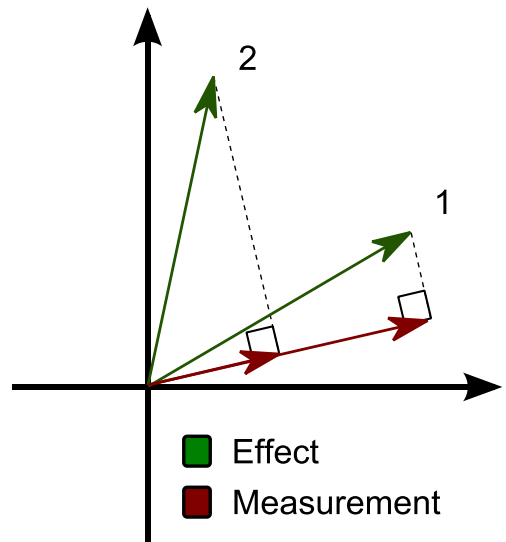


Figure 11: The two effect arrows are exactly the same length and the measurement size depends on the similarity between the effect and the measurement. Effect 1 indicates a larger similarity and thus its measurement will appear larger than effect 2.

5.6.4.2 *The anchor based method*

Anchor based methods are currently the gold standard for deciding the MCID for a measurement [254]. The anchor refers to an external criterion that defines success for the intervention, e.g. patient's experience of improvement or ability to return to work. By comparing the success group with the other, an estimate of the MCID can be attained.

The patient's experience is most commonly applied where a global rating of change (GRoC) scale is applied [255]. A GRoC scale contains a sequence of questions that aims to separate patients that have done worse from those who have improved. The questions are ordered and can range from 3 to 101 alternatives. A simple 5-alternative GRoC scale could be [256]:

1. Got a lot better
2. Got a little bit better
3. Unchanged
4. Got a little worse
5. Got a lot worse

The MCID from the scale above would be the difference between the 3rd and the 2nd group. Note that the difference between the 4th and the 3rd group above is likely different than the one between the 3rd and the 2nd group.

There is currently no consensus on choosing the number of groups in a GRoC scale and which ones should be compared [254]. Preston and Colman showed that a test with 101 categories has reasonable reliability [257], although the large number may lead to answers of poorer quality [258]. Preston and Colman also suggested that about 10 alternatives are optimal. Note that depending on the number of groups, the distance between neighboring ones differs, thus a GRoC scale with more groups will generate smaller MCIDs.

Walters and Brazier [259] compared MCIDs for the EQ-5D with the SF-6D index for 8 different cohorts. They used a 5-level GRoC scale where they compared those that performed slightly better to those that performed slightly worse. Apart from concluding that the MCID was different for the EQ-5D and the SF-6D indices, they found a large variety between different patient cohorts, e.g. MCID for the EQ-5D index in the knee osteoarthritis was 0.1 while in chronic obstructive pulmonary disease it was 0.0. The measurement and similarity of effect seem thus also to affect the MCID for anchor based methods (Figure 11).

Kerujente et al. reported recently results for SF-36 after joint arthroplasties using an alternative anchor method; by asking patients if they would be willing to undergo surgery again, they could compare patient's estimates and deduce a clinically important difference [94]. Although this is an interesting approach, it

does not truly estimate the minimal value. It seems also quite likely that patients will suffer from a recall bias perhaps larger than a satisfaction anchor. As the authors state, the main benefit with this approach is that it is difficult to identify patients for successful interventions such as arthroplasties who are only modestly satisfied, i.e. almost everyone will reply very satisfied, and the answer of interest for the MCID calculation will be rare, requiring very large cohorts.

5.6.4.3 *Risk and MCID*

Apart from MCID being dependent on diagnoses, it also most likely relates to the intervention risk. Estimating the MCID for patients going through with a THR may be affected both by high-risk high-demand and an IKEA effect bias [260], where the proudness of patients accomplishment, going through with the surgery and the intense rehab, cause them to overestimate their improvement, i.e. pushing the MCID downwards. In addition, the high-risk high-demand bias may cause patients to expect a large improvement as the risk associated with surgery is perceived to be high, thus pushing for higher MCID estimates. It is difficult to deduce which of these is dominant, but it is likely that at least one of them will impact the MCID estimate.

5.6.4.4 *Changing management or estimating base level*

Studies I-V all focus on patient factors not subjected to change as the definition of MCID suggests. It is therefore difficult to apply the MCID concept to patient factors; they are mostly useful in assessing patients' base level and their potential to achieve MCID with the surgery. They can also help us identify risk groups, such as the women in Charnley class C identified in study III.

In study V we did find a potential factor that could cause a change in management. Although our finding needs further investigating, we believe that it may be sensible to avoid discontinuation of antidepressant soon before surgery, and that the risk is almost negligible unless the patient is experiencing some side effects. The risk in considering a change in management is important, as suggested by the MCID's definition:

$$MCID \leq \frac{\textit{Estimated improvement}}{\textit{Risk}}$$

If the risk increases, the effect size from the change should also be greater. For instance changing a paragraph in pre-operative patient information would require an estimated effect equal to the MCID. Conversely, switching surgical technique to minimally invasive surgery would require a considerably larger improvement.

5.7 BALANCING COMPLEXITY

One of the main aims with statistics is summarizing and presenting a comprehensive outline of the relationships in the data. For large data sets with multiple outcomes, a conflict can arise between the overview and the details, inhibiting wider understanding of the results [261]. E.g. in study II the main emphasis is on the negative impact of age after the late 60s, although the supplemental analysis exhibits more patterns available in the data. The simpler form of presenting the age relationship was chosen in order to introduce the complex idea of non-linear relationships.

Apart from opting for simpler relations, the studies in this thesis rely heavily on graphics. Graphs can present complex relationships and allow comparison between multiple values [262, p. 87]. The difficulties in understanding patient-reported outcomes have been approached by visually contrasting different groups. Contrasting them is an alternative approach to the MCID; by using references familiar to orthopaedic surgeons it is possible to convey the effect of the studied variable even though the outcome measure may be unfamiliar to him/her.

6 CONCLUSIONS

6.1 STUDY I

We failed to validate any of the scores for re-operations after total hip arthroplasties, although the Elixhauser score may be useful for estimating the comorbidities relevant for the risk of re-operation within 2 years. The comorbidities were rare and the associated risk increase was small, thereby it is undoubtedly best suited when studying large samples and not individual patients.

6.2 STUDY II

There is a non-linear relationship for age and HRQoL in patients receiving THR that results in residual confounding if treated as a simple linear term or categorically in the regression. The implication of this is important, as age is a common confounder for which adjustment is necessary. The same applies to the preoperative EQ-5D index and EQ VAS.

6.3 STUDY III

The self-administered Charnley classification is a reliable instrument with interesting properties easy to utilize in everyday clinical practice. There is also strong evidence that women in Charnley class C fail to improve their mobility as much as men.

6.4 STUDY IV

There are clear similarities in how basic predictors influence patient-reported outcomes in patients with THR in Sweden and Denmark. Apparent cultural, social, and other such differences among these countries are not reflected in these predictors.

6.5 STUDY V

Antidepressants have a negative influence on PROMs 1 year after THR independent of the pre-operative EQ-5D anxiety/depression dimension. We also found that discontinuation with treatment prior to surgery is associated with poorer outcomes in the pain and satisfaction dimensions.

7 FUTURE PERSPECTIVES

The variety of reported results suggests that there is a need for more structured reporting and analysis. While many studies have their own unique angle, these should most likely be accommodated with standard metrics. It would be beneficial if supplements contained basic patient factors such as sex and age accompanied by unadjusted estimates, adjusted only according to a standard model, and fully adjusted according to the study protocol. This is an important piece as comparing effect estimates between populations permits to move the debate beyond simple factors such as gender and age.

In large registry studies the use of categories is a misguided desire for simplification. Simplification can be achieved through tables that contrast non-linear estimates and that can easily substitute the need for categories (see Table 2 under the results for study II). If cut-off values are of interest, an alternative can be piecewise linear regression [263], [264]. With sample sizes well above 100,000 patients and small effect estimates, the choice of statistical methods is increasingly important.

Understanding complex systems can further be enhanced by interactive graphics. Data-Driven Documents (D3) have gained a lot of interest and are currently often used by major newspapers such as the New York Times [265], [266]. Today, 20 years after the journals started publishing PDFs online, the next step of interactive graphics could revolutionize the understanding of research.

8 ACKNOWLEDGEMENTS

I want to thank the many people that have made this possible:

Olof Sköldenberg, principal supervisor: You have supported me tremendously with your enthusiasm and constantly positive attitude. Your ability to quickly grasp new concepts has been invaluable. You have an amazing ability of knowing what a PhD-student needs, you have pushed me and supported me exactly when I needed it the most.

André Stark, co-supervisor: Your confidence and belief in me has been truly inspiring. I have always felt welcome with any question and you have always found time for them.

Göran Garellick, co-supervisor: You did an amazing job with getting me started; understanding the registry is challenging and I would have been lost without you. You have also throughout the process supported me, no matter how unconventional my ideas were.

Johan Kärrholm, co-supervisor: Your insights and knowledge have been very helpful. You have always kept an open mind with a great interest for new things; I have truly enjoyed our shared joy for R and epidemiological discussions.

Ola Rolfson, co-author: Your profound understanding of the patient-reported outcomes has been crucial. You have also always supported me in my ideas and provided important input.

Meridith Greene, co-author: Your help with phrasing and outside perspective on orthopaedics has been very constructive.



Swedish Hip Arthroplasty Register

Kristina Annerbrink, co-author: Thank you for being a wonderful friend and also for helping out on such short notice with study V.

Paolo Frumento, co-author: Thank you for introducing me to statistics at a more fundamental level.

Henrik Malchau, co-author: Thank you for interesting conversations and input.

Other co-authors: Thank you all for your help and input.

Peter Herberts, founder of SHAR: Thank you for all the interesting conversation about the registry's past and future.

Maria Elmberg and everyone at the Clinical Epidemiology Unit who helped with "*Forskarskola i epidemiologi, generation 9*": Your introduction to the fine art of epidemiological research both excellent and extremely useful.

Matteo Bottai, statistician: Thank you for teaching me statistics and showing me its wonders.

Anders Oden, statistician: You introduced me to statistics, with your help I discovered the fascinating non-linear relations.

Szilard Nemes, statistician: Thank you for our interesting conversations regarding statistical methods.

Torbjörn Ahl, colleague: Thank you for introducing me to Göran Garellick and making all this possible.

Gustaf Neander, head of the Department of Orthopaedics at Danderyd Hospital: Thank you for always supporting me and providing the time to finish this thesis.

Heléne Sjöo and Paula-Therese Kelly Pettersson, research nurses: You are the backbone of our research at the Department of Orthopaedics. Your constant enthusiasm and willingness to help out is extraordinary.

All my colleagues: I feel privileged working with all of you, not only for your orthopaedic talents but also our wonderful friendship.

All the personnel at the Department of Orthopaedics at Danderyd Hospital: Your ability of keeping track of and helping us doctors is unprecedented. Without you the hospital would fall apart.

Kajsa Erikson, Karin Lindborg, Karin Pettersson, Karin Davidsson, and everyone else at the Swedish Hip Arthroplasty Register: Thank you all for helping out with the practicalities. You are truly the cornerstone of the registry.

Department of Clinical Sciences, Danderyd Hospital: Thank you for believing in this thesis.

Coursera, edX, John Hopkins, and Stanford: Thank you for providing top of the line statistical courses. A special thanks goes to Andrew Ng, Trevor Hastie, Rob Tibshirani, Brian Caffo, and Jeffrey Leek for taking their time to teach outstanding courses in statistics, data analysis, and machine learning.

Everyone at Cross Validated: I want to thank everyone, especially prof. Frank Harrell, for helping me out whenever I ran into questions of statistical nature.

All colleagues and patients in Sweden who meticulously have filled out forms and collected data: Without you this would not have been possible.

Peter and Sofia Moström: Your warmth, friendship, and hospitality have made my trips to Gothenburg extra memorable.

Magdalena Kay, Bogdana, and John Carpenter: Thank you for lending your time and amazing English talent.

Ulli and Ponka Nylund: the best possible in-laws, even though you live too far away.

Maria Gordon, Johan Malmström, Teodor, and Arvid: Thank you Maria for the sibling rivalry, constantly pushing me to become a better person. Thank you Johan for taking such wonderful care of my beloved sister and my two fantastic nephews. I also want to thank Teo for teaching my son everything he will ever need to know about cars, and Arvid for keeping a watchful eye on them both.

My parents: Your unconditional love and support is beyond words.

Filip Gordon: Thank you for closing my laptop lid whenever you can, and urging me to live in the moment.

Daniela Gordon, my wife: You have supported me throughout the thesis, both with invaluable understanding of epidemiology and loving support. You are at the center of my universe; there is nothing more valuable to me than our shared time. I hope we will have even more of it in the future.

9 SAMMANFATTNING PÅ SVENSKA

Bakgrund

Grundläggande patientfaktorers betydelse för risken för omoperation och patientrapporterade utfall efter höftledsplastiker är dåligt kartlagda.

Syften

Målen med denna avhandling var att undersöka:

1. Om samsjuklighetsmått utvecklade för dödlighetsutfall även går att applicera på omoperationer.
2. Hur ålder påverkar hälsorelaterad livskvalitet (HRQoL).
3. Om kön/ålder interagerar med den själv-administrerad Charnley-klassifikationens inverkan på HRQoL .
4. Generaliserbarheten mellan länder av HRQoL effekter.
5. Om antidepressiv användning påverkar patientrapporterade utfall .

Metoder

Alla studier baseras på Svenska Höftprotesregistrets (SHAR) databas. För studie I har data från SHAR korsmatchats med Patientregistret. Genom matchningen kunde 3 olika samsjuklighetsindex beräknas: Charlsons index, Elixhausers index, och Royal College of Surgeons (RCS) Charlson index. De tre indexen jämfördes sedan med hjälp av överlevnadsanalys avseende omoperationer mellan 0 till 2 år och 2 till 12 år.

I studie II och III användes SHARs databas med patientrapporterade utfall där HRQoL mätts med EQ-5D och EQ-VAS. I studie II modelleras ålder med hjälp av linjär regression i kombination med splines. I studie III användes linjär regression med interaktionsvariabler som utvärderades med hjälp av ANOVA-tester, undergruppsanalyser och EQ-5D dimensions-specifika analyser.

I studie IV länkade vi SHARs databas med patientrapporterade utfall till Patientregistret för beräkning av Charlsons index. Denna kohort slogs sedan ihop med en mindre dansk kohort som också länkats till den danska motsvarigheten till Patientregistret för samma beräkning. Därefter undersöktes kön, ålder, samsjuklighetseffekt med hjälp av interaktionstermer såsom i studie III.

I studie V, korsmatchades SHARs databas med patientrapporterade utfall mot Läkemedelsregistret. Vi beräknade användningen av antidepressiva läkemedel med hjälp av reguljära uttryck (regular expressions). Mått för behandlingsefterlevnad, förändring i behandling, och indikation hämtades från förskrivningstexten.

Resultat

Studie I

I den tidigt postoperativa perioden, 0-2 år, visade bara Elixhausers index en signifikant ökad risk med ökad risk för både 1-2 och ≥ 3 sjukdomstillstånd. Den prediktiva C-statistiken under denna period var dock dålig även för Elixhausers index, 0,52. Inget av samsjuklighetsmåten visade sig vara av värde mellan 2-12 år.

Studie II

Ålder uppvisade ett icke-linjärt samband med HRQoL. Åldern hade knappt mätbar inverkan på värdena fram till 60-70 års ålder, varefter den hade en negativ inverkan på både EQ-5D index och EQ-VAS.

Studie III

Vi fann att kvinnor i Charnley klass C hade ett sämre EQ-5D utfall jämfört med män. Denna effekt berodde främst på det faktum att kvinnor inte förbättrades i rörlighetsdimensionen, endast 40 % av kvinnorna blev bättre medan 50 % av männen förbättrades. Ålder interagerade däremot inte med Charnley klass. Vi fann också att klassificeringen utförs bäst utan delning av B-klassen eller aggregerandet av klasser.

Studie IV

Danska patienter hade övergripande högre HRQoL jämfört med svenska patienter (p-värde $< 0,001$). En regressionsanalys kunde dock inte påvisa att koefficienterna för kön, ålder eller Charlsons index skiljer sig åt mellan länderna (EQ - 5D-index p-värde = 0,83 och EQ VAS p-värde = 0,41) 1 år efter THR.

Studie V

Antidepressiva läkemedel användes av 9 % av patienterna (n = 954). Patienter som använder antidepressiva medel rapporterade även sämre HRQoL, mer smärta, och upplevde mindre tillfredsställelse 1 år efter operation. Den negativa effekten av preoperativ antidepressiv användning var oberoende av patient-rapporterade ångest eller depression (mätt enligt EQ-5D) 1 år efter THR. Utsättning av behandlingen var också negativt associerat med smärta och tillfredsställelse efter 1 år.

Slutsatser

Studie I

Vi misslyckades med att validera något av samsjuklighetsmåten för omoperationer efter höftledsprotoser, även om Elixhausers index kan vara användbart för att uppskatta samsjuklighet relevant för reoperationsrisken inom 2 år. Riskökningen associerade samsjuklighet var liten och är utan tvekan bäst lämpad för studier av stora populationer och inte enskilda patienter.

Studie II

Det finns ett icke-linjärt förhållande mellan ålder och hälsorelaterad livskvalitet ett år efter en höftledsplastik; detta innebär att man inte kommer att justera bort all den effekt som en patients ålder medför om den behandlas som en enkel linjär variabel eller delas upp i kategorier. Innebörden av detta är viktig eftersom ålder är en vanlig confounder. Detsamma gäller för de preoperativa HRQoL värdena.

Studie III

Den självadministrerade Charnleyklassifikationen är ett tillförlitligt instrument med intressanta egenskaper som är lätt att använda i den kliniska vardagen. Det finns dessutom mycket som talar för att kvinnor i Charnley klass C misslyckas med att förbättra sin rörlighet lika mycket som män.

Studie IV

Det finns klara likheter i hur grundläggande prediktorer påverkar patientrapporterade utfall efter höftledsplastiker mellan svenska och danska patienter. Skenbara kulturella, sociala och andra sådana skillnader mellan dessa länder tycks inte återspeglas i dessa prediktorer.

Studie V

Antidepressiva medel har en negativ inverkan på patientrapporterade utfall 1 år efter en höftledsplastik, oberoende av den preoperativa ångest eller depression mätt enligt EQ-5D:s femte dimension. Vi fann också att om behandlingen avbröts före kirurgi var detta förknippat med sämre resultat gällande smärtlindring och tillfredsställelse.

10 REFERENCES

- [1] F. H. Albee, “Arthritis deformans of the hip.: A preliminary report of a new operation.,” *J. Am. Med. Assoc.*, vol. L, no. 24, pp. 1977–1980, Jun. 1908.
- [2] H. Osmond-Clarke, “Half a Century of Orthopaedic Progress in Great Britain,” *J. Bone Joint Surg. Br.*, vol. 32-B, no. 4, pp. 620–675, Nov. 1950.
- [3] I. D. Learmonth, C. Young, and C. Rorabeck, “The operation of the century: total hip replacement,” *Lancet*, vol. 370, no. 9597, pp. 1508–19, Oct. 2007.
- [4] J. Charnley, “Arthroplasty of the hip: A new operation,” *The Lancet*, vol. 277, no. 7187, pp. 1129–1132, May 1961.
- [5] A. Oksuzyan, B. Jeune, K. Juel, J. W. Vaupel, and K. Christensen, “Changes in hospitalisation and surgical procedures among the oldest-old: a follow-up study of the entire Danish 1895 and 1905 cohorts from ages 85 to 99 years,” *Age Ageing*, vol. 42, no. 4, pp. 476–481, Jul. 2013.
- [6] A. B. Pedersen, J. A. Baron, S. Overgaard, and S. P. Johnsen, “Short- and long-term mortality following primary total hip replacement for osteoarthritis: A Danish nationwide epidemiological study,” *J. Bone Joint Surg. Br.*, vol. 93-B, no. 2, pp. 172–177, Feb. 2011.
- [7] J. Kärrholm, G. Garellick, and P. Herberts, “Mortality after total hip replacement,” in *Annual Report 2006. Swedish Hip Arthroplasty Register*, 2007, pp. 75–77.
- [8] S. G. Memtsoudis, M. Pumberger, Y. Ma, Y.-L. Chiu, G. Fritsch, P. Gerner, L. Poultsides, and A. G. D. Valle, “Epidemiology and risk factors for perioperative mortality after total hip and knee arthroplasty,” *J. Orthop. Res.*, vol. 30, no. 11, pp. 1811–1821, Nov. 2012.
- [9] H. D. Taylor, D. A. Dennis, and H. S. Crane, “Relationship between mortality rates and hospital patient volume for medicare patients undergoing major orthopaedic surgery of the hip, knee, spine, and femur,” *J. Arthroplasty*, vol. 12, no. 3, pp. 235–242, Apr. 1997.
- [10] O. Sköldenberg, *Uncemented femoral stems. Studies on periprosthetic bone remodelling and prevention of bone loss in total hip arthroplasty*. Stockholm, Sweden: Karolinska Institutet, 2010.
- [11] P. Cram, X. Lu, P. J. Kaboli, M. S. Vaughan-Sarrazin, X. Cai, B. R. Wolf, and Y. Li, “Clinical Characteristics and Outcomes of Medicare Patients Undergoing Total Hip Arthroplasty, 1991-2008,” *JAMA J. Am. Med. Assoc.*, vol. 305, no. 15, pp. 1560–1567, Apr. 2011.

- [12] L. P. Hunt, Y. Ben-Shlomo, E. M. Clark, P. Dieppe, A. Judge, A. J. MacGregor, J. H. Tobias, K. Vernon, and A. W. Blom, "90-day mortality after 409 096 total hip replacements for osteoarthritis, from the National Joint Registry for England and Wales: a retrospective analysis," *The Lancet*, vol. 382, no. 9898, pp. 1097–1104, Sep. 2013.
- [13] A. B. Pedersen, F. Mehnert, H. T. Sorensen, C. Emmeluth, S. Overgaard, and S. P. Johnsen, "The risk of venous thromboembolism, myocardial infarction, stroke, major bleeding and death in patients undergoing total hip and knee replacement A 15-year retrospective cohort study of routine clinical practice," *Bone Jt. J.*, vol. 96-B, no. 4, pp. 479–485, Apr. 2014.
- [14] G. Garellick, J. Kärrholm, C. Rogmark, O. Rolfson, and P. Herberts, *Annual Report 2011 Swedish Hip Arthroplasty Register*. 2012.
- [15] S. A. Lie, L. B. Engesaeter, L. I. Havelin, H. K. Gjessing, and S. E. Vollset, "Mortality after total hip replacement: 0-10-year follow-up of 39,543 patients in the Norwegian Arthroplasty Register," *Acta Orthop.*, vol. 71, no. 1, pp. 19–27, Jan. 2000.
- [16] L. Pulido, J. Parvizi, M. Macgibeny, P. F. Sharkey, J. J. Purtill, R. H. Rothman, and W. J. Hozack, "In Hospital Complications After Total Joint Arthroplasty," *J. Arthroplasty*, vol. 23, no. 6, Supplement, pp. 139–145, Sep. 2008.
- [17] L. M. de Vries, M. C. J. M. Sturkenboom, J. A. N. Verhaar, J. H. Kingma, and B. H. C. Stricker, "Complications after hip arthroplasty and the association with hospital procedure volume," *Acta Orthop.*, vol. 82, no. 5, pp. 545–552, Oct. 2011.
- [18] R. Jimenez-Garcia, M. Villanueva-Martinez, C. Fernandez-de-las-Penas, V. Hernandez-Barrera, A. Rios-Luna, P. C. Garrido, A. L. de Andres, I. Jimenez-Trujillo, J. S. R. Montero, and A. Gil-de-Miguel, "Trends in primary total hip arthroplasty in Spain from 2001 to 2008: Evaluating changes in demographics, comorbidity, incidence rates, length of stay, costs and mortality," *BMC Musculoskelet. Disord.*, vol. 12, p. 43, Feb. 2011.
- [19] J. A. Singh, M. R. Jensen, W. S. Harmsen, S. E. Gabriel, and D. G. Lewallen, "Cardiac and thromboembolic complications and mortality in patients undergoing total hip and total knee arthroplasty," *Ann. Rheum. Dis.*, vol. 70, no. 12, pp. 2082–2088, Dec. 2011.
- [20] J. Kärrholm, "The Swedish Hip Arthroplasty Register (www.shpr.se)," *Acta Orthop.*, vol. 81, pp. 3–4, Feb. 2010.
- [21] L. Pulido, E. Ghanem, A. Joshi, J. J. Purtill, and J. Parvizi, "Periprosthetic Joint Infection: The Incidence, Timing, and Predisposing Factors," *Clin. Orthop.*, vol. 466, no. 7, pp. 1710–1715, Jul. 2008.

- [22] H. Dale, A. M. Fenstad, G. Hallan, L. I. Havelin, O. Furnes, S. Overgaard, A. B. Pedersen, J. Kärrholm, G. Garellick, P. Pulkkinen, A. Eskelinen, K. Mäkelä, and L. B. Engesaeter, "Increasing risk of prosthetic joint infection after total hip arthroplasty," *Acta Orthop.*, vol. 83, no. 5, pp. 449–458, Oct. 2012.
- [23] J. V. Lindgren, M. Gordon, P. Wretenberg, J. Kärrholm, and G. Garellick, "Incidence of deep periprosthetic joint infection following primary total hip arthroplasty: Including a new method of evaluation and surveillance," *Acta Orthop.*, vol. 85, p. ??–??, 2014.
- [24] G. J. Della Rocca, K. S. M. Leung, and H.-C. Pape, "Periprosthetic Fractures: Epidemiology and Future Projections," *J. Orthop. Trauma June 2011*, 2011.
- [25] H. Lindahl, H. Malchau, P. Herberts, and G. Garellick, "Periprosthetic Femoral Fractures: Classification and Demographics of 1049 Periprosthetic Femoral Fractures from the Swedish National Hip Arthroplasty Register," *J. Arthroplasty*, vol. 20, no. 7, pp. 857–865, Oct. 2005.
- [26] R. E. Cook, P. J. Jenkins, P. J. Walmsley, J. T. Patton, and C. M. Robinson, "Risk factors for Periprosthetic Fractures of the Hip: A Survivorship Analysis," *Clin. Orthop.*, vol. 466, no. 7, pp. 1652–1656, Jul. 2008.
- [27] R. Y. Woo and B. F. Morrey, "Dislocations after total hip arthroplasty," *J. Bone Joint Surg. Am.*, vol. 64, no. 9, pp. 1295–1306, Dec. 1982.
- [28] B. Kristiansen, L. Jørgensen, and P. Hölmich, "Dislocation following total hip arthroplasty," *Arch. Orthop. Trauma. Surg.*, vol. 103, no. 6, pp. 375–377, Apr. 1985.
- [29] M. Soong, H. E. Rubash, and W. Macaulay, "Dislocation After Total Hip Arthroplasty," *J. Am. Acad. Orthop. Surg.*, vol. 12, no. 5, pp. 314–321, Sep. 2004.
- [30] M. Khatod, T. Barber, E. Paxton, R. Namba, and D. Fithian, "An Analysis of the Risk of Hip Dislocation with a Contemporary Total Joint Registry," *Clin. Orthop.*, vol. 447, pp. 19–23, Jun. 2006.
- [31] G. Garellick, C. Rogmark, J. Kärrholm, and O. Rolfsson, *Svenska Höftprotesregistret Årsrapport 2012*, vol. 2012. 2013.
- [32] S. S. Jameson, D. Lees, P. James, I. Serrano-Pedraza, P. F. Partington, S. D. Muller, R. M. D. Meek, and M. R. Reed, "Lower rates of dislocation with increased femoral head size after primary total hip replacement: A five-year analysis of NHS patients in England," *J. Bone Joint Surg. Br.*, vol. 93-B, no. 7, pp. 876–880, Jul. 2011.
- [33] B. C. Werner, "Instability after total hip arthroplasty," *World J. Orthop.*, vol. 3, no. 8, p. 122, 2012.

- [34] L. Pulido, C. Restrepo, and J. Parvizi, "Late Instability Following Total Hip Arthroplasty," *Clin. Med. Amp Res.*, vol. 5, no. 2, pp. 139–142, Jun. 2007.
- [35] J. W. Harkess and J. R. J. Crockarell, "Arthroplasty of the hip, Complications," in *Campbell's operative orthopaedics.*, W. C. Campbell, S. T. Canale, and J. H. Beaty, Eds. Philadelphia, PA: Mosby/Elsevier, 2008, pp. 392–429.
- [36] D. R. Haynes, T. N. Crotti, A. E. Potter, M. Loric, G. J. Atkins, D. W. Howie, and D. M. Findlay, "The osteoclastogenic molecules RANKL and RANK are associated with periprosthetic osteolysis," *J. Bone Joint Surg. Br.*, vol. 83, no. 6, pp. 902–911, Aug. 2001.
- [37] D. R. Haynes, "Bone lysis and inflammation," *Inflamm. Res. Off. J. Eur. Histamine Res. Soc. Al*, vol. 53, no. 11, pp. 596–600, Nov. 2004.
- [38] R. Anand, S. E. Graves, R. N. de Steiger, D. C. Davidson, P. Ryan, L. N. Miller, and K. Cashman, "What Is the Benefit of Introducing New Hip and Knee Prostheses?," *J. Bone Jt. Surg.*, vol. 93, no. Supplement 3, pp. 51–54, Dec. 2011.
- [39] T. A. Brennan, L. L. Leape, N. M. Laird, L. Hebert, A. R. Localio, A. G. Lawthers, J. P. Newhouse, P. C. Weiler, and H. H. Hiatt, "Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I," *N. Engl. J. Med.*, vol. 324, no. 6, pp. 370–376, Feb. 1991.
- [40] E. Bosma, E. J. Veen, and J. A. Roukema, "Incidence, nature and impact of error in surgery," *Br. J. Surg.*, vol. 98, no. 11, pp. 1654–1659, Jun. 2011.
- [41] R. M. Wilson, W. B. Runciman, R. W. Gibberd, B. T. Harrison, L. Newby, and J. D. Hamilton, "The Quality in Australian Health Care Study," *Med. J. Aust.*, vol. 163, no. 9, pp. 458–471, Nov. 1995.
- [42] E. J. Thomas, D. M. Studdert, H. R. Burstin, E. J. Orav, T. Zeena, E. J. Williams, K. M. Howard, P. C. Weiler, and T. A. Brennan, "Incidence and types of adverse events and negligent care in Utah and Colorado," *Med. Care*, vol. 38, no. 3, pp. 261–271, Mar. 2000.
- [43] T. Schiøler, H. Lipczak, B. L. Pedersen, T. S. Mogensen, K. B. Bech, A. Stockmarr, A. R. Svenning, and A. Frølich, "[Incidence of adverse events in hospitals. A retrospective study of medical records]," *Ugeskr. Laeger*, vol. 163, no. 39, pp. 5370–5378, Sep. 2001.
- [44] C. Vincent, "Adverse events in British hospitals: preliminary retrospective record review," *BMJ*, vol. 322, no. 7285, pp. 517–519, Mar. 2001.
- [45] M. Unbeck, O. Muren, and U. Lillkrona, "Identification of adverse events at an orthopedics department in Sweden," *Acta Orthop.*, vol. 79, no. 3, pp. 396–403, 2008.

- [46] M. Unbeck, K. Schildmeijer, P. Henriksson, U. Jürgensen, O. Muren, L. Nilsson, and K. Pukk Härenstam, “Is detection of adverse events affected by record review methodology? an evaluation of the ‘Harvard Medical Practice Study’ method and the ‘Global Trigger Tool,’” *Patient Saf. Surg.*, vol. 7, no. 1, p. 10, 2013.
- [47] J. Charnley, “The long-term results of low-friction arthroplasty of the hip performed as a primary intervention,” *J. Bone Joint Surg. Br.*, vol. 54-B, no. 1, pp. 61–76, Feb. 1972.
- [48] S. E. Gabriel and S.-L. T. Normand, “Getting the Methods Right — The Foundation of Patient-Centered Outcomes Research,” *N. Engl. J. Med.*, vol. 367, no. 9, pp. 787–790, 2012.
- [49] N. J. Devlin and J. Appleby, *Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making*. The King’s Fund 2010, 2010.
- [50] E. A. Lingard, K. Muthumayandi, and J. P. Holland, “Comparison of patient-reported outcomes between hip resurfacing and total hip replacement,” *J Bone Jt. Surg Br*, vol. 91-B, no. 12, pp. 1550–1554, Dec. 2009.
- [51] S. Berger, “For Young and Active, a Hip Replacement Alternative,” *The New York Times*, 17-Oct-2006.
- [52] P. N. Maffulli, “Young at hip,” *The Guardian*, 09-Feb-2006.
- [53] A. Garratt, L. Schmidt, A. Mackintosh, and R. Fitzpatrick, “Quality of life measurement: bibliographic study of patient assessed health outcome measures,” *BMJ*, vol. 324, no. 7351, p. 1417, Jun. 2002.
- [54] EuroQol Group, “EuroQol—a new facility for the measurement of health-related quality of life,” *Health Policy Amst. Neth.*, vol. 16, no. 3, pp. 199–208, Dec. 1990.
- [55] A. L. Stewart and J. E. Ware, *Measuring Functioning and Well-being: The Medical Outcomes Study Approach*. Duke University Press, 1992.
- [56] J. Ware Jr, M. Kosinski, and S. D. Keller, “A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity,” *Med. Care*, vol. 34, no. 3, pp. 220–233, Mar. 1996.
- [57] J. E. Ware, *How to Score and Interpret Single-item Health Status Measures: A Manual for Users of the of the SF-8 Health Survey:(with a Supplement on the SF-6 Health Survey)*. QualityMetric, Incorporated, 2001.
- [58] E. Ashby, M. P. W. Grocott, and F. S. Haddad, “Outcome measures for orthopaedic interventions on the hip,” *J. Bone Joint Surg. Br.*, vol. 90-B, no. 5, pp. 545–549, May 2008.
- [59] W. H. Harris, “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.*, vol. 51, no. 4, pp. 737–755, Jun. 1969.
- [60] P. Soderman and H. Malchau, “Is the Harris Hip Score System Useful to Study the Outcome of Total Hip Replacement?,” *Clin. Orthop.*, vol. 384, pp. 189–197, Mar. 2001.
- [61] H. L. Hoeksma, C. H. M. V. den Ende, H. K. Ronday, A. Heering, F. C. Breedveld, and J. Dekker, “Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in osteoarthritis of the hip,” *Ann. Rheum. Dis.*, vol. 62, no. 10, pp. 935–938, Oct. 2003.
- [62] N. N. Mahomed, D. C. Arndt, B. J. McGrory, and W. H. Harris, “The Harris hip score: Comparison of patient self-report with surgeon assessment,” *J. Arthroplasty*, vol. 16, no. 5, pp. 575–580, Aug. 2001.
- [63] D. N. Bellamy and W. W. Buchanan, “A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee,” *Clin. Rheumatol.*, vol. 5, no. 2, pp. 231–241, Jun. 1986.
- [64] N. Bellamy, W. W. Buchanan, C. H. Goldsmith, J. Campbell, and L. W. Stitt, “Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee,” *J. Rheumatol.*, vol. 15, no. 12, pp. 1833–1840, Dec. 1988.
- [65] N. Bellamy, “The WOMAC Knee and Hip Osteoarthritis Indices: development, validation, globalization and influence on the development of the AUSCAN Hand Osteoarthritis Indices,” *Clin. Exp. Rheumatol.*, vol. 23, no. 5 Suppl 39, pp. S148–153, Oct. 2005.
- [66] J. Dawson, R. Fitzpatrick, A. Carr, and D. Murray, “Questionnaire on the perceptions of patients about total hip replacement,” *J. Bone Joint Surg. Br.*, vol. 78, no. 2, pp. 185–190, Mar. 1996.
- [67] D. W. Murray, R. Fitzpatrick, K. Rogers, H. Pandit, D. J. Beard, A. J. Carr, and J. Dawson, “The use of the Oxford hip and knee scores,” *J. Bone Joint Surg. Br.*, vol. 89-B, no. 8, pp. 1010–1014, Aug. 2007.
- [68] A. Nilsson and A. Bremander, “Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire,” *Arthritis Care Res.*, vol. 63, no. S11, pp. S200–S207, Nov. 2011.
- [69] H. Behrend, K. Giesinger, J. M. Giesinger, and M. S. Kuster, “The ‘Forgotten Joint’ as the Ultimate Goal in Joint Arthroplasty: Validation of a New Patient-Reported Outcome Measure,” *J. Arthroplasty*, vol. 27, no. 3, pp. 430–436.e1, Mar. 2012.

- [70] J. M. Giesinger, M. S. Kuster, B. Holzner, and K. Giesinger, "Development Of A Computer-Adaptive Version Of The Forgotten Joint Score," *J. Arthroplasty*, vol. 28, no. 3, pp. 418–422, Mar. 2013.
- [71] K. J. Rothman, S. Greenland, and T. L. Lash, *Modern epidemiology*. Philadelphia: Wolters Kluwer / Lippincott Williams & Wilkins, 2008.
- [72] S. S. Coughlin, "Recall bias in epidemiologic studies," *J. Clin. Epidemiol.*, vol. 43, no. 1, pp. 87–91, 1990.
- [73] K. Korkeila, S. Suominen, J. Ahvenainen, A. Ojanlatva, P. Rautava, H. Helenius, and M. Koskenvuo, "Non-response and related factors in a nation-wide health survey," *Eur. J. Epidemiol.*, vol. 17, no. 11, pp. 991–999, Nov. 2001.
- [74] J. Kim, J. H. Lonner, C. L. Nelson, and P. A. Lotke, "Response bias: effect on outcomes evaluation by mail surveys after total knee arthroplasty," *J. Bone Joint Surg. Am.*, vol. 86-A, no. 1, pp. 15–21, Jan. 2004.
- [75] A. G. J. Bot, J. A. Anderson, V. Neuhaus, and D. Ring, "Factors Associated With Survey Response in Hand Surgery Research," *Clin. Orthop. Relat. Res.*, vol. 471, no. 10, pp. 3237–3242, Oct. 2013.
- [76] P. Dieppe, A. Judge, S. Williams, I. Ikwueke, K.-P. Guenther, M. Floeren, J. Huber, T. Ingvarsson, I. Learmonth, L. S. Lohmander, A. Nilsson, W. Puhl, D. Rowley, R. Theiler, K. Dreinhofer, and E. Study Group, "Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in European orthopaedic centres," *BMC Musculoskelet. Disord.*, vol. 10, no. 1, p. 19, 2009.
- [77] G. K. Chammout, S. S. Mukka, T. Carlsson, G. F. Neander, A. W. Helge Stark, and O. G. Sköldenberg, "Total Hip Replacement Versus Open Reduction and Internal Fixation of Displaced Femoral Neck FracturesA Randomized Long-Term Follow-up Study," *J. Bone Jt. Surg.*, vol. 94, no. 21, pp. 1921–1928, Nov. 2012.
- [78] O. Leonardsson, *Arthroplasty for Femoral Neck Fracture. Results of a nationwide implementation*, vol. 2012:45. Dept. of Orthopaedics, Inst. for Clinical Sciences, Malmö, 2012.
- [79] E. Jämsen, L. J. Virta, M. Hakala, M. J. Kauppi, A. Malmivaara, and M. U. K. Lehto, "The decline in joint replacement surgery in rheumatoid arthritis is associated with a concomitant increase in the intensity of anti-rheumatic therapy: A nationwide register-based study from 1995 through 2010," *Acta Orthop.*, vol. 84, no. 4, pp. 331–337, Aug. 2013.
- [80] F. E. Harrell, *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer, 2001.

- [81] H. Malchau, P. Herberts, and L. Ahnfelt, "Prognosis of total hip replacement in Sweden. Follow-up of 92,675 operations performed 1978-1990," *Acta Orthop. Scand.*, vol. 64, no. 5, pp. 497-506, Oct. 1993.
- [82] P. Münger, C. Röder, U. Ackermann-Liebrich, and A. Busato, "Patient-related risk factors leading to aseptic stem loosening in total hip arthroplasty: a case-control study of 5,035 patients," *Acta Orthop.*, vol. 77, no. 4, pp. 567-574, Aug. 2006.
- [83] S. Overgaard, "Primær hoftealloplastik: Overlevelseskurver Effekt af alder og køn, alle operationstyper," in *Årsrapport 2011*, Dansk Hoftealloplastik Register – Årsrapport 2011 Dansk Hoftealloplastik Register, 2011, pp. 52-53.
- [84] C. Wainwright, J.-C. Theis, N. Garneti, and M. Melloh, "Age at hip or knee joint replacement surgery predicts likelihood of revision surgery," *J Bone Jt. Surg Br*, vol. 93-B, no. 10, pp. 1411-1415, Oktober 2011.
- [85] S. Graves, J. Lynch, L. Griffith, S. Miller, M. Lorimer, Y.-L. Liu, K. Cashman, T. Stanford, G. O'Donahue, D. Davidsson, R. de Steiger, A. Tomkins, R. Vial, and L. Miller, "Outcome by Patient Characteristics," in *Australian Orthopaedic Association National Joint Replacement Registry.*, AOA, 2013, pp. 52-62.
- [86] J. J. Prokopetz, E. Losina, R. L. Bliss, J. Wright, J. A. Baron, and J. N. Katz, "Risk factors for revision of primary total hip arthroplasty: a systematic review," *BMC Musculoskelet. Disord.*, vol. 13, no. 1, p. 251, Dec. 2012.
- [87] A. L. Malkani, K. L. Ong, E. Lau, S. M. Kurtz, B. J. Justice, and M. T. Manley, "Early- and late-term dislocation risk after primary hip arthroplasty in the Medicare population," *J. Arthroplasty*, vol. 25, no. 6 Suppl, pp. 21-25, Sep. 2010.
- [88] N. F. SooHoo, E. Farng, J. R. Lieberman, L. Chambers, and D. S. Zingmond, "Factors That Predict Short-term Complication Rates After Total Hip Arthroplasty," *Clin. Orthop. Relat. Res.*, vol. 468, no. 9, pp. 2363-2371, Apr. 2010.
- [89] O. Ethgen, O. Bruyère, F. Richy, C. Dardennes, and J.-Y. Reginster, "Health-Related Quality of Life in Total Hip and Total Knee Arthroplasty A Qualitative and Systematic Review of the Literature," *J. Bone Jt. Surg.*, vol. 86, no. 5, pp. 963-974, May 2004.
- [90] F. X. McGuigan, W. J. Hozack, L. Moriarty, K. Eng, and R. H. Rothman, "Predicting quality-of-life outcomes following total joint arthroplasty: Limitations of the SF-36 health status questionnaire," *J. Arthroplasty*, vol. 10, no. 6, pp. 742-747, Dec. 1995.

- [91] P. D. Birdsall, J. H. Hayes, R. Cleary, I. M. Pinder, C. G. Moran, and J. L. Sher, "Health outcome after total knee replacement in the very elderly," *J. Bone Joint Surg. Br.*, vol. 81-B, no. 4, pp. 660–662, Jul. 1999.
- [92] P. L. Santaguida, G. A. Hawker, P. L. Hudak, R. Glazier, N. N. Mahomed, H. J. Kreder, P. C. Coyte, and J. G. Wright, "Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review," *Can. J. Surg. J. Can. Chir.*, vol. 51, no. 6, pp. 428–436, Dec. 2008.
- [93] O. Rolfson, J. Kärrholm, L. E. Dahlberg, and G. Garellick, "Patient-reported outcomes in the Swedish Hip Arthroplasty Register: Results of a nationwide prospective observational study," *J. Bone Joint Surg. Br.*, vol. 93-B, no. 7, pp. 867–875, Jul. 2011.
- [94] J. C. Keurentjes, M. Fiocco, and R. G. Nelissen, "Willingness to undergo surgery again validated clinically important differences in health-related quality of life after total hip replacement or total knee replacement surgery," *J. Clin. Epidemiol.*, vol. 67, no. 1, pp. 114–120, Jan. 2014.
- [95] P. Rissanen, S. Aro, H. Sintonen, P. Slätis, and P. Paavolainen, "Quality of life and functional ability in hip and knee replacements: A prospective study," *Qual. Life Res.*, vol. 5, no. 1, pp. 56–64, Feb. 1996.
- [96] R. Benroth and S. Gawande, "Patient-reported health status in total joint replacement," *J. Arthroplasty*, vol. 14, no. 5, pp. 576–580, Aug. 1999.
- [97] C. A. Jones, D. C. Voaklander, D. W. Johnston, and M. E. Suarez-Almazor, "The effect of age on pain, function, and quality of life after total hip and knee arthroplasty," *Arch. Intern. Med.*, vol. 161, no. 3, pp. 454–460, Feb. 2001.
- [98] J. M. Quintana, A. Escobar, U. Aguirre, I. Lafuente, and J. C. Arenaza, "Predictors of health-related quality-of-life change after total hip arthroplasty," *Clin. Orthop.*, vol. 467, no. 11, pp. 2886–2894, Nov. 2009.
- [99] B. M. Lawless, M. Greene, J. Slover, Y.-M. Kwon, and H. Malchau, "Does Age or Bilateral Disease Influence the Value of Hip Arthroplasty?," *Clin. Orthop. Relat. Res.*, vol. 470, no. 4, pp. 1073–1078, Apr. 2012.
- [100] F. Krupic, T. Eisler, T. Eliasson, G. Garellick, M. Gordon, and J. Kärrholm, "No influence of immigrant background on the outcome of total hip arthroplasty," *Acta Orthop.*, pp. 1–7, Jan. 2013.
- [101] A. Rothwell, J. Taylor, P. Devane, M. Wright, H. Leslie, H. Griffin, P. Larmer, T. Hobbs, and Chris Frampton, "Fourteen year report: January 1999 to December 2012," The New Zealand Joint Registry, Annual report 14, Nov. 2013.
- [102] G. Hallan, S. A. Lie, O. Furnes, L. B. Engesaeter, S. E. Vollset, and L. I. Havelin, "Medium- and long-term performance of 11 516 uncemented

- primary femoral stems from the Norwegian arthroplasty register,” *J Bone Jt. Surg Br*, vol. 89-B, no. 12, pp. 1574–1580, Dec. 2007.
- [103] Inacio MS, Ake CF, Paxton EW, and et al, “Sex and risk of hip implant failure: Assessing total hip arthroplasty outcomes in the united states,” *JAMA Intern. Med.*, vol. 173, no. 6, pp. 435–441, Mar. 2013.
- [104] A. J. Smith, P. Dieppe, K. Vernon, M. Porter, and A. W. Blom, “Failure rates of stemmed metal-on-metal hip replacements: analysis of data from the National Joint Registry of England and Wales,” *The Lancet*, vol. 379, no. 9822, pp. 1199–1204, Mar. 2012.
- [105] K. L. Ong, S. M. Kurtz, E. Lau, K. J. Bozic, D. J. Berry, and J. Parvizi, “Prosthetic Joint Infection Risk After Total Hip Arthroplasty in the Medicare Population,” *J. Arthroplasty*, vol. 24, no. 6, Supplement 1, pp. 105–109, Sep. 2009.
- [106] H. Dale, G. Hallan, G. Hallan, B. Espehaug, L. I. Havelin, and L. B. Engesaeter, “Increasing risk of revision due to deep infection after hip arthroplasty,” *Acta Orthop.*, vol. 80, no. 6, pp. 639–645, Dec. 2009.
- [107] A. B. Pedersen, J. E. Svendsson, S. P. Johnsen, A. Riis, and S. Overgaard, “Risk factors for revision due to infection after primary total hip arthroplasty,” *Acta Orthop.*, vol. 81, no. 5, pp. 542–547, Oct. 2010.
- [108] R. S. Namba, M. C. S. Inacio, and E. W. Paxton, “Risk factors associated with surgical site infection in 30 491 primary total hip replacements,” *J. Bone Jt. Surg. - Br. Vol.*, vol. 94-B, no. 10, pp. 1330–1338, Sep. 2012.
- [109] M. Khatod, G. Cafri, R. S. Namba, M. C. S. Inacio, and E. W. Paxton, “Risk factors for Total Hip Arthroplasty Aseptic Revision,” *J. Arthroplasty*, Jan. 2014.
- [110] J. L. Conroy, S. L. Whitehouse, S. E. Graves, N. L. Pratt, P. Ryan, and R. W. Crawford, “Risk Factors for Revision for Early Dislocation in Total Hip Arthroplasty,” *J. Arthroplasty*, vol. 23, no. 6, pp. 867–872, Sep. 2008.
- [111] D. J. W. McMinn, K. I. E. Snell, J. Daniel, R. B. C. Treacy, P. B. Pynsent, and R. D. Riley, “Mortality and implant revision rates of hip arthroplasty in patients with osteoarthritis: registry based cohort study,” *BMJ*, vol. 344, no. jun14 1, pp. e3319–e3319, Jun. 2012.
- [112] J. D. Slover, M. G. Walsh, and J. D. Zuckerman, “Sex and Race Characteristics in Patients Undergoing Hip and Knee Arthroplasty in an Urban Setting,” *J. Arthroplasty*, vol. 25, no. 4, pp. 576–580, Jun. 2010.
- [113] H.-Y. Shi, L.-W. Mau, J.-K. Chang, and J.-W. Wang, “Responsiveness of the Harris Hip Score and the SF-36: five years after total hip arthroplasty,” *Qual. Life Res.*, vol. 18, no. 8, pp. 1053–1060, Oct. 2009.
- [114] A. Laupacis, R. Bourne, C. Rorabeck, D. Feeny, C. Wong, P. Tugwell, K. Leslie, and R. Bullas, “The effect of elective total hip replacement on

- health-related quality of life,” *J. Bone Jt. Surg.*, vol. 75, no. 11, pp. 1619–1626, Nov. 1993.
- [115] G. H. Smith, S. Johnson, J. A. Ballantyne, E. Dunstan, and I. J. Brenkel, “Predictors of excellent early outcome after total hip arthroplasty,” *J. Orthop. Surg.*, vol. 7, p. 13, 2012.
- [116] S. Hajat, R. Fitzpatrick, R. Morris, B. Reeves, M. Rigge, O. Williams, D. Murray, and P. Gregg, “Does waiting for total hip replacement matter? Prospective cohort study,” *J. Health Serv. Res. Policy*, vol. 7, no. 1, pp. 19–25, Jan. 2002.
- [117] A. Judge, M. K. Javaid, N. K. Arden, J. Cushnaghan, I. Reading, P. Croft, P. A. Dieppe, and C. Cooper, “Clinical tool to identify patients who are most likely to achieve long-term improvement in physical function after total hip arthroplasty,” *Arthritis Care Res.*, vol. 64, no. 6, pp. 881–889, Jun. 2012.
- [118] C. Baumann, A. Rat, G. Osnowycz, D. Mainard, C. Cuny, and F. Guillemin, “Satisfaction with care after total hip or knee replacement predicts self-perceived health status after surgery,” *BMC Musculoskelet. Disord.*, vol. 10, no. 1, p. 150, 2009.
- [119] M. Stevens, N. Paans, R. Wagenmakers, J. van Beveren, J. J. A. M. van Raay, K. van der Meer, R. Stewart, S. K. Bulstra, I. H. F. Reininga, and I. van den Akker-Scheek, “The Influence of Overweight/Obesity on Patient-Perceived Physical Functioning and Health-Related Quality of Life After Primary Total Hip Arthroplasty,” *Obes. Surg.*, vol. 22, no. 4, pp. 523–529, Apr. 2012.
- [120] C. J. Lavernia, J. C. Alcerro, J. S. Contreras, and M. D. Rossi, “Patient perceived outcomes after primary hip arthroplasty: does gender matter?,” *Clin. Orthop.*, vol. 469, no. 2, pp. 348–354, Feb. 2011.
- [121] K. Bjorgul, W. M. Novicoff, and K. J. Saleh, “Evaluating comorbidities in total hip and knee arthroplasty: available instruments,” *J. Orthop. Traumatol.*, vol. 11, pp. 203–209, Nov. 2010.
- [122] E. W. Paxton, O. Furnes, R. S. Namba, M. C. S. Inacio, A. M. Fenstad, and L. I. Havelin, “Comparison of the Norwegian Knee Arthroplasty Register and a United States Arthroplasty Registry,” *J. Bone Jt. Surg. Am.*, vol. 93, no. Supplement_3, pp. 20–30, Dec. 2011.
- [123] B. M. Jolles, P. Zangger, and P.-F. Leyvraz Prof., “Factors predisposing to dislocation after primary total hip arthroplasty,” *J. Arthroplasty*, vol. 17, no. 3, pp. 282–288, Apr. 2002.
- [124] M. A. Rauh and K. A. Krackow, “In-hospital deaths following elective total joint arthroplasty,” *Orthopedics*, vol. 27, no. 4, pp. 407–411, Apr. 2004.

- [125] S. Ridgeway, "Infection of the surgical site after arthroplasty of the hip," *J. Bone Jt. Surg. - Br. Vol.*, vol. 87-B, no. 6, pp. 844–850, Jun. 2005.
- [126] M. T. A. Sharabiani, P. Aylin, and A. Bottle, "Systematic review of comorbidity indices for administrative data," *Med. Care*, vol. 50, no. 12, pp. 1109–1118, Dec. 2012.
- [127] M. E. Charlson, P. Pompei, K. L. Ales, and C. R. MacKenzie, "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation," *J. Chronic Dis.*, vol. 40, no. 5, pp. 373–383, 1987.
- [128] A. Elixhauser, C. Steiner, D. R. Harris, and R. M. Coffey, "Comorbidity measures for use with administrative data," *Med. Care*, vol. 36, no. 1, pp. 8–27, Jan. 1998.
- [129] J. N. Armitage and J. H. van der Meulen, "Identifying co-morbidity in surgical patients using administrative data with the Royal College of Surgeons Charlson Score," *Br. J. Surg.*, vol. 97, no. 5, pp. 772–781, May 2010.
- [130] S. P. Johnsen, H. T. Sørensen, U. Lucht, K. Søballe, S. Overgaard, and A. B. Pedersen, "Patient-related predictors of implant failure after primary total hip replacement in the initial, short- and long-terms. A nationwide Danish follow-up study including 36,984 patients," *J. Bone Joint Surg. Br.*, vol. 88, no. 10, pp. 1303–1308, Oct. 2006.
- [131] J. M. Paterson, J. I. Williams, H. J. Kreder, N. N. Mahomed, N. Gunraj, X. Wang, and A. Laupacis, "Provider volumes and early outcomes of primary total joint replacement in Ontario," *Can. J. Surg.*, vol. 53, no. 3, p. 175, Jun. 2010.
- [132] D. M. Urquhart, F. S. Hanna, S. L. Brennan, A. E. Wluka, K. Leder, P. A. Cameron, S. E. Graves, and F. M. Cicuttini, "Incidence and Risk Factors for Deep Surgical Site Infection After Primary Total Hip Arthroplasty: A Systematic Review," *J. Arthroplasty*, vol. 25, no. 8, pp. 1216–1222.e3, Dec. 2010.
- [133] D. L. Groll, T. To, C. Bombardier, and J. G. Wright, "The development of a comorbidity index with physical function as the outcome," *J. Clin. Epidemiol.*, vol. 58, no. 6, pp. 595–602, Jun. 2005.
- [134] A.-C. Rat, F. Guillemin, G. Osnowycz, J.-P. Delagoutte, C. Cuny, D. Mainard, and C. Baumann, "Total hip or knee replacement for osteoarthritis: Mid- and long-term quality of life," *Arthritis Care Res.*, vol. 62, no. 1, pp. 54–62, Jan. 2010.
- [135] M. Mariconda, O. Galasso, G. G. Costa, P. Recano, and S. Cerbasi, "Quality of life and functionality after total hip arthroplasty: a long-term follow-up study," *BMC Musculoskelet. Disord.*, vol. 12, no. 1, p. 222, 2011.

- [136] O. Rolfson, L. E. Dahlberg, J.-A. Nilsson, H. Malchau, and G. Garellick, "Variables determining outcome in total hip replacement surgery," *J. Bone Joint Surg. Br.*, vol. 91-B, no. 2, pp. 157–161, Feb. 2009.
- [137] G. Hulleberg, A. Aamodt, B. Espehaug, and P. Benum, "A clinical and radiographic 13-year follow-up study of 138 Charnley hip arthroplasties in patients 50–70 years old: Comparison of university hospital data and registry data," *Acta Orthop.*, vol. 79, no. 5, pp. 609–617, Jan. 2008.
- [138] M. Dunbar, O. Robertsson, and L. Ryd, "What's all that noise? The effect of co-morbidity on health outcome questionnaire results after knee arthroplasty," *Acta Orthop.*, vol. 75, no. 2, pp. 119–126, Jan. 2004.
- [139] M. Ostendorf, H. F. van Stel, E. Buskens, A. J. P. Schrijvers, L. N. Marting, A. J. Verbout, and W. J. A. Dhert, "Patient-reported outcome in total hip replacement. A comparison of five instruments of health status," *J. Bone Joint Surg. Br.*, vol. 86, no. 6, pp. 801–808, Aug. 2004.
- [140] J. J. Callaghan, S. H. Dysart, C. F. Savory, and W. J. Hopkinson, "Assessing the results of hip replacement. A comparison of five different rating systems," *J. Bone Joint Surg. Br.*, vol. 72-B, no. 6, pp. 1008–1009, Nov. 1990.
- [141] G. Garellick, H. Malchau, and P. Herberts, "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.*, vol. 80-B, no. 4, pp. 600–606, Jul. 1998.
- [142] C. J. Lavernia, A. Laoruengthana, J. S. Contreras, and M. D. Rossi, "All-Patient Refined Diagnosis-Related Groups in Primary Arthroplasty," *J. Arthroplasty*, vol. 24, no. 6, Supplement 1, pp. 19–23, Sep. 2009.
- [143] C. Röder, L. p. Staub, P. Eichler, M. Widmer, D. Dietrich, S. Eggli, and U. Müller, "Avoiding misclassification bias with the traditional Charnley classification: Rationale for a fourth Charnley class BB," *J. Orthop. Res.*, vol. 24, no. 9, pp. 1803–1808, 2006.
- [144] M. M. Vissers, J. B. Bussmann, J. A. N. Verhaar, J. J. V. Busschbach, S. M. A. Bierma-Zeinstra, and M. Reijman, "Psychological Factors Affecting the Outcome of Total Hip and Knee Arthroplasty: A Systematic Review," *Semin. Arthritis Rheum.*, vol. 41, no. 4, pp. 576–588, Feb. 2012.
- [145] O. Stundner, M. Kirksey, Y. L. Chiu, M. Mazumdar, L. Poultsides, P. Gerner, and S. G. Memtsoudis, "Demographics and Perioperative Outcome in Patients with Depression and Anxiety Undergoing Total Joint Arthroplasty: A Population-Based Study," *Psychosomatics*, vol. 54, no. 2, pp. 149–157, Mar. 2013.
- [146] K. J. Bozic, E. Lau, S. Kurtz, K. Ong, H. Rubash, T. P. Vail, and D. J. Berry, "Patient-Related Risk Factors for Periprosthetic Joint Infection and

- Postoperative Mortality Following Total Hip Arthroplasty in Medicare Patients,” *J. Bone Jt. Surg. Am.*, vol. 94, no. 9, pp. 794–800, May 2012.
- [147] R. A. Malinzak, M. A. Ritter, M. E. Berend, J. B. Meding, E. M. Olberding, and K. E. Davis, “Morbidly Obese, Diabetic, Younger, and Unilateral Joint Arthroplasty Patients Have Elevated Total Joint Arthroplasty Infection Rates,” *J. Arthroplasty*, vol. 24, no. 6, Supplement, pp. 84–88, Sep. 2009.
- [148] E. Jämsen, P. Nevalainen, A. Eskelinen, K. Huotari, J. Kalliovalkama, and T. Moilanen, “Obesity, Diabetes, and Preoperative Hyperglycemia as Predictors of Periprosthetic Joint Infection A Single-Center Analysis of 7181 Primary Hip and Knee Replacements for Osteoarthritis,” *J. Bone Jt. Surg.*, vol. 94, no. 14, p. e101 1–9, Jul. 2012.
- [149] J. S. Everhart, E. Altneu, and J. H. Calhoun, “Medical Comorbidities Are Independent Preoperative Risk Factors for Surgical Infection After Total Joint Arthroplasty,” *Clin. Orthop. Relat. Res.*, vol. 471, no. 10, pp. 3112–3119, Oct. 2013.
- [150] A. Lübbecke, R. Stern, G. Garavaglia, L. Zurcher, and P. Hoffmeyer, “Differences in outcomes of obese women and men undergoing primary total hip arthroplasty,” *Arthritis Care Res.*, vol. 57, no. 2, pp. 327–334, Mar. 2007.
- [151] D. Haverkamp, M. N. Klinkenbijn, M. P. Somford, G. H. R. Albers, and H. M. van der Vis, “Obesity in total hip arthroplasty—does it really matter?: A meta-analysis,” *Acta Orthop.*, vol. 82, no. 4, pp. 417–422, Aug. 2011.
- [152] A. M. Davis, A. M. Wood, A. C. M. Keenan, I. J. Brenkel, and J. A. Ballantyne, “Does body mass index affect clinical outcome post-operatively and at five years after primary unilateral total hip replacement performed for osteoarthritis? A multivariate analysis of prospective data,” *J. Bone Joint Surg. Br.*, vol. 93-B, no. 9, pp. 1178–1182, Sep. 2011.
- [153] P. Aranda Villalobos, J. L. Navarro-Espigares, E. Hernández-Torres, J. L. Martínez-Montes, M. Villalobos, and M. Arroyo-Morales, “Body Mass Index as Predictor of Health-Related Quality-of-Life Changes After Total Hip Arthroplasty: A Cross-Over Study,” *J. Arthroplasty*, vol. 28, no. 4, pp. 666–670, Apr. 2013.
- [154] M. M. Dowsey, D. Liew, J. D. Stoney, and P. F. M. Choong, “The impact of obesity on weight change and outcomes at 12 months in patients undergoing total hip arthroplasty,” *Med. J. Aust.*, vol. 193, no. 1, 2010.
- [155] C. A. Jones, V. Cox, G. S. Jhangri, and M. E. Suarez-Almazor, “Delineating the impact of obesity and its relationship on recovery after total joint arthroplasties,” *Osteoarthritis Cartilage*, vol. 20, no. 6, pp. 511–518, Jun. 2012.

- [156] J. G. Andrew, J. Palan, H. V. Kurup, P. Gibson, D. W. Murray, and D. J. Beard, "Obesity in total hip replacement," *J Bone Jt. Surg Br*, vol. 90-B, no. 4, pp. 424–429, Apr. 2008.
- [157] B. R. Shah and J. E. Hux, "Quantifying the Risk of Infectious Diseases for People With Diabetes," *Diabetes Care*, vol. 26, no. 2, pp. 510–513, Feb. 2003.
- [158] Ata A, Lee J, Bestle SL, Desemone J, and Stain SC, "Postoperative hyperglycemia and surgical site infection in general surgery patients," *Arch. Surg.*, vol. 145, no. 9, pp. 858–864, Sep. 2010.
- [159] P. K. H. Chan, I. J. Brenkel, and J. Aderinto, "The outcome of total hip and knee arthroplasty in diabetics," *Curr. Orthop.*, vol. 19, no. 1, pp. 59–67, Feb. 2005.
- [160] A. B. Pedersen, F. Mehnert, S. P. Johnsen, and H. T. Sorensen, "Risk of revision of a total hip replacement in patients with diabetes mellitus: A population-based follow up study," *J Bone Jt. Surg Br*, vol. 92-B, no. 7, pp. 929–934, Jul. 2010.
- [161] M. H. J. Marchant, N. A. Viens, C. Cook, T. P. Vail, and M. P. Bolognesi, "The Impact of Glycemic Control and Diabetes Mellitus on Perioperative Outcomes After Total Joint Arthroplasty," *J. Bone Jt. Surg.*, vol. 91, no. 7, pp. 1621–1629, Jul. 2009.
- [162] J. M. Queally, A. Abdulkarim, and K. J. Mulhall, "Total hip replacement in patients with neurological conditions," *J Bone Jt. Surg Br*, vol. 91-B, no. 10, pp. 1267–1273, Oktober 2009.
- [163] R. M. D. Meek, D. B. Allan, G. McPhillips, L. Kerr, and C. R. Howie, "Epidemiology of dislocation after total hip arthroplasty," *Clin. Orthop.*, vol. 447, pp. 9–18, Jun. 2006.
- [164] E. Jämsen, T. Puolakka, M. Peltola, A. Eskelinen, and M. U. K. Lehto, "Surgical outcomes of primary hip and knee replacements in patients with Parkinson's disease a nationwide registry-based case-controlled study," *Bone Jt. J.*, vol. 96-B, no. 4, pp. 486–491, Apr. 2014.
- [165] K. J. Bozic, E. Lau, K. Ong, V. Chan, S. Kurtz, T. P. Vail, H. E. Rubash, and D. J. Berry, "Risk factors for early revision after primary total hip arthroplasty in Medicare patients," *Clin. Orthop.*, vol. 472, no. 2, pp. 449–454, Feb. 2014.
- [166] J. A. Singh and D. G. Lewallen, "Medical Comorbidity is Associated with Persistent Index Hip Pain after Total Hip Arthroplasty," *Pain Med.*, vol. 14, no. 8, pp. 1222–1229, Aug. 2013.
- [167] Luppino FS, de Wit LM, Bouvy PF, and et al, "Overweight, obesity, and depression: A systematic review and meta-analysis of longitudinal studies," *Arch. Gen. Psychiatry*, vol. 67, no. 3, pp. 220–229, Mar. 2010.

- [168] S. L. Samson and A. J. Garber, "Metabolic Syndrome," *Endocrinol. Metab. Clin. North Am.*, vol. 43, no. 1, pp. 1–23, Mar. 2014.
- [169] H. Nåsell, *The impact of smoking on orthopaedic patients*. Karolinska Institutet, 2011.
- [170] M. E. Greene, O. Rolfson, S. Nemes, M. Gordon, H. Malchau, and G. Garellick, "Education Attainment is Associated With Patient-reported Outcomes: Findings From the Swedish Hip Arthroplasty Register," *Clin. Orthop.*, Feb. 2014.
- [171] B. Espehaug, L. I. Havelin, L. B. Engesaeter, N. Langeland, and S. E. Vollset, "Patient-related risk factors for early revision of total hip replacements. A population register-based case-control study of 674 revised hips," *Acta Orthop. Scand.*, vol. 68, no. 3, pp. 207–215, Jun. 1997.
- [172] National Board of Health and Welfare, "In English - the National Patient Register," *Socialstyrelsen*, 04-Mar-2014. [Online]. Available: <http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/ine-english>. [Accessed: 04-Mar-2014].
- [173] J. F. Ludvigsson, E. Andersson, A. Ekblom, M. Feychting, J.-L. Kim, C. Reuterwall, M. Heurgren, and P. O. Olausson, "External review and validation of the Swedish national inpatient register," *BMC Public Health*, vol. 11, p. 450, 2011.
- [174] National Board of Health and Welfare, "Läkemedelsregistret," 04-Mar-2014. [Online]. Available: <http://www.socialstyrelsen.se/register/halsodataregister/lakemedelsregistret>. [Accessed: 04-Mar-2014].
- [175] A. Paulsen, A. B. Pedersen, S. Overgaard, and E. M. Roos, "Feasibility of 4 patient-reported outcome measures in a registry setting," *Acta Orthop.*, vol. 83, no. 4, pp. 321–327, Aug. 2012.
- [176] K. Burström, S. Sun, U.-G. Gerdtham, M. Henriksson, M. Johannesson, L.-Å. Levin, and N. Zethraeus, "Swedish experience-based value sets for EQ-5D health states," *Qual. Life Res.*, pp. 1–12, Aug. 2013.
- [177] K. U. Wittrup-Jensen, J. Lauridsen, C. Gudex, and K. M. Pedersen, "Generation of a Danish TTO value set for EQ-5D health states," *Scand. J. Public Health*, vol. 37, no. 5, pp. 459–466, Jul. 2009.
- [178] H. Brenner and M. Blettner, "Controlling for continuous confounders in epidemiologic research," *Epidemiology*, vol. 8, no. 4, pp. 429–434, 1997.
- [179] D. R. Cox, "Regression models and life tables," *JR Stat Soc B*, vol. 34, no. 2, pp. 187–220, 1972.
- [180] J. Ranstam, J. Kärrholm, P. Pulkkinen, K. Mäkelä, B. Espehaug, A. B. Pedersen, F. Mehnert, and O. Furnes, "Statistical analysis of arthroplasty data: I. Introduction and background," *Acta Orthop.*, vol. 82, no. 3, pp. 253–257, Jun. 2011.

- [181] M. Pintilie, *Competing Risks: A Practical Perspective*, 1st ed. Wiley, 2006.
- [182] J. P. Fine and R. J. Gray, “A Proportional Hazards Model for the Subdistribution of a Competing Risk,” *J. Am. Stat. Assoc.*, vol. 94, no. 446, pp. 496–509, Jun. 1999.
- [183] J. Ranstam, J. Kärrholm, P. Pulkkinen, K. Mäkelä, B. Espehaug, A. B. Pedersen, F. Mehnert, and O. Furnes, “Statistical analysis of arthroplasty data: II. Guidelines,” *Acta Orthop.*, vol. 82, no. 3, pp. 258–267, Jun. 2011.
- [184] B. Gray, *cmprsk: Subdistribution Analysis of Competing Risks*. 2013.
- [185] D. Schoenfeld, “Partial residuals for the proportional hazards regression model,” *Biometrika*, vol. 69, no. 1, pp. 239–241, Apr. 1982.
- [186] P. M. Grambsch and T. M. Therneau, “Proportional hazards tests and diagnostics based on weighted residuals,” *Biometrika*, vol. 81, no. 3, pp. 515–526, Sep. 1994.
- [187] F. Cribari-Neto and M. da G. A. Lima, “Heteroskedasticity-consistent interval estimators,” *J. Stat. Comput. Simul.*, vol. 79, no. 6, pp. 787–803, 2009.
- [188] H. White, “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity,” *Econom. J. Econom. Soc.*, pp. 817–838, 1980.
- [189] J. G. MacKinnon and H. White, “Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties,” *J. Econom.*, vol. 29, no. 3, pp. 305–325, Sep. 1985.
- [190] J. S. Long and L. H. Ervin, “Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model,” *Am. Stat.*, vol. 54, no. 3, pp. 217–224, Aug. 2000.
- [191] R. Koenker and G. Bassett Jr, “Regression quantiles,” *Econom. J. Econom. Soc.*, pp. 33–50, 1978.
- [192] R. C. Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2013.
- [193] J. F. E. Harrell, *rms: Regression Modeling Strategies*. 2013.
- [194] H. Wickham, *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.
- [195] Y. Xie, *Dynamic Documents with R and Knitr*. New York, NY: CRC Press, 2013.
- [196] Python Software Foundation, “7.2. re — Regular expression operations — Python v2.7.6 documentation,” *Python v2.7.6 documentation*, 01-Apr-2014. [Online]. Available: <https://docs.python.org/2/library/re.html>. [Accessed: 01-Apr-2014].
- [197] H. Quan, B. Li, C. M. Couris, K. Fushimi, P. Graham, P. Hider, J.-M. Januel, and V. Sundararajan, “Updating and Validating the Charlson

- Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries,” *Am. J. Epidemiol.*, vol. 173, no. 6, pp. 676–682, Mar. 2011.
- [198] C. van Walraven, P. C. Austin, A. Jennings, H. Quan, and A. J. Forster, “A Modification of the Elixhauser Comorbidity Measures Into a Point System for Hospital Death Using Administrative Data,” *Med. Care*, vol. 47, no. 6, pp. 626–633, Jun. 2009.
- [199] V. Sundararajan, H. Quan, P. Halfon, K. Fushimi, J.-C. Luthi, B. Burnand, and W. A. Ghali, “Cross-National Comparative Performance of Three Versions of the ICD-10 Charlson Index,” *Med. Care*, vol. 45, pp. 1210–1215, Dec. 2007.
- [200] K. Lai, E. R. Bohm, C. Burnell, and D. R. Hedden, “Presence of Medical Comorbidities in Patients With Infected Primary Hip or Knee Arthroplasties,” *J. Arthroplasty*, vol. 22, no. 5, pp. 651–656, Aug. 2007.
- [201] D. Fender, J. H. P. van der Meulen, and P. J. Gregg, “Relationship between outcome and annual surgical experience for the charnley total hip replacement. Results from a regional hip register,” *J. Bone Joint Surg. Br.*, vol. 85, no. 2, pp. 187–190, Mar. 2003.
- [202] K. J. Bozic, J. Maselli, P. S. Pekow, P. K. Lindenauer, T. P. Vail, and A. D. Auerbach, “The Influence of Procedure Volumes and Standardization of Care on Quality and Efficiency in Total Joint Replacement Surgery,” *J. Bone Jt. Surg. Am.*, vol. 92, no. 16, pp. 2643–2652, Nov. 2010.
- [203] M. Manley, K. Ong, E. Lau, and S. M. Kurtz, “Effect of Volume on Total Hip Arthroplasty Revision Rates in the United States Medicare Population,” *J. Bone Jt. Surg. Am.*, vol. 90, no. 11, pp. 2446–2451, Nov. 2008.
- [204] H. J. Kreder, R. A. Deyo, T. Koepsell, M. F. Swiontkowski, and W. Kreuter, “Relationship between the volume of total hip replacements performed by providers and the rates of postoperative complications in the state of Washington,” *J. Bone Joint Surg. Am.*, vol. 79, no. 4, pp. 485–494, Apr. 1997.
- [205] M. Pennington, R. Grieve, J. S. Sekhon, P. Gregg, N. Black, and J. H. van der Meulen, “Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis,” *BMJ*, vol. 346, no. feb27 1, p. f1026, Feb. 2013.
- [206] A. Judge, N. K. Arden, R. N. Batra, G. Thomas, D. Beard, M. K. Javaid, C. Cooper, D. Murray, D. Murray, G. Andrew, P. Gibson, J. Nolan, A. Hamer, M. Fordyce, K. Tuson, D. J. Beard, A. Potter, A. McGovern, K. Reilly, C. Jenkins, A. Cooper, C. Darrah, L. Cawton, P. Inaparthi, and C. Pitchfork, “The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-

- replacement surgery: a prospective cohort study,” *BMJ Open*, vol. 3, no. 3, Jan. 2013.
- [207] J. Cushnaghan, D. Coggon, I. Reading, P. Croft, P. Byng, K. Cox, P. Dieppe, and C. Cooper, “Long-term outcome following total hip arthroplasty: A controlled longitudinal study,” *Arthritis Care Res.*, vol. 57, no. 8, pp. 1375–1380, 2007.
- [208] C. Y. Ng, J. A. Ballantyne, and I. J. Brenkel, “Quality of life and functional outcome after primary total hip replacement: A five-year follow-up,” *J. Bone Joint Surg. Br.*, vol. 89-B, no. 7, pp. 868–873, Jul. 2007.
- [209] K. Grosse Frie, J. van der Meulen, and N. Black, “Relationship between patients’ reports of complications and symptoms, disability and quality of life after surgery,” *Br. J. Surg.*, vol. 99, no. 8, pp. 1156–1163, 2012.
- [210] P. Royston, D. G. Altman, and W. Sauerbrei, “Dichotomizing continuous predictors in multiple regression: a bad idea,” *Stat. Med.*, vol. 25, no. 1, pp. 127–141, Jan. 2006.
- [211] J. S. Williams, “Assessing the suitability of fractional polynomial methods in health services research: a perspective on the categorization epidemic,” *J. Health Serv. Res. Policy*, vol. 16, no. 3, pp. 147–152, Jul. 2011.
- [212] D. G. Altman, B. Lausen, W. Sauerbrei, and M. Schumacher, “Dangers of using ‘optimal’ cutpoints in the evaluation of prognostic factors,” *J. Natl. Cancer Inst.*, vol. 86, no. 11, pp. 829–835, 1994.
- [213] R. Miller and D. Siegmund, “Maximally selected chi square statistics,” *Biometrics*, vol. 38, no. 4, pp. 1011–1016, 1982.
- [214] B. Lausen and M. Schumacher, “Evaluating the effect of optimized cutoff values in the assessment of prognostic factors,” *Comput. Stat. Data Anal.*, vol. 21, no. 3, pp. 307–326, Mar. 1996.
- [215] S. G. Hilsenbeck and G. M. Clark, “Practical p-value adjustment for optimally selected cutpoints,” *Stat. Med.*, vol. 15, no. 1, pp. 103–112, 1996.
- [216] S. W. Lagakos, “Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable,” *Stat. Med.*, vol. 7, no. 1–2, pp. 257–274, 1988.
- [217] L. Shan, B. Shan, D. Graham, and A. Saxena, “Total hip replacement: a systematic review and meta-analysis on mid-term quality of life,” *Osteoarthritis Cartilage*, vol. 22, no. 3, pp. 389–406, Mar. 2014.
- [218] A. Simm, N. Nass, B. Bartling, B. Hofmann, R.-E. Silber, and A. Navarrete Santos, “Potential biomarkers of ageing,” *Biol. Chem.*, vol. 389, no. 3, pp. 257–265, Mar. 2008.

- [219] C. Röder, J. Parvizi, S. Eggli, D. J. Berry, M. E. Müller, and A. Busato, “Demographic factors affecting long-term outcome of total hip arthroplasty,” *Clin. Orthop.*, no. 417, pp. 62–73, Dec. 2003.
- [220] H. K. Vincent, A. P. Alfano, L. Lee, and K. R. Vincent, “Sex and Age Effects on Outcomes of Total Hip Arthroplasty After Inpatient Rehabilitation,” *Arch. Phys. Med. Rehabil.*, vol. 87, no. 4, pp. 461–467, Apr. 2006.
- [221] W. A. Macrae, “Chronic post-surgical pain: 10 years on,” *Br. J. Anaesth.*, vol. 101, no. 1, pp. 77–86, Jul. 2008.
- [222] K. Burström, M. Johannesson, and F. Diderichsen, “Swedish population health-related quality of life results using the EQ-5D,” *Qual. Life Res.*, vol. 10, no. 7, pp. 621–635, Aug. 2001.
- [223] J. Sørensen, M. Davidsen, C. Gudex, K. M. Pedersen, and H. Brønnum-Hansen, “Danish EQ-5D population norms,” *Scand. J. Public Health*, vol. 37, no. 5, pp. 467–474, Jul. 2009.
- [224] A. J. Smith, V. Wylde, J. R. Berstock, A. D. Maclean, and A. W. Blom, “Surgical approach and patient-reported outcomes after total hip replacement,” *Hip Int. J. Clin. Exp. Res. Hip Pathol. Ther.*, vol. 22, no. 4, pp. 355–361, Aug. 2012.
- [225] J. V. Lindgren, P. Wretenberg, J. Kärrholm, G. Garellick, and O. Rolfson, “Patient reported outcome is influenced by surgical approach in total hip replacement a study of the Swedish Hip Arthroplasty Register including 42,233 patients,” *Bone Jt. J.*, vol. 96-B, p. ??–??, 2014.
- [226] J. A. Lieberman, J. Greenhouse, R. M. Hamer, K. R. Krishnan, C. B. Nemeroff, D. V. Sheehan, M. E. Thase, and M. B. Keller, “Comparing the Effects of Antidepressants: Consensus Guidelines for Evaluating Quantitative Reviews of Antidepressant Efficacy,” *Neuropsychopharmacology*, vol. 30, no. 3, pp. 445–460, Jan. 2005.
- [227] S. Henriksson and G. Isacson, “Increased antidepressant use and fewer suicides in Jämtland county, Sweden, after a primary care educational programme on the treatment of depression,” *Acta Psychiatr. Scand.*, vol. 114, no. 3, pp. 159–167, 2006.
- [228] S. Henriksson, R. Asplund, G. Boëthius, T. Hällström, and G. Isacson, “Infrequent use of antidepressants in depressed individuals (an interview and prescription database study in a defined Swedish population 2001–2002),” *Eur. Psychiatry*, vol. 21, no. 6, pp. 355–360, Sep. 2006.
- [229] T. Duivenvoorden, M. M. Vissers, J. A. N. Verhaar, J. J. V. Busschbach, T. Gosens, R. M. Bloem, S. M. A. Bierma-Zeinstra, and M. Reijman, “Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study,” *Osteoarthritis Cartilage*, vol. 21, no. 12, pp. 1834–1840, Dec. 2013.

- [230] S. McDonald, S. E. Hetrick, and S. Green, “Pre-operative education for hip or knee replacement,” in *Cochrane Database of Systematic Reviews*, John Wiley & Sons, Ltd, 2004.
- [231] M. Hossain, D. J. Parfitt, D. J. Beard, C. Darrah, J. Nolan, D. W. Murray, and G. Andrew, “Does pre-operative psychological distress affect patient satisfaction after primary total hip arthroplasty?,” *BMC Musculoskelet. Disord.*, vol. 12, no. 1, p. 122, 2011.
- [232] T. Rosemann, M. Backenstrass, K. Joest, A. Rosemann, J. Szecsenyi, and G. Laux, “Predictors of depression in a sample of 1,021 primary care patients with osteoarthritis,” *Arthritis Care Res.*, vol. 57, no. 3, pp. 415–422, 2007.
- [233] Lin EB, Katon W, Von Korff M, and et al, “Effect of improving depression care on pain and functional outcomes among older adults with arthritis: A randomized controlled trial,” *JAMA*, vol. 290, no. 18, pp. 2428–2429, Nov. 2003.
- [234] S. J. Linton and S. Bergbom, “Understanding the link between depression and pain,” *Scand. J. Pain*, vol. 2, no. 2, pp. 47–54, Apr. 2011.
- [235] A. Tylee and P. Gandhi, “The Importance of Somatic Symptoms in Depression in Primary Care,” *Prim. Care Companion J. Clin. Psychiatry*, vol. 7, no. 4, p. 167, 2005.
- [236] S. Henriksson, G. Boëthius, J. Håkansson, and G. Isacson, “Indications for and outcome of antidepressant medication in a general population: a prescription database and medical record study, in Jämtland county, Sweden, 1995,” *Acta Psychiatr. Scand.*, vol. 108, no. 6, pp. 427–431, 2003.
- [237] P. Söderman, H. Malchau, P. Herberts, and O. Johnell, “Are the findings in the Swedish National Total Hip Arthroplasty Register valid?: A comparison between the Swedish National Total Hip Arthroplasty Register, the National Discharge Register, and the National Death Register,” *J. Arthroplasty*, vol. 15, no. 7, pp. 884–889, Oct. 2000.
- [238] A. M. Jurek, S. Greenland, G. Maldonado, and T. R. Church, “Proper interpretation of non-differential misclassification effects: expectations vs observations,” *Int. J. Epidemiol.*, vol. 34, no. 3, pp. 680–687, Jun. 2005.
- [239] O. Rolfson, R. Salomonsson, L. Dahlberg, and G. Garellick, “Reliability and response rate of an internet-based follow-up questionnaire for measuring patient-reported outcome after total hip arthroplasty,” *J. Bone Joint Surg. Br.*, vol. 94-B, no. SUPP XXXVII, pp. 272–272, Sep. 2012.
- [240] J. F. Ludvigsson, P. Otterblad-Olausson, B. U. Pettersson, and A. Ekblom, “The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research,” *Eur. J. Epidemiol.*, vol. 24, no. 11, pp. 659–667, 2009.

- [241] M. Rosén and T. Hakulinen, “The Quality of Data in Registers,” in *Handbook of epidemiology*, W. Ahrens and I. Pigeot, Eds. Berlin: Springer, 2005, p. 237.
- [242] E. Whiskey, U. Werneke, and D. Taylor, “A systematic review and meta-analysis of *Hypericum perforatum* in depression: a comprehensive clinical review,” *Int. Clin. Psychopharmacol.*, vol. 16, no. 5, pp. 239–252, Sep. 2001.
- [243] K. Johnell, J. Fastbom, M. Rosén, and A. Leimanis, “Inappropriate Drug Use in the Elderly: A Nationwide Register-Based Study,” *Ann. Pharmacother.*, vol. 41, no. 7–8, pp. 1243–1248, Jul. 2007.
- [244] V. R. Basili and B. T. Perricone, “Software Errors and Complexity: An Empirical Investigation,” *Commun ACM*, vol. 27, no. 1, pp. 42–52, Jan. 1984.
- [245] G. J. Myers, C. Sandler, and T. Badgett, *The Art of Software Testing*. John Wiley & Sons, 2011.
- [246] H. Wickham, “testthat: Get Started with Testing,” *R J.*, vol. 3, pp. 5–10, 2011.
- [247] T. Herndon, M. Ash, and R. Pollin, “Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff,” *Camb. J. Econ.*, p. bet075, Dec. 2013.
- [248] J. Leek, “Statisticians and computer scientists – if there is no code, there is no paper | Simply Statistics,” *Simply Statistics*, 23-Jan-2013. .
- [249] M. Delgado-Rodríguez and J. Llorca, “Bias,” *J. Epidemiol. Community Health*, vol. 58, no. 8, pp. 635–641, 2004.
- [250] R. Jaeschke, J. Singer, and G. H. Guyatt, “Measurement of health status: Ascertaining the minimal clinically important difference,” *Control. Clin. Trials*, vol. 10, no. 4, pp. 407–415, Dec. 1989.
- [251] K. W. Wyrwich and F. D. Wolinsky, “Identifying meaningful intra-individual change standards for health-related quality of life measures,” *J. Eval. Clin. Pract.*, vol. 6, no. 1, pp. 39–49, 2000.
- [252] G. R. Norman, J. A. Sloan, and K. W. Wyrwich, “Interpretation of Changes in Health-related Quality of Life: The Remarkable Universality of Half a Standard Deviation. [Miscellaneous Article],” *Med. Care May 2003*, vol. 41, no. 5, pp. 582–592, 2003.
- [253] K. W. M. Wyrwich, N. A. Nienaber, W. M. Tierney, and F. D. Wolinsky, “Linking Clinical Relevance and Statistical significance in Evaluating Intra-Individual Changes in Health-Related Quality of Life,” *Med. Care May 1999*, vol. 37, no. 5, pp. 469–478, 1999.
- [254] A. G. Copay, B. R. Subach, S. D. Glassman, D. W. Polly Jr., and T. C. Schuler, “Understanding the minimum clinically important difference: a

- review of concepts and methods,” *Spine J.*, vol. 7, no. 5, pp. 541–546, Sep. 2007.
- [255] S. J. Kamper, C. G. Maher, and G. Mackay, “Global Rating of Change Scales: A Review of Strengths and Weaknesses and Considerations for Design,” *J. Man. Manip. Ther.*, vol. 17, no. 3, pp. 163–170, 2009.
- [256] S. S. Farivar, H. Liu, and R. D. Hays, “Half standard deviation estimate of the minimally important difference in HRQOL scores?,” *Expert Rev. Pharmacoecon. Outcomes Res.*, vol. 4, no. 5, pp. 515–523, Oct. 2004.
- [257] C. C. Preston and A. M. Colman, “Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences,” *Acta Psychol. (Amst.)*, vol. 104, no. 1, pp. 1–15, Mar. 2000.
- [258] J. A. Krosnick, “Survey Research,” *Annu. Rev. Psychol.*, vol. 50, no. 1, pp. 537–567, 1999.
- [259] S. J. Walters and J. E. Brazier, “Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D,” *Qual. Life Res.*, vol. 14, no. 6, pp. 1523–1532, Aug. 2005.
- [260] M. I. Norton, D. Mochon, and D. Ariely, “The ‘IKEA Effect’: When Labor Leads to Love,” Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 1777100, Mar. 2011.
- [261] B. C. K. Choi, “Future challenges for diagnostic research: striking a balance between simplicity and complexity,” *J. Epidemiol. Community Health*, vol. 56, no. 5, pp. 334–335, May 2002.
- [262] S. Few, *Show me the numbers: designing tables and graphs to enlighten*. Burlingame, Calif.: Analytics Press, 2012.
- [263] J. D. Toms and M. L. Lesperance, “Piecewise regression: a tool for identifying ecological thresholds,” *Ecology*, vol. 84, no. 8, pp. 2034–2041, Aug. 2003.
- [264] V. M. R. Muggeo, “Estimating regression models with unknown break-points,” *Stat. Med.*, vol. 22, no. 19, pp. 3055–3071, Oct. 2003.
- [265] M. Bostock, V. Ogievetsky, and J. Heer, “D3: Data-Driven Documents,” *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, pp. 2301–2309, Dec. 2011.
- [266] J. Ashkenas, M. Bloch, S. Carter, and A. Cox, “The Facebook Offering: How It Compares,” *The New York Times*, 17-May-2012. [Online]. Available: http://www.nytimes.com/interactive/2012/05/17/business/dealbook/how-the-facebook-offering-compares.html?_r=0. [Accessed: 07-Apr-2014].