From DEPARTMENT OF LEARNING, INFORMATICS, MANAGEMENT AND ETHICS
Karolinska Institutet, Stockholm, Sweden

HEALTH-RELATED QUALITY OF LIFE OF DIFFERENT PATIENT GROUPS IN SWEDISH NATIONAL QUALITY REGISTERS

Longitudinal studies of EQ-5D data and application of different value sets

Fitsum Sebsibe Teni

Karolinska Institutet

Stockholm 2021

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-133-6

HEALTH-RELATED QUALITY OF LIFE OF DIFFERENT PATIENT GROUPS IN SWEDISH NATIONAL QUALITY REGISTERS

Longitudinal studies of EQ-5D data and application of different value sets

THESIS FOR LICENTIATE DEGREE

By

Fitsum Sebsibe Teni

To be publicly defended at Karolinska Institutet, Widerström building, room Utsikten, 4th floor.

Friday 26 March 2021, at 9:00 AM

Principal Supervisor:
Kristina Burström, associate professor
Karolinska Institutet
Department of Learning, Informatics, Management and Ethics
Stockholm Centre for Healthcare Ethics
Health Outcomes and Economic Evaluation Research Group

Co-supervisor:
Professor Ola Rolfson
University of Gothenburg
Sahlgrenska Academy
Institute of Clinical Sciences
Department of Orthopaedics

Examination Board:
Knut Stavem, professor
University of Oslo, Norway
Faculty of Medicine
Institute of Clinical Medicine

Jahangir Khan, associate professor
University of Gothenburg
Sahlgrenska Academy
Institute of Medicine
School of Public Health and Community Medicine

Katarina Steen Carlsson, associate professor
Lund University
Faculty of Medicine
Department of Clinical Sciences, Malmö
Health Economics Unit

ABSTRACT

Background: EQ-5D data on patients' health-related quality of life (HRQoL) are collected in roughly 40% of about 100 National Quality Registers (NQRs) in Sweden. The aim of the thesis was to increase knowledge on the use of the EQ-5D-3L in assessing HRQoL of patients in different NQRs and to compare different value sets used to describe HRQoL.

Methods: In Study I, data on patients from 11 NQRs with spine surgery, hip arthroplasty, knee arthroplasty, ankle surgery, anterior cruciate ligament, osteoarthritis, fractures, heart failure, respiratory failure, psoriasis, and rheumatology were included. A total of 266,241 patients from the NQRs and 49,169 members of the general population were included. In Study II, 69,290 patients in the Swedish Hip Arthroplasty Register at baseline and 1-year follow-up, and 21,305 patients at 6-year follow-up were included. Data on demographic and clinical characteristics and EQ-5D-3L data were retrieved in both studies. In Study I, descriptive analyses, Paretian Classification of Health Change (PCHC) and a two-level random intercept model of the impact of diagnoses on EQ VAS scores were performed. EQ-5D-3L indices were calculated using eight value sets from Sweden, Germany, Denmark, and the UK. One-way analysis of variance was used to assess the discriminative ability of the value sets across American Society of Anaesthesiologists (ASA) classes.

Results: In Study I, the pain/discomfort dimension was the dimension with the highest proportion of problems reported both in most registers and in the general population. The highest proportion of improved category of PCHC were reported by patients from the ankle, hip, knee, and spine registers. The two-level random intercept models of EQ VAS score, as predicted by diagnoses, showed patients in most registers, with the exception of fractures, had lower scores than the general population at baseline and at 1-year follow-up. In Study II, all value sets were able to discriminate HRQoL among the ASA classes, and showed the predictive ability of ASA classes on HRQoL.

Conclusion: Both studies demonstrated the importance of the EQ-5D-3L instrument in providing HRQoL data to complement clinical data. The studies documented HRQoL of different patient groups and demonstrated the consistency of the EQ-5D-3L dimensions and the EQ VAS score. All the value sets used to summarize EQ-5D-3L data showed that ASA classes predicted HRQoL and demonstrated ability to differentiate across ASA classes.

Key words: American Society of Anesthesiologists (ASA) physical status classification system, EQ-5D, EQ-5D-3L, EQ VAS, Health-Related Quality of Life (HRQoL), National Quality Register (NQR), Sweden

LIST OF SCIENTIFIC PAPERS

I. Teni FS, Rolfson O, Devlin N, Parkin D, Nauclér E, Burström K,.The Swedish Quality Register (SWEQR) Study Group. A longitudinal study of patients' health-related quality of life using EQ-5D-3L in 11 Swedish National Quality Registers. *Manuscript submitted.*

II. Teni FS, Burström K, Berg J, Leidl R, Rolfson O. Predictive ability of the American Society of Anaesthesiologists physical status classification system on health-related quality of life of patients after total hip replacement: comparisons across eight EQ-5D-3L value sets. *BMC Musculoskeletal Disorders*. 2020; 21:441.

RELATED PUBLICATION

Burström K, Teni FS, Gerdtham UG, Leidl R, Helgesson G, Rolfson O, Henriksson M. Experience-Based Swedish TTO and VAS Value Sets for EQ-5D-5L Health States. *Pharmacoeconomics*. 2020; 38(8):839–856.

CONTENTS

LIST OF ABBREVIATIONS

ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
ASA	American Society of Anesthesiologists physical status classification system
BMI	Body mass index
BOA	Better management of patients with osteoarthritis
HRQoL	Health-related quality of life
IQR	Interquartile range
LTOT	Long-term oxygen therapy
MAE	Mean absolute error
NQR	National Quality Register
OLS	Ordinary least squares
PCHC	Paretian classification of health change
PsoReg	Swedish Registry for Systematic Psoriasis Treatment
QALYs	Quality-adjusted life years
RMSE	Root mean square error
RSE	Robust standard error
SD	Standard deviation
SF-36	36-Item Short-Form Health Survey
SFR	Swedish Fracture Register
SHAR	Swedish Hip Arthroplasty Register
SKAR	Swedish Knee Arthroplasty Register
SOASP	Supported Osteoarthritis Self-Management Programme
Swedankle	Swedish Ankle Registry
SwedeHF	Swedish Heart Failure Registry
Swedevox	Swedish National Register for Respiratory Failure
Swespine	Swedish Spine Register
THR	Total hip replacement
TTO	Time trade-off
UK	United Kingdom
VAS	Visual analogue scale
XBase	Swedish National Anterior Cruciate Ligament Register

1. BACKGROUND

1.1. PATIENT-REPORTED OUTCOMES

Patient-reported outcomes (PROs) provide important insight into outcomes that matter the most to patients. So, their incorporation in clinical trials as well as routine care helps to address views of patients [1]. PROs are described by the United States Food and Drug Administration as information on the health status of patients received from them directly, without interpretation by physicians or others [2]. PROs provide information on patients' perspective on their health condition, not available from other sources [3]. Several benefits of PROs have been described at both individual and population levels. At an individual level, PROs are helpful in informing clinical decisions, in consultation with patients, and focusing on the needs of patients in clinical care. They are also useful in prioritizing resource allocation for patients who need them most [4]. PROs have also been used in screening for diseases as well as tracking progress made by patients [5]. At population level, PROs can inform activities to improve the quality of clinical care provided to patients and in planning treatment [4]. PROs data can also be employed to compare quality of care across various health care providers [6]. Routine collection of data on PROs in registries is also associated with important functions, including assessment of the natural history of diseases and the effectiveness of treatments, as well as in the assessment of quality of care [7].

1.2. HEALTH-RELATED QUALITY OF LIFE

Health-related quality of life (HRQoL) has been defined in many ways with no universally accepted definition. However, definitions of HRQoL are described as having two common aspects: multidimensionality and the inclusion of objective and subjective perspectives [8]. Karimi and Brazier (2016) identified four definitions of HRQoL in the literature, including "how well a person functions in their life and his or her perceived wellbeing in physical, mental and social domains of health" [9–11]. Another definition describes quality of life as a broad concept concerning various factors that affect an individual's life. Hence, HRQoL is described as a concept focused on the factors related to an individual's health [10,12]. A third definition describes HRQoL as "those aspects of self-perceived wellbeing that are related to or affected by the presence of disease or treatment". The fourth definition describes HRQoL as the values or utilities of different health states [10,13].

HRQoL instruments are categorized as generic or condition-specific. Generic measures are employed to capture the physical, psychological, and social components of health defined by the World Health Organization [14]. They can be employed in any population to measure

HRQoL and to compare the burden of diseases across different patient groups [9,15]. Generic measures are useful in clinical and research applications. Clinically, generic measures provide information on the impact of treatments on the health and quality of life of patients. In this way the measures provide information complementary to other clinical data. In research, generic measures complement data from clinical trials by incorporating patients' perspectives on their health. They can also be used in studies on cost-effectiveness of treatments [14]. The most commonly used generic HRQoL measures include EQ-5D and the 36-item Short-Form Health Survey (SF-36), among others. Condition-specific instruments are used to measure changes in dimensions of relevance for specific conditions. They are developed to detect small treatment changes in a manner more sensitive than generic measures. There are numerous condition-specific HRQoL measures available for different diseases [16].

Data on HRQoL are applied in resource allocation as they play a role in the calculation of quality-adjusted life years (QALYs). QALYs is a single measure combining the impact of interventions on survival (life years gained) and on HRQoL over this duration [17]. QALYs are used as outcome measures in economic evaluations, mainly cost-utility analyses where costs and outcomes in the form of QALYs are compared among alternative interventions aiding resource allocation decisions [18].

1.3. EQ-5D

EQ-5D is a brief generic HRQoL instrument widely used to describe and value health across different disease areas. It covers physical, emotional, and social functioning aspects presented in a five-dimension questionnaire [19,20]. The five dimensions, mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, form the descriptive system. The visual analogue scale (EQ VAS), in which respondents report their overall health status, forms a second component of the EQ-5D questionnaire. It is a vertical line ranging from zero, labelled 'worst imaginable health state', to 100, 'best imaginable health state' [19,20]. There are four versions of the EQ-5D, namely EQ-5D-3L, EQ-5D-5L, EQ-5D-Y-3L, and EQ-5D-Y-5L [21–24]. The EQ-5D-3L version has three severity levels in its descriptive system. These are *no problems, some/moderate problems, and confined to bed/unable to/extreme problems.* A self-reported version of EQ-5D-3L is available in more than 180 languages worldwide [21]. The descriptive system of the EQ-5D-5L has five levels of severity: *no problems, slight problems, moderate problems, severe problems, and unable to/extreme problems.* It has been introduced to improve the sensitivity of the EQ-5D-3L [19]. The EQ-5D-Y (EQ-5D-Y-3L) is a version of the EQ-5D-3L questionnaire, used among children and

adolescents. The self-completed version is available in over 90 languages [23]. A five-level severity version of the EQ-5D-Y (EQ-5D-Y-5L) has been developed recently [24].

A health state summarizes the health status of a respondent based on the level of severity of problems reported on the five dimensions in the EQ-5D descriptive system. For example, a person who reported *some problems* regarding mobility, *no problems* with self-care, *some problems* with usual activities, *extreme problems* regarding pain/discomfort, and *moderate problems* in the anxiety/depression dimensions is said to have a health state of 21232. Presented in this manner, a total of 243 unique health states are possible [21].

1.3.1. EQ-5D-3L VALUE SETS

The health status of respondents described as a health state can be summarized into a single EQ-5D-3L index, incorporating the weights provided to the severity level. A set of such weights is known as value sets. Value sets are formulas developed to transform health states into an EQ-5D-3L index, by attaching specific weights to the different severity levels [21]. EQ-5D-3L value sets are developed by valuation studies in the general population using health state valuation methods such as time trade-off (TTO) and the visual analogue scale (VAS). Thus far, more than 30 countries have developed EQ-5D-3L value sets. TTO method of valuation has been employed in most of these value sets [21].

In the TTO valuation, respondents are asked to compare living for a specific period of time with a certain health state to living for shorter duration in full health [25]. The TTO question can be presented through iterative comparisons with shorter durations. The point at which a respondent becomes indifferent between the compared states signifies the value a respondent assigns to the health state in question [26]. Another approach to presenting a TTO question involves a horizontal line marked with duration of time for which respondents indicate the duration at which they are indifferent between living in a specific health state for 10 years, versus in full health. This method has been used in different studies where respondents valued their own health through the TTO method [27–29].

In the VAS valuation method, the EQ-5D VAS is employed for valuation of a list of specified EQ-5D health states. This provides the stated preference of respondents for a specified health states by using the 0 (‘worst imaginable health state’) to 100 (‘best imaginable health state’) range [30]. The EQ VAS, on the other hand, is used by individuals to report on their own health state. However, EQ VAS has also been used to develop value sets in several studies where respondents value their own health states [27,28,31–36]. In economic evaluations, the

use of different value sets has been shown to impact the calculation of QALYs in different studies, leading to different outcomes in various economic evaluation studies [37–39].

1.3.2. PERSPECTIVES IN VALUATION

Hypothetical perspectives of the general public are commonly employed in valuation exercises conducted to develop value sets. In these valuations, respondents are presented with specified health states to value [40,41]. Experience-based health state valuation involves respondents valuing their own health state [42,43]. This perspective has been applied in different studies involving health state valuations [27,29]. Most of the EQ-5D-3L value sets developed so far have employed hypothetical perspectives [21]. However, there is an increasing discussion about and interest in experience-based value sets [40]. Accordingly, several country level value sets have been developed using this perspective as well [27,28,31–33]. There are discussions on the use of hypothetical or experience-based perspectives [44]. Hypothetical and or experience-based perspectives have been shown to lead to different values, making the choice of perspective to consider an important issue to discuss [45].

Different recommendations have been made regarding the choice of valuation perspective when the value sets are aimed to be used for the purposes of economic evaluation. For instance, in Sweden the Dental and Pharmaceutical Benefits Agency recommends the use of experience-based perspective in its guidelines from 2003 and in the updated version in 2017 [46,47]. However, in the UK valuation studies among the general population using hypothetical perspectives are recommended for value sets developed to calculate QALYs for economic evaluation [48].

1.4. NATIONAL QUALITY REGISTERS IN SWEDEN

In Sweden, there are about 100 National Quality Registers (NQRs) which collect data on patients' diagnoses, treatments, and outcomes [49]. These individual-based data are collected and used to improve quality of health care, to monitor compliance with guidelines and for research. The clinical benefits of making use of the data from the NQRs have been established in different clinical areas, showing improvements in health care [50]. In more than 40% of the NQRs, PROs are also collected using the EQ-5D [50].

Most of the NQRs in Sweden were initiated by health professionals. The first NQR dates back to 1975, with the initiation of the Swedish Knee Arthroplasty Register (SKAR). Most of the NQRs had a completeness rate of higher than 80% in 2020 [51]. Furthermore, the NQRs employ validation approaches on the data collected in the registers [50].

1.5. RATIONALE

Studies comparing HRQoL of different patient groups at time points in a comprehensive manner are not common. So, the present thesis helps to fill this gap in the literature. It has been conducted to provide information on the HRQoL of patients in different NQRs in different clinical stages and time points in a manner that reflects the viewpoint of patients regarding their condition, complementing objective clinical information.

There is also research gap regarding the use of different EQ-5D-3L value sets in in summarizing data on HRQoL into an EQ-5D-3L index. Data on HRQoL collected using EQ-5D-3L are commonly presented using value sets chosen from the several EQ-5D-3L value sets available. Value sets could vary based on the method of valuation used and perspective taken, as well as other factors, such as study setting and modelling approaches used. In addition, the extent to which the use of different value sets shows the discriminative abilities of clinical variables categorizing patients into different groups and their predictive abilities remains to be studied. Among such classifications is the American Society of Anaesthesiologists physical status classification system (ASA). It provides a way to make a simple categorization of patients undergoing surgery to determine their operative risk [52].

So, comparison of how using different value sets to summarize HRQoL affects the resulting indices and choices made based on them is important. This has implications on how the discriminative and predictive abilities of ASA classes are shown when using different EQ-5D-3L value sets. Hence, the present thesis was conducted to assess this through making the comparisons across EQ-5D-3L value sets.

2. AIM

The overall aim of the thesis was, using longitudinal studies, to increase knowledge on the use of the EQ-5D-3L in assessing the HRQoL of patients in different National Quality Registers (NQRs) in Sweden, and to compare different value sets from four countries in Europe employed to describe HRQoL.

The specific objectives were:

- To compare problems reported in the five EQ-5D-3L dimensions and EQ VAS scores at baseline and at 1-year follow-up among different patient groups and specific diagnoses in 11 NQRs, and to compare with data from the general population (Study I).
- To compare the discriminative abilities of EQ-5D-3L value sets from Sweden, Germany, Denmark, and the UK in relation to the American Society of Anaesthesiologists physical status classification system (ASA) classes, and to compare the predictive ability of ASA class on HRQoL using these value sets based on data from patients who underwent total hip replacement (THR) in Sweden (Study II).

3. METHODS

3.1. NATIONAL QUALITY REGISTERS IN THE THESIS

A total of 11 NQRs were included in the present thesis. All were employed in Study I, as the aim was to assess the HRQoL of patients from different NQRs at different time points and compare this with the general population. The registers spanned over a variety of disease groups, ranging from those focused mainly on diseases and interventions in the musculoskeletal system, to those focused on others such as heart failure, rheumatology, psoriasis, and respiratory failure. General information on the registers, including year of establishment, diseases areas and interventions covered, enrollment in terms of number of patients, and timeline for EQ-5D data collection are provided in Table 1. The Swedish Hip Arthroplasty Register (SHAR) was used in study II.

3.2. GENERAL POPULATION DATA

In Study I, HRQoL data from the 11 registers were compared with data from the general population in Sweden. This was a survey employed in the development of the experience-based Swedish EQ-5D-3L TTO and VAS value sets [27].

3.3. PARTICIPANTS

A large number of participants were included in both studies, owing to the rich data provided in the Swedish NQRs which have a high completion rate. In Study I, a total of 266,241 patient records from the 11 NQRs with complete data at baseline and at 1-year follow-up, as well as 49,169 from the general population were included. In Study II, 69,290 patient records from the SHAR with complete data at baseline and at 1-year follow-up, and 21,305 records at 6-year follow-up were used.

3.4. STUDY DESIGN

In order to assess HRQoL at baseline and at follow-up stages as well as the changes in between a longitudinal study design was employed in both studies. Accordingly, the HRQoL of patients was assessed at baseline and at 1-year follow-up in both studies. In addition, assessment of the 6-year follow-up was performed in Study II.

Table 1: General information and EQ-5D data collection in the 11 National Quality Registers included in this project [50,53]

Register		Diagnosis/ condition	Intervention[a]	Start year	Unique patients[b]	New entries per year[b]	Follow-up times for EQ-5D
Intervention-based registers	Swedish Spine Register (Swespine)	Spinal stenosis, disk hernia, and other spinal diagnoses	Spine surgery	1993	130,000	10,000	Before surgery, 1, 2, 5, and 10 years after
	Swedish Hip Arthroplasty Register (SHAR)	Hip osteoarthritis and other hip joint diagnoses	Hip replacement	1979	370,000	25,000	Before surgery, 1, 6 and 10 years after
	Swedish Knee Arthroplasty Register (SKAR)	Knee osteoarthritis and other knee joint diagnoses	Knee replacement	1975	220,000	16,000	Before surgery, 1-year after
	Swedish Ankle Registry (Swedankle)	Osteoarthritis and inflammatory conditions in the ankle	Total ankle arthroplasty, ankle arthrodesis procedures	1997	4,000	400	Before surgery, 6 months, 1 and 2 years after
	Swedish National Anterior Cruciate Ligament Register (XBase)	Cruciate ligament injuries	Cruciate ligament surgery	2005	46,000	4,000	Before surgery, 1, 2, 5 and 10 years after
	Better management of patients with osteoarthritis (BOA)	Hip, knee, hand osteoarthritis	Supported Osteoarthritis Self-Management Programme (SOASP)	2008	110,000	18,000	First visit, 3 and 12 months; 1, 2, 3, 4, 5, 6, and 7 years after (100 patients per year are randomized to continued follow-ups)
	Swedish Fracture Register (SFR)	All types of fractures including vertebral/ spinal fractures	Surgical and non-surgical treatments	2011	430,000	82,000	A week before injury (recall) and 1-year after
Diagnosis-based registers	Swedish Heart Failure Registry (SwedeHF)	Chronic heart failure	Pharmacological treatment, physical activity	2003	92,000	9,500	At new visit, within 6 months and 1-year, once every year
	Swedish National Register for Respiratory Failure (Swedevox)	Long-term oxygen therapy (LTOT) due to respiratory failure	LTOT	1987	20,000	1,200	At treatment start, 1-year after
	Swedish Registry for Systematic Psoriasis Treatment (PsoReg)	Psoriasis	Systemic treatment for psoriasis	2007	6,500	500	At each revisit due to psoriasis/ visit to the skin clinic/ telephone conversation with a dermatologist
	Swedish Rheumatology Quality Register (SRQ)	Rheumatic diseases	Medical treatment, rehabilitation	1995	80,000	7,000	At visit

[a] In this research project, unless stated otherwise, the term intervention refers to procedures such as surgeries or other forms of treatments provided to patients in the registers.
[b] Information on the number of unique patients and new entries per year was gathered through communication with the respective registers.

3.5. DATA

In the present thesis, demographic, clinical and PROs data were retrieved from the registers and general population data. In Study I, data on age, sex, height, and weight (body mass index (BMI)), the EQ-5D-3L dimensions, and EQ VAS were retrieved from the 11 registers and data on the general population. Furthermore, data on diagnosis and interventions were retrieved from the registers. The data collected spanned the period from 2001 to 2019.

In Study II, a similar set of data components were retrieved from the SHAR. In addition, data on the laterality of the arthroplasty (right or left side), ASA class of patients, as well as Charnley class were retrieved in Study II. As the study focused on the relationship between ASA class and HRQoL, the data retrieved and included in the study are from the years 2008 (when collection of data on ASA class was initiated in the SHAR) to 2016 [54].

3.6. AMERICAN SOCIETY OF ANESTHESIOLOGISTS PHYSICAL STATUS CLASSIFICATION SYSTEM

In Study II, the predictive effect of ASA class on HRQoL at baseline, and at 1-year and 6-year follow-ups was assessed. The ASA classification system was introduced in 1941 to assess the operative risk of patients prior to surgery [52]. After various updates at different times, ASA class currently categorizes patients preoperatively into six groups based on their operative risk [55]. A person in ASA class I is described as a normal healthy patient, for example a patient that is healthy and non-smoking, with no or minimal use of alcohol. Patients categorized in ASA class II are those with mild systemic disease. Such patients are described as not having considerable functional limitations. A patient with severe systemic disease is categorized in ASA class III. Such patients have considerable functional limitations. Patients in ASA class IV are those with severe systemic disease, which presents a constant threat to life. Examples of such patients are those with ongoing cardiac ischemia or recent myocardial infarction. Patients categorized in ASA class V are patients in a state in which they are not expected to survive without an operation, for example a patient experiencing massive trauma. Patients that are declared brain dead are categorized in ASA class VI [55]. In Study II, records of patients categorized in ASA class I to IV were included.

3.7. VALUE SETS EMPLOYED IN THE THESIS

In Study II, eight EQ-5D-3L value sets were compared to assess their ability to discriminate across ASA classes and show the predictive effect of ASA class on HRQoL. Four of the value sets were developed using the TTO method, and the remaining four were developed

through the VAS method. The value sets are, TTO and VAS from Sweden [27], Germany [32,56], Denmark [41,57], and the UK [41,58].

3.7.1. TIME TRADE-OFF VALUE SETS

Among the TTO value sets used in Study II, the value set from Sweden was developed through an experience-based approach in a study of more than 45,000 participants from the general population in which they valued their own health states. In the study, a total of 148 health states were valued by the participants [27]. The remaining three TTO value sets from Germany, Denmark, and the UK were all developed through a hypothetical perspective, in which participants valued specific health states [56–58]. In the development of the German value set, a study conducted between 1978 and 1998 in northern Germany was used. In the study, 339 participants valued 15 health states each, selected from a total of 36 health states [56]. The value set from Denmark comes from a study conducted in 2000, involving 1,332 participants from the general population. A total of 46 health states were valued, with 16 health states presented to each respondent [57]. The UK EQ-5D-3L TTO value set was developed based on a survey with 3,395 respondents among the general population in the UK. A total of 42 health states were valued, with 13 health states valued by each respondent [58]. The TTO valuation approach employed in this study was also used in the development of the above hypothetical value sets in Germany and Denmark [56,57].

3.7.2. VISUAL ANALOGUE SCALE VALUE SETS

Two of the four VAS value sets in Study II, the Swedish and German EQ-5D-3L value sets, were developed using an experience-based perspective in the valuation studies [27,32]. The value sets from Denmark and the UK were developed using hypothetical perspectives [41]. The Swedish VAS value set was developed using data from the same survey used to develop the TTO value set involving respondents from the general population in 2004 and 2006 [27]. The German experience-based VAS value set was developed based on data from general population surveys in 2006 and 2007 with roughly 2,000 participants [32]. The Danish VAS value set was developed using a study conducted among the general population in Denmark in 1999 to 2000. The total number of participants in the valuation was 1,179 [41]. The VAS value set from the UK was developed in the same study in which the TTO value set was developed. The study was conducted in 1993, covering a total of 45 health states, with each participant valuing 15 health states [41].

3.8. DATA ANALYSIS

In the thesis, a number statistical analyses were performed, including proportion of problems reported on the EQ-5D-3L dimensions; mean and median values with standard deviation (SD) and interquartile range (IQR) respectively; as well as ordinary least squares regressions (OLS) and prediction error measures.

3.8.1. PARETIAN CLASSIFICATION OF HEALTH CHANGE

In Study I, Paretian Classification of Health Change (PCHC), which employs the Pareto principle, was used to assess the longitudinal change in HRQoL as reported in the five dimensions of the EQ-5D questionnaire. Accordingly, four to five categories of PCHC are used to characterize longitudinal change in the EQ-5D dimensions [59,60]. A change is categorized as *improved* when at least one of the dimensions improves without any other worsening. In contrast, a change could be categorized as *worsened* when at least one of the dimensions worsens without any other showing improvement. When there is a mix of improvements and worsening across the dimensions, it is categorized as *mixed.* When there is no change in any of the five dimensions, the categorization, *no change*, is used. A fifth group, *no problem* categorizes a health state which shows no problems at both time points (i.e. health state 11111 at both time points) [59,60].

PCHC provides a way to show, in a simple and clear manner, the overall change in health state related to specific interventions. In studies on the UK National Health Services PRO measures program, differences in the *improved* group after surgical interventions, such as hip and knee arthroplasties, were shown as compared to other interventions, such as hernia and cataract surgeries [60]. However, there are disadvantages to PCHC in relation to the fact that the magnitude of change is not taken into account. In addition, if the proportion of changes categorized as *mixed* dominates, the benefit of using PCHC is reduced [59,60].

3.8.2. EQ VAS SCORE

In Study I, analysis of covariance (ANCOVA) was performed to compare the outcome variable EQ VAS score across categorical variable sex controlled for age. Accordingly, EQ VAS scores were compared between men and women, controlling for age among patients in each register at baseline, at 1-year follow-up, and the change in EQ VAS score, as well as among the general population. ANCOVA allows accounting for a third variable, called covariates, when assessing a relationship between an independent and a dependent variable. To do this, ANCOVA controls the relationship between the covariate and the dependent variable statistically [61].

In comparing how diagnoses across the different patient groups predict EQ VAS score, a two-level random intercept model was performed. In such models, regression analysis takes two levels into account in the model. The dependent variable (EQ VAS score) is clustered into different categories based on another variable (register). In such a random intercept model, the intercept is allowed to vary across the categories of the variable [62]. In Study I, a two-level random intercept model was used in comparing the predictive effect of different diagnoses on EQ VAS score, the outcome variable, across the registers. This helped to account for the difference in EQ VAS score across the diagnoses, categorized by register. The model takes the difference across registers into account as one level.

3.8.3. BASELINE CHARACTERISTICS OF PARTICIPANTS (STUDY II)

In Study II, Cochran-Armitage test was employed with the aim of assessing the uniformity of distribution of categorical demographic and clinical characteristics of patients in the SHAR across ASA classes at baseline. The Cochran-Armitage test is useful in assessing linear trends in two-way tables between a binary variable and an ordinal variable [63,64]. The Kendall rank correlation test was used to assess the uniformity of the distribution of the proportion of patients, in two ordinal variables (demographic/clinical characteristic with ASA class) are compared. This test is used in assessing the relationship between two ordinal variables [65].

Effect sizes for the statistical tests - Cochran-Armitage and Kendall rank correlation test - were determined by calculating the values for Cramer's V and Kendall tau, respectively. Cramer's V is useful in providing the magnitude of two-by-two chi-square tests, adding more information than just statistical significance found in the chi-square tests [66]. Similarly, the Kendall tau coefficient provides the magnitude of correlation for ordinal variables [65].

3.8.4. ANALYSIS ON DISCRIMINATIVE ABILITY OF ASA CLASSES

One-way analysis of variance (ANOVA) is used to compare mean values of a continuous variable across two or more groups, assessing statistical significance of the difference [67]. In Study II, one-way ANOVA was employed to assess the discriminative abilities of the different EQ-5D-3L value sets across ASA classes. In addition, the Kruskal Wallis test was conducted to confirm whether the one-way ANOVA was robust, as the assumption of homogeneity of variance was violated. The effect size measures for the tests one-way ANOVA test was performed using eta squared [68]. The Kruskal Wallis test provides an important alternative in comparing the mean value of an outcome variable across several groups, when the distribution is non-normal or unknown [69].

3.8.5. PREDICTIVE ABILITY OF ASA CLASSES ON HRQOL

In Study II, ordinary least squares (OLS) models were employed in the assessment of the predictive ability of ASA classes on HRQoL of patients in the SHAR, as described by the different EQ-5D-3L value sets. Owing to the presence of heteroscedasticity, robust standard error (RSE) (heteroscedasticity-consistent) was used instead of standard error. RSE is useful in remedying the problems of violation of the assumption of equal variance [70]. Both unadjusted models and those adjusted for demographic, clinical and PROs were performed. In the assessment of prediction error, mean absolute error (MAE) and root mean square error (RMSE) were employed, compared across EQ-5D-3L value sets. MAE describes the mean of the difference between predicted and observed values in the regression model [71]. Root mean square error is the square root of the mean of squared error between observed and predicted values [72]. To take account of the variation in the interval of the different EQ-5D-3L value sets, normalized versions of MAE and RMSE were also calculated.

3.9. ETHICAL CONSIDERATIONS

Study I was approved by the Regional Ethics Review Board in Gothenburg (# 1185-18/2019-00812 and # 2020-04369) for the data retrieved from the NQRs and by the Swedish Ethical Authority (# 2020-03090) for the data from the general Swedish population survey. Study II was approved by the Regional Ethics Review Board in Gothenburg (# 271–14). The general population data are based on information from individuals who agreed to participate after having been informed about the study. The ethical review committees assessing the applications for the studies in the thesis, based on data from NQRs, waived the need for informed consent by individual patients whose data were used in the study, considering the issues of studies on NQRs discussed below.

In the present thesis, informed consent is among the ethical issues relevant to the studies, owing to the use of data from many NQRs. Informed consent from participants in studies is an important component in the conduct of ethically sound research. According to the Declaration of Helsinki, researchers are required to seek the consent of research participants, after first providing them with sufficient comprehensible information. The information should address the benefits and risks associated with participation in the research [73].

In the present thesis, as regards the data retrieved from registers in Study I and Study II, it was impractical to secure informed consent from each individual for which data were used. A number of challenges related to seeking informed consent in register-based research have been discussed [74]. The large sample size of the data from the registers is one of the reasons which made it both difficult and costly to seek consent from individuals. Another reason

relates to the possibility that the participation of individuals in the research might be reduced considerably after seeking informed consent from each patient. As data collected over several years are retrieved from registers, there is a possibility that patients may have already passed away, making it impossible to get informed consent [74].

In mitigating the problem of not securing informed consent for the data from registers, it is important that patients are informed about the possibility of their data being used for research purposes when they are registered in NQRs. At this stage, patients are also informed that they can opt out of the registers at any time and have their records deleted, according to the Swedish Patient Data Act [75].

4. RESULTS

4.1. HRQOL OF PATIENTS IN 11 NQRS (STUDY I)

In most of the registers, the highest proportions of problems were reported in the pain/discomfort dimension at baseline and at 1-year follow-up, and in the general population, with the exception of patients with respiratory failure. The highest proportions of severe problems were also reported in the pain/discomfort dimension. A similar pattern in problems reported was also shown across the EQ-5D-3L dimensions for the different diagnoses within the registers.

PCHC, which shows the pattern of changes in patients' health states across time, showed the highest proportions of *improved* category recorded among patients with diagnoses in the intervention-based registers. These include patients in the ankle, hip, knee, and spine registers. The exceptions were patients in the respiratory and heart failure as well as fracture (where baseline data are collected with a recall of the time before injury) registers.

Over the 1-year follow-up, improvements in EQ VAS score were recorded in almost all the registers, with the exception of patients with respiratory failure and those with fractures. In most registers - with the exception of fracture and respiratory failure patients at baseline, as well as BOA, fracture, heart failure, and respiratory failure patients at 1-year follow-up - women had lower mean EQ VAS scores, controlling for age based on ANCOVA tests. A similar pattern was observed by diagnosis at both time points. Mean EQ VAS scores by diagnosis at baseline, at 1-year follow-up, and the change over the period, had similar patterns as shown in the registers. Overall, mean EQ VAS scores were closer within a register than across registers.

On the basis of the results of a two-level random intercept model, patients with diagnoses in almost all registers showed lower EQ VAS scores at baseline in comparison with the general population, with the exception of those in the fracture register. Similarly, diagnoses in most registers remained at lower EQ VAS scores than the general population at 1-year follow-up. Mean EQ VAS score in the different PCHC categories showed an increase over the follow-up period for those in the *improved* group, and a decrement in the *worsened* group. Furthermore, mean EQ VAS scores in nine selected health states were consistent with the severity of the health states both at baseline and 1-year follow-up.

4.2. PREDICTIVE ABILITY OF ASA CLASSES ON HRQOL (STUDY II)

At both baseline and 1-year follow-up, the highest proportion of any problems, as well as severe problems, were reported in the pain/discomfort dimension. A similar pattern of reported problems was also shown across ASA classes. Mean EQ VAS score among SHAR patients showed improvement, from 56.0 at baseline to 76.7 at 1-year follow-up. The mean and median EQ VAS scores in the different ASA classes showed lower EQ VAS scores in higher ASA classes.

Indicating the discriminative abilities of the different value sets, mean EQ-5D-3L indices, calculated based on the value sets, differed across ASA classes. The highest indices were found in ASA class I, decreasing over the classes, with the lowest found in ASA Class IV. This was shown to be statistically significant at both baseline and 1-year follow-up. The EQ-5D-3L index based on the Swedish TTO showed the highest indices at baseline and at 1-year follow-up, with indices comparable to the Swedish TTO shown by the other value sets at 1-year follow-up.

In assessing the predictive effect of ASA classes on HRQoL, at baseline and at 1-year follow-up, OLS models of all EQ-5D-3L indices based on all value sets and EQ VAS scores showed that ASA class predicts HRQoL regardless of the value set employed to calculate them. Both at baseline and 1-year follow-up, decrements associated with ASA classes were consistent. Larger decrements were found in the higher ASA classes. The pattern was similar, both in unadjusted models and models adjusted for age group, sex, BMI, laterality, Charnley class, and hip pain levels.

The R^2 values, indicating the amount of variation in HRQoL explained by ASA class membership, rose from 1-2% at baseline to 4-5% at 1-year follow-up. The prediction error measures, MAE and RMSE, and the normalized versions of the measures showed small variations across the different OLS models. Overall, the findings at 6-year follow-up were generally comparable to those at 1-year follow-up.

5. DISCUSSION

The two studies demonstrated the usefulness of the EQ-5D-3L instrument in reflecting the view of patients on their overall health through reporting their HRQoL. In Study I, the EQ-5D-3L has been used to compare the HRQoL of patients across 11 NQRs. Similarly, in Study II, the EQ-5D-3L has been used to show how the clinical status of patients predicts their HRQoL, through assessing the relationship between ASA classes and HRQoL of patients in the SHAR. In doing this, a comparison of the discriminative abilities of the different EQ-5D-3L value sets across ASA classes was performed, thereby assessing the implication of which value set is used in calculating the EQ-5D-3L indices.

5.1. MAIN FINDINGS AND COMPARISON TO OTHER STUDIES

5.1.1. HRQOL OF PATIENTS IN 11 NQRS (STUDY I)

The proportion of problems reported in the pain/discomfort dimension, highest among patients and in the general population, corresponded to pain symptom experienced in different patient groups. Accordingly, larger proportions were reported among patients in the intervention-based registers, which include spine, hip and knee registers where pain symptoms occur in the diseases [76–79]. The proportion of problems reported in the different registers could also be related to the types of symptoms of the diseases in the different registers. This could be exemplified by the lower proportion of problems in the pain/discomfort and mobility dimensions in the psoriasis register as compared to patients in the spine register, owing to the difference in the symptoms and disease conditions. This showed an expected difference in line with the difference in the nature of the diseases and how this affects the pertinent EQ-5D-3L dimensions.

From baseline to 1-year follow-up, the problems reported reduced in most registers, with the greatest improvement found among patients in the intervention-based registers, including spine, hip, and knee. Similarly, PCHC categories across different diagnoses demonstrated a similar pattern with a higher proportion in the *improved* group among patients with diagnoses in many of the intervention-based registers. PCHC was used as important information, to show the change in the pattern of problems reported over time, in studies of diseases, such as heart failure, prostate cancer, diabetes mellitus, and overactive bladder [80–84]. The change in the problems reported in the different dimensions and PCHC categories followed the characteristics of diseases and the respective interventions made.

EQ VAS provides information on a respondent's view of their own overall health. In this study, mean EQ VAS score showed improvement over the period up to the one-year follow-

up among patients across most of the registers. Among the exceptions were patients in the fracture register, which relates to the data collection approach in which HRQoL status prior to the occurrence of the fracture was reported for baseline using recall technique. The improvement in mean EQ VAS score seemed to be in line with improvements following interventions at baseline. Mean EQ VAS scores among patients in most registers, at both time points, were lower than in the general population [27]. The improvement in EQ VAS scores reflected the expected change from baseline to 1-year follow-up while remaining lower than the relatively healthier general population.

Mean EQ VAS scores also showed consistent pattern to the problems reported on the EQ-5D-3L dimensions, with lower mean EQ VAS scores among patients who reported higher proportions of problems. The change in mean EQ VAS score was shown to be consistent with changes in problems reported on the EQ-5D-3L dimensions as well. This is in line with previous studies demonstrating such consistency [85,86]. The pattern of consistency between the EQ-5D-3L dimensions and EQ VAS scores was similar across PCHC categories. Mean EQ VAS scores of selected health states showed similar consistency with the severity of the health states. A similar pattern was reported in another study, in which mean EQ VAS scores were in line with the severity of health states [87]. Another study of patients with Parkinson's disease showed the importance of the EQ-5D-3L dimensions and the EQ VAS in documenting the HRQoL of patients and assessing the severity of diseases [88].

5.1.2. PREDICTIVE ABILITY OF ASA CLASSES ON HRQOL (STUDY II)

The fact that the highest proportion of problems are reported in the pain/discomfort dimension demonstrate that pain is an important symptom among patients with osteoarthritis of the hip [89]. This is particularly the case at baseline and remains so to a greater extent at 1-year follow-up relative to problems in other dimensions. The study showed that EQ-5D-3L indices calculated based on all the EQ-5D-3L values sets and on the patient-reported EQ VAS score differentiated among the ASA classes. Among the value sets, Swedish TTO had higher EQ-5D-3L indices in all ASA classes than in the other value sets at baseline. At 1-year follow-up the difference decreased, which could be due to the difference between the Swedish experience-based TTO value set and the hypothetical German, Danish, and UK TTO value sets [41,56–58]. The difference could be attributed to the fact that in the hypothetical value sets, much lower indices are assigned to severe health states than in the Swedish TTO; however, healthier states are valued much closer to the Swedish TTO. Despite variation in the indices resulting from the different value sets, discriminative abilities of ASA class remains demonstrated in all of them.

EQ-5D-3L indices based on all value sets and on patient reported EQ VAS score showed the predictive effect of ASA classes on HRQoL. This demonstrated that ASA class influences HRQoL both at baseline and even to a higher extent postoperatively. A similar finding was reported in a study which assessed a relationship between ASA class and functional recovery following surgery for hip fracture [90]. Higher decrements in HRQoL among more severe ASA classes were also shown in yet another study on surgical interventions for hip fracture [91]. In the comparison of the different value sets, decrements in HRQoL across ASA classes were the greatest among the hypothetical TTO value sets at baseline. At 1-year follow-up, the highest decrements were in the Danish VAS and the value sets from the UK. The difference could be attributed to the wider interval in the EQ-5D-3L index of the hypothetical value sets.

The R^2 of all the regression models of the different value sets showed improvement from baseline to 1-year follow-up, explaining more variation by membership to ASA classes. Larger R^2 values are shown in the patient-reported EQ VAS score and VAS value sets, which could be due to the VAS method of valuation study. In terms of the prediction error, the Swedish TTO had the lowest MAE and RMSE at both time points, while normalized MAE and RMSE were lowest among Danish VAS at baseline and Danish TTO at 1-year follow-up.

The findings of the study regarding the comparison of value sets in relation to discriminative and predictive abilities of ASA classes HRQoL showed varying indices and error measures. However, both the discriminative and predictive abilities of ASA were demonstrated across the value sets. So, the findings do not indicate to a suggestion of a particular value set but showed that an understanding of features of value sets used to calculate HRQoL is important.

5.2. STRENGTHS AND LIMITATIONS

In this thesis there are a number of strengths in relation to the data sets and methods employed. One of the important strengths of the studies concerns the large size of patient records in the registers and from the general population. This contributed to addressing the research questions in each study with increased power. Another strength of the thesis is related to the comprehensive study across many patient groups in Study I. This has contributed to the literature by documenting HRQoL in several patient groups at once. Another strength of the thesis is that a longitudinal study design was employed in both studies. In comparison to cross-sectional studies, longitudinal ones provides an advantage of assessing change over time through measurements taken at repeated follow-ups [92]. Accordingly, in this thesis, longitudinal design helped in showing HRQoL status at baseline and follow-up stages, as well as the changes over the follow-up period.

The studies also have a number of limitations that should be taken into consideration when interpreting the findings. Variations in the mode of data collection on PROs across the NQRs, and over time, constitute a limitation which could lead to differences in HRQoL measures. The inclusion of only patient records with complete data both at baseline and 1-year follow-up in both studies which excludes patients with only baseline or 1-year follow-up data is another limitation. This could potentially lead to bias with a possible variation in the measurement of HRQoL between the groups included and those not included. However, analysis of this in Study I showed that findings concerning patients with data only at baseline or at 1-year follow-up only are comparable to those with complete baseline and 1-year follow-up data. Another limitation relates to the time difference in the time of the studies which developed the value sets compared in Study II. The time of data collection for the value sets development studies ranged from the 1990s to 2000s. This could contribute to differences in valuations which are unaccounted for in the study. The use of different regression models in the development of the value sets that were compared could also lead to differences in the indices which was not assessed in this study. Such variation due to modeling has been reported in a study comparing different value sets [93].

5.3. LESSONS LEARNED

Several important lessons are learned from the thesis regarding the EQ-5D-3L questionnaire and the value sets used to summarize such data into a single index. One is the importance of the EQ-5D-3L questionnaire in providing complementary information to the clinical data by adding the views of patients on their health status. Another lesson concerns the consistency between problems reported on the EQ-5D-3L dimensions and the EQ VAS components, as demonstrated in the different patient groups. This highlights the importance of the often underused EQ VAS component.

Yet another important lesson is that the thesis demonstrated that it is crucial to have a clear understanding of the value sets employed to calculate EQ-5D-3L indices. This is because several features such as valuation methods employed, perspectives taken, the setting in which the value set was developed, as well as others factors such as modeling approaches need to be considered. This is important as choice of value sets might have implications on indices calculated which could be used to inform resource allocation decisions through economic evaluations.

Furthermore, the thesis can exemplify the way in which data sets from the Swedish NQRs, which hold a wealth of information on different patient groups, could be used in a manner

that adds evidence to the literature in several areas. The collaboration across several registers in conducting a large study demonstrated in the present thesis also indicates to the potential of producing research outputs that could provide useful inputs for practical applications.

5.4. IMPLICATIONS OF THE THESIS

The information from the two studies in the present thesis demonstrate the importance of the EQ-5D-3L questionnaire in documenting the HRQoL of patients, as well as approaches for summarizing and presenting HRQoL data. Among the specific implications of the findings in Study I in the provision of crucial information by documenting data on the HRQoL of several patient groups and diagnoses within the registers, using large data sets collected alongside routine clinical practice. This could be helpful in improving the care provided to patients, by using HRQoL data together with clinical data.

The study also demonstrated the consistency between the EQ-5D-3L dimensions and EQ VAS scores with proportions of problems reported going in line with mean EQ VAS scores. In relation to this, the study added to the evidence that EQ VAS score is an important measure of self-reported health. This is demonstrated by its consistency with the EQ-5D-3L descriptive system, as well as the nature of diagnoses and patient groups. Thus, the EQ VAS score reflects patients' health status, which reinforces the use of EQ VAS to compare HRQoL within and between patient groups. Accordingly, there is a need for further investigation into EQ VAS in reporting HRQoL and as a basis to explore patient value sets. This will give insight on exploration of potential value sets through experience-based perspectives. Such studies could contribute to discussions of public health importance on approaches to health care resource allocation.

Study II addressed the aspect of choice of approach in summarizing the data on the EQ-5D-3L descriptive system using value sets and how the choice affects the resulting HRQoL measurement. Accordingly, the study's findings provided information on how the use of different EQ-5D-3L value sets impacts the measurement of HRQoL in showing the discriminative abilities of ASA classes. In relation to this, the study demonstrated the predictive ability of ASA classes on HRQoL, adding useful information for clinical practice which could help in determining care required for groups of patients in different ASA classes.

6. CONCLUSIONS

In this thesis, the importance of the EQ-5D-3L instrument in measuring the HRQoL of patients to provide the perspective of patients on their health as an addition to objective clinical data. Study I documented comprehensive information on the HRQoL of patients in the 11 NQRs, with comparisons made with data from the general population. The fact that HRQoL reported across patient groups varied based on the symptoms manifested by the conditions was demonstrated in the study. In relation to this, the consistent relationship between problems reported on the EQ-5D-3L dimensions and EQ VAS scores was demonstrated across the different patient groups and in the general population.

In Study II, the discriminative ability of all value sets investigated across ASA classes was demonstrated. Hence the HRQoL reported using each of the value sets showed lower values in more severe ASA classes and vice versa. In relation to this, the study also added to the evidence on the predictive ability of ASA classes on HRQoL at baseline as well as at follow-ups. This was demonstrated in all the value sets employed to calculate EQ-5D-3L indices. The findings do not indicate to a specific value set among the compared ones but the importance of a clearer understanding of the value sets employed in calculating EQ-5D-3L index as it can be affected by the different features of the value set.

7. SUGGESTIONS FOR FUTURE RESEARCH

This thesis investigated the HRQoL of different patient groups and compared value sets employed to summarize such HRQoL data. Evidence from such large studies is not commonly available in the literature. Therefore, future research should focus on more such studies, especially regarding documenting HRQoL of different patient groups in various settings using the EQ-5D instrument in its different versions. This will help document HRQoL of a wide range of patient groups to understand the views of patients on their health.

The consistency of EQ VAS scores with the problems reported in the EQ-5D-3L dimensions was demonstrated in the thesis. This indicates the importance of exploring the relationship between the EQ-5D-3L dimensions and EQ VAS score to assess their potential for developing patient VAS value sets. This will add to the literature on experience-based value sets based on patient HRQoL data. Based on findings from such studies the potential of experience-based value sets to be applied for decision making could be explored.

In terms of comparing the implications of using different value sets on the resulting EQ-5D-3L indices, studies comparing more value sets and using other patient groups would provide a fuller picture of differences across value sets. Evidence from such studies will be useful in making informed choice when using EQ-5D-3L value sets.

Conducting the above suggested studies on HRQoL data collected using the five-level version of the EQ-5D questionnaire would be useful to document differences from the three-level version. In addition, owing to the increased use of the EQ-5D-5L questionnaire, evidence from studies on this version will provide important information on the characteristics of value sets and the potential of patient value sets.

8. ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to all who supported me throughout this thesis! In particular I would like to thank:

My principal supervisor Kristina Burström for the opportunity to study in this research project on the EQ-5D questionnaire. I am thankful for the lessons I learned from your expertise and wealth of experience on the EQ-5D. It has been a very nice learning experience to work on my MSc thesis, and now on the licentiate degree project under your supervision.

Ola Rolfson, my co-supervisor, for the opportunity to study in this research project encompassing many registers. Thank you very much! Beginning from the MSc thesis for which you were my co-supervisor as well, I have benefited immensely from your expertise, insights, comments and suggestions over the years in the licentiate project.

Anita Berglund, my mentor, for taking your time to support me through my journey in the thesis.

Emma Nauclèr, for facilitating access to the data sets for the thesis, and your advice on statistical analyses. I would also like to thank Jonatan Nåtman and Johanna Vinblad, for facilitating access to data sets from the different registers.

Nancy Devlin and David Parkin, for your insightful comments and suggestions at the different stages of the thesis.

Co-authors of the studies in the thesis, Jenny Berg, Nancy Devlin, Reiner Leidl, Emma Nauclèr, David Parkin and members of the Swedish Quality Register Study Group, Allan Abbott, Magnus Ekström, Magnus Forssblad, Peter Fritzell, Åsa Jonsson, Mikael Landén, Michael Möller, Malin Regardt, Björn Rosengren, Marcus Schmitt-Egenolf, Johanna Vinblad, and Annette W-Dahl, for your crucial feedback in different stages of the thesis. I would like to thank the National Quality Registers, spine, hip, knee, ankle, cruciate ligament injury, BOA, fracture, heart failure, respiratory failure, psoriasis and rheumatology registers for the collaboration and providing access to the data sets employed in the thesis.

Asli Kulane, Nicola Orsini, Megan Doheny and Mimmi Åström, for taking your time to read and comment on the compilation thesis summary (‘kappa’) at the pre-licentiate seminar.

Members of the Health Outcomes and Economic Evaluation Research Group, Jenny Berg, Kristina Burström, Shaofan Chen, Haiting Jiang, Sophie Langenskiöld, Peter Lindgren, Jonas Lundkvist, Ola Rolfson, Sun Sun, and Mimmi Åström, for the lessons I learned through discussions at the group meetings and the valuable comments I received on the thesis.

Bo Burström, for the opportunity to attend and present studies at the research group meetings at the Equity and Health Policy (EHP). I would also like to thank all the members of the Research Group for the lively and informative meetings and the feedback I received.

Members of the Stockholm Centre for Healthcare Ethics, especially Gert Helgesson, Head of the Centre, Annelie Jonsson, Niklas Juth and Tomas Månsson, for the opportunity to work and study at the centre and for the lessons from seminars at the centre. I am grateful to Niklas Juth for agreeing to chair the licentiate seminar.

Carl Johan Sundberg, Head of the Department of LIME, for the opportunity to be a member of LIME, and the support to undertake and complete my studies at the department. I would also like to thank Maria Herdland, Marie Lind, and Therese Wahlström, for your support during the programme. I am grateful to Tomas Månsson and Ingrid Smedberg for your continued support in the doctoral education through the changes in the programme until completion. I would like to thank Ludvig Andersson, Erik Attoff and Ludwig Berglund for the support on IT related issues.

Mimmi Åström, for your continued support beginning from the MSc thesis project to the Licentiate project. I am grateful for your kindness and wish you all the best in your PhD project!

Shaofan Chen, for sharing your experience and the nice conversations while sharing an office. Congratulations again for the successful completion of your PhD. I would also like to thank Javad Jafari, for the nice time sharing an office and I wish you all the best in your PhD project.

Colleagues at LIME, particularly Jayne Alfredsson, Stefano Bonacina, Filip Gedin, Eva Hagel, Matilda Hagman, Ulrika Schüldt Håård, Daniel Olsson, Vibeke Sparring and Mesfin Tessma, and fellow doctoral students, Cecilia Dahlgren, Olivia Ernstsson, Fanny Goude, Max Kleijberg, and Sofi Varg, for the nice conversations and discussions.

My friends, Lulu Gumbo, Bethelhem Hailu, Joseph Kazibwe, Michael Mammo, Habtamu Sisay, Phuong Tran, and Befikadu Legesse Wubishet, for your support and encouragement throughout the programme.

My family, Fikirte Girma, Sebsibe Teni, Samrawit Sebsibe, Brook Sebsibe, Yosef Alemu and Gelane, for always being there. Thank you!

I gratefully acknowledge the financial support from the EuroQol Research Foundation and Stockholm County Council.

9. REFERENCES

1. Tamura MK. Introduction to Patient-Reported Outcomes Perspectives Series. CJASN. American Society of Nephrology; 2017;12:1881–1881.

2. US Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. [Internet]. 2009. Available from: https://www.fda.gov/media/77832/download

3. Rothman ML, Beltran P, Cappelleri JC, Lipscomb J, Teschendorf B, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. Patient-reported outcomes: conceptual issues. Value Health. 2007;10 Suppl 2:S66-75.

4. PRO secretariat. The Danish National Work on Patient Reported Outcomes [Internet]. 2019. Available from: https://pro-danmark.dk/da/pro-english

5. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. The Lancet. Elsevier; 2009;374:369–70.

6. Snyder CF, Jensen RE, Segal JB, Wu AW. Patient-reported outcomes (PROs): putting the patient perspective in patient-centered outcomes research. Med Care. 2013;51:S73–9.

7. Gliklich RE, Dreyer NA, Leavy MB. Use of Patient-Reported Outcomes in Registries. Registries for Evaluating Patient Outcomes: A User's Guide. 3rd ed. Rockville (MD): Agency for Healthcare Research and Quality; 2014. p. 93–125.

8. de Wit M, Hajos T. Health-related quality of life. Encyclopedia of Behavioral Medicine. Springer; 2013. p. 929–31.

9. Hays R, Reeve B. Measurement and Modeling of Health-Related Quality of Life. International Encyclopedia of Public Health. 2008.

10. Karimi M, Brazier J. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? PharmacoEconomics. 2016;34:645–9.

11. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA. 1995;273:59–65.

12. Torrance GW. Utility approach to measuring health-related quality of life. Journal of Chronic Diseases. 1987;40:593–600.

13. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. Soc Sci Med. 1995;41:1383–94.

14. Maciejewski M. Generic measures. Understanding Health Care Outcomes Research. 2nd ed. Jones and Bartlett Publishers; 2006. p. 123–64.

15. Kaplan RM, Ries AL. Quality of Life: Concept and Definition. COPD: Journal of Chronic Obstructive Pulmonary Disease. Taylor & Francis; 2007;4:263–71.

16. Atherly A. Condition-specifc measures. Understanding Health Care Outcomes Research. 2nd ed. Jones and Bartlett Publishers; 2006. p. 165–83.

17. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010;96:5–21.

18. Robinson R. Cost-utility analysis. BMJ. 1993;307:859–62.

19. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. Appl Health Econ Health Policy. 2017;15:127–37.

20. Gusi N, Olivares PR, Rajendram R. The EQ-5D Health-Related Quality of Life Questionnaire. Handbook of Disease Burdens and Quality of Life Measures. 2010.

21. EuroQol Research Foundation. EQ-5D-3L User Guide [Internet]. 2018 [cited 2020 Nov 9]. Available from: https://euroqol.org/publications/user-guides.

22. EuroQol Research Foundation. EQ-5D-5L User Guide [Internet]. 2019 [cited 2020 Nov 9]. Available from: https//euroqol.org/publications/user-guides.

23. EuroQol Research Foundation. EQ-5D-Y User Guide [Internet]. 2020. Available from: https://euroqol.org/publications/user-guides.

24. Kreimeier S, Åström M, Burström K, Egmar A-C, Gusi N, Herdman M, et al. EQ-5D-Y-5L: developing a revised EQ-5D-Y with increased response categories. Qual Life Res. 2019;28:1951–61.

25. Torrance GW, Thomas WH, Sackett DL. A Utility Maximization Model for Evaluation of Health Care Programs. Health Serv Res. 1972;7:118–33.

26. Lugnér AK, Krabbe PFM. An overview of the time trade-off method: concept, foundation, and the evaluation of distorting factors in putting a value on health. Expert Review of Pharmacoeconomics & Outcomes Research. Taylor & Francis; 2020;20:331–42.

27. Burström K, Sun S, Gerdtham U-G, Henriksson M, Johannesson M, Levin L-Å, et al. Swedish experience-based value sets for EQ-5D health states. Qual Life Res. 2014;23:431–42.

28. Burström K, Teni FS, Gerdtham U-G, Leidl R, Helgesson G, Rolfson O, et al. Experience-Based Swedish TTO and VAS Value Sets for EQ-5D-5L Health States. Pharmacoeconomics. 2020;38:839–56.

29. Burström K, Johannesson M, Diderichsen F. A comparison of individual and social time trade-off values for health states in the general population. Health Policy. 2006;76:359–70.

30. EuroQoL Group. Terminology [Internet]. EQ-5D. 2020. Available from: https://euroqol.org/support/terminology/

31. Sun S, Chen J, Kind P, Xu L, Zhang Y, Burström K. Experience-based VAS values for EQ-5D-3L health states in a national general population health survey in China. Qual Life Res. 2015;24:693–703.

32. Leidl R, Reitmeir P. A value set for the EQ-5D based on experienced health states: development and testing for the German population. Pharmacoeconomics. 2011;29:521–34.

33. Leidl R, Reitmeir P. An Experience-Based Value Set for the EQ-5D-5L in Germany. Value in Health. 2017;20:1150–6.

34. Gutacker N, Patton T, Shah K, Parkin D. Using EQ-5D Data to Measure Hospital Performance: Are General Population Values Distorting Patients' Choices? Med Decis Making. SAGE Publications Inc STM; 2020;40:511–21.

35. Nemes S, Burström K, Zethraeus N, Eneqvist T, Garellick G, Rolfson O. Assessment of the Swedish EQ-5D experience-based value sets in a total hip replacement population. Qual Life Res. 2015;24:2963–70.

36. Wu XY, Ohinmaa A, Johnson JA, Veugelers PJ. Assessment of children's own health status using visual analogue scale and descriptive system of the EQ-5D-Y: linkage between two systems. Qual Life Res. 2014;23:393–402.

37. Bernert S, Fernández A, Haro JM, König H-H, Alonso J, Vilagut G, et al. Comparison of Different Valuation Methods for Population Health Status Measured by the EQ-5D in Three European Countries. Value in Health. 2009;12:750–8.

38. Knies S, Evers SMAA, Candel MJJM, Severens JL, Ament AJHA. Utilities of the EQ-5D. Pharmacoeconomics. 2009;27:767–79.

39. Dongen JM van, Ben ÂJ, Finch AP, Rossenaar MMM, Biesheuvel-Leliefeld KEM, Apeldoorn AT, et al. Assessing the Impact of EQ-5D Country-specific Value Sets on Cost-utility Outcomes. Med Care. Lippincott Williams and Wilkins; 2020;59:82–90.

40. Versteegh MM, Brouwer WBF. Patient and general public preferences for health states: A call to reconsider current guidelines. Soc Sci Med. 2016;165:66–74.

41. Szende A, Oppe M, Devlin N, editors. EQ-5D Value Sets: Inventory, Comparative Review and User Guide. Springer Netherlands; 2007.

42. Cubi-Molla P, Shah K, Burström K. Experience-Based Values: A Framework for Classifying Different Types of Experience in Health Valuation Research. Patient. 2018;11:253–70.

43. Brazier J, Rowen D, Karimi M, Peasgood T, Tsuchiya A, Ratcliffe J. Experience-based utility and own health state valuation for a health state classification system: why and how to do it. Eur J Health Econ. 2018;19:881–91.

44. Stamuli E. Health outcomes in economic evaluation: who should value health? British Medical Bulletin. 2011;97:197–210.

45. Rand-Hendriksen K, Augestad LA, Kristiansen IS, Stavem K. Comparison of hypothetical and experienced EQ-5D valuations: relative weights of the five dimensions. Qual Life Res. 2012;21:1005–12.

46. The Pharmaceutical Benefits Board. General guidelines for economic evaluations from the Pharmaceutical Benefits Board [Internet]. The Pharmaceutical Benefits Board; 2003. Available from: https://www.tlv.se/download/18.2e53241415e842ce95514e9/1510316396792/Guidelines-for-economic-evaluations-LFNAR-2003-2.pdf

47. The Dental and Pharmaceutical Benefits Agency (TLV). General guidelines for economic evaluations from The Dental and Pharmaceutical Benefits Agency (updated 2017) (Swedish) [Internet]. 2017. Available from: https://www.tlv.se/download/18.467926b615d084471ac3230c/1510316374332/TLVAR_2017_1.pdf

48. National Institute for Health and Care Excellence. Incorporating economic evaluation. Developing NICE guidelines: the manual [Internet]. 2014. p. 124–59. Available from: https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869

49. Nationella Kvalitetsregister. About the National Quality Registries [Internet]. Nationella Kvalitetsregister.Sveriges kommuner och regioner.; 2019 [cited 2020 Jul 22]. Available from: http://kvalitetsregister.se/englishpages/aboutqualityregistries.2422.html

50. Emilsson L, Lindahl B, Köster M, Lambe M, Ludvigsson JF. Review of 103 Swedish Healthcare Quality Registries. J Intern Med. 2015;277:94–136.

51. Socialstyrelsen. Täckningsgrader för Nationella kvalitetsregister (Coverage rates for National Quality Registers) [Internet]. 2020. Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/statistik/2020-12-7049.pdf

52. Saklad M. Grading of patients for surgical procedures. Anesthesiology. 1941;2:281–4.

53. Sveriges Kommuner Och Landsting. Nationella Kvalitetsregister [Internet]. [cited 2019 Apr 23]. Available from: http://kvalitetsregister.se/hittaregister.1815.html

54. Kärrholm J, Lindahl H, Malchau H, Mohaddes M, Nemes S, Rogmark C, et al. The Swedish Hip Arthroplasty Register: Annual Report 2016 For Year 2016 [Internet]. Swedish Hip Arthroplasty Register; 2016. Available from: https://registercentrum.blob.core.windows.net/shpr/r/Annual-Report-2016-B1eWEH-mHM.pdf

55. American Society of Anesthesiologists. ASA Physical Status Classification System [Internet]. 2020. Available from: https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system

56. Greiner W, Claes C, Busschbach JJV, von der Schulenburg J-MG. Validating the EQ-5D with time trade off for the German population. Eur J Health Econ. 2005;6:124–30.

57. Wittrup-Jensen KU, Lauridsen J, Gudex C, Pedersen KM. Generation of a Danish TTO value set for EQ-5D health states. Scand J Public Health. 2009;37:459–66.

58. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35:1095–108.

59. Devlin NJ, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. Health Econ. 2010;19:886–905.

60. Devlin N, Parkin D, Janssen B. Analysis of EQ-5D Profiles. Methods for Analysing and Reporting EQ-5D Data. Springer; 2020. p. 23–49.

61. Philippas D. Analysis of Covariance (ANCOVA). In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research [Internet]. Dordrecht: Springer Netherlands; 2014 [cited 2021 Jan 9]. p. 157–61. Available from: https://doi.org/10.1007/978-94-007-0753-5_82

62. Park S, Lake ET. Multilevel Modeling of a Clustered Continuous Outcome. Nurs Res. 2005;54:406–13.

63. Buonaccorsi JP, Laake P, Veierød MB. On the power of the Cochran-Armitage test for trend in the presence of misclassification. Stat Methods Med Res. 2014;23:218–43.

64. Agresti A. Categorical Data Analysis. 3 edition. Hoboken, NJ: Wiley; 2012.

65. Gust L, D'journo XB. The use of correlation functions in thoracic surgery research. J Thorac Dis. 2015;7:E11–5.

66. Kim H-Y. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod. 2017;42:152–5.

67. Kirkwood BR, Sterne JAC. Means, standard deviations and standard errors. Essential Medical Statistics. 2nd ed. Blackwell; 2003.

68. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol [Internet]. 2013 [cited 2021 Jan 10];4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3840331/

69. Fan C, Zhang D, Zhang C-H. On Sample Size of the Kruskal–Wallis Test with Application to a Mouse Peritoneal Cavity Study. Biometrics. 2011;67:213–24.

70. Hayes AF, Cai L. Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. Behavior Research Methods. 2007;39:709–22.

71. Sammut C, Webb GI, editors. Mean Absolute Error. Encyclopedia of Machine Learning [Internet]. Boston, MA: Springer US; 2010 [cited 2021 Jan 10]. p. 652–652. Available from: https://doi.org/10.1007/978-0-387-30164-8_525

72. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research. Inter-Research Science Center; 2005;30:79–82.

73. World Medical Association. WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects [Internet]. 2018. Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

74. Ludvigsson JF, Håberg SE, Knudsen GP, Lafolie P, Zoega H, Sarkkola C, et al. Ethical aspects of registry-based research in the Nordic countries. Clin Epidemiol. 2015;7:491–508.

75. Socialdepartementet. Patientdatalag (2008:355) (Patient Data Act) [Internet]. Sveriges Riksdag. 2008 [cited 2021 Jan 18]. Available from: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/patientdatalag-2008355_sfs-2008-355

76. Melancia JL, Francisco AF, Antunes JL. Spinal stenosis. Handb Clin Neurol. 2014;119:541–9.

77. Humphreys SC, Eck JC. Clinical Evaluation and Treatment Options for Herniated Lumbar Disc. AFP. 1999;59:575.

78. Lespasio MJ, Sultan AA, Piuzzi NS, Khlopas A, Husni ME, Muschler GF, et al. Hip Osteoarthritis: A Primer. Perm J. 2018;22.

79. Lespasio MJ, Piuzzi NS, Husni ME, Muschler GF, Guarino A, Mont MA. Knee Osteoarthritis: A Primer. Perm J. 2017;21.

80. Jonsson Å, Orwelius L, Dahlstrom U, Kristenson M. Evaluation of the usefulness of EQ-5D as a patient-reported outcome measure using the Paretian classification of health change among patients with chronic heart failure. Journal of Patient-Reported Outcomes. 2020;4:50.

81. Devlin N, Herdman M, Pavesi M, Phung D, Naidoo S, Beer TM, et al. Health-related quality of life effects of enzalutamide in patients with metastatic castration-resistant prostate cancer: an in-depth post hoc analysis of EQ-5D data from the PREVAIL trial. Health Qual Life Outcomes. 2017;15.

82. Pavesi M, Devlin N, Hakimi Z, Nazir J, Herdman M, Hoyle C, et al. Understanding the effects on HR-QoL of treatment for overactive bladder: a detailed analysis of EQ-5D clinical trial data for mirabegron. J Med Econ. 2013;16:866–76.

83. McClure NS, Sayah FA, Ohinmaa A, Johnson JA. Minimally Important Difference of the EQ-5D-5L Index Score in Adults with Type 2 Diabetes. Value Health. 2018;21:1090–7.

84. Hakimi Z, Herdman M, Pavesi M, Devlin N, Nazir J, Hoyle C, et al. Using EQ-5D-3L and OAB-5D to assess changes in the health-related quality of life of men with lower urinary tract symptoms associated with benign prostatic hyperplasia. Qual Life Res. 2017;26:1187–95.

85. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. Qual Life Res. 2014;23:977–89.

86. Whynes DK. Does the correspondence between EQ-5D health state description and VAS score vary by medical condition? Health Qual Life Outcomes. 2013;11:155.

87. Heijink R, Reitmeir P, Leidl R. International comparison of experience-based health state values at the population level. Health Qual Life Outcomes. 2017;15.

88. Schrag A, Selai C, Jahanshahi M, Quinn NP. The EQ-5D—a generic quality of life measure—is a useful instrument to measure quality of life in patients with Parkinson's disease. Journal of Neurology, Neurosurgery & Psychiatry. BMJ Publishing Group Ltd; 2000;69:67–73.

89. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis & Rheumatism. 1991;34:505–14.

90. Chen L-H, Liang J, Chen M-C, Wu C-C, Cheng H-S, Wang H-H, et al. The relationship between preoperative American Society of Anesthesiologists Physical Status Classification scores and functional recovery following hip-fracture surgery. BMC Musculoskelet Disord. 2017;18:410.

91. Sprague S, Bhandari M, Heetveld MJ, Liew S, Scott T, Bzovsky S, et al. Factors associated with health-related quality of life, hip function, and health utility after operative management of femoral neck fractures. Bone Joint J. 2018;100-B:361–9.

92. Payne G, Payne J. Longitudinal and Cross-sectional Studies. Key Concepts in Social Research. SAGE; 2004. p. 143–7.

93. Liu L, Li S, Wang M, Chen G. Comparison of EQ-5D-5L health state utilities using four country-specific tariffs on a breast cancer patient sample in mainland China. Patient Prefer Adherence. 2017;11:1049–56.