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

MAKING USE OF PATIENT-REPORTED OUTCOME MEASURES IN HEALTH CARE: THE CASE OF EQ-5D IN THE SWEDISH NATIONAL QUALITY REGISTRIES

Olivia Ernstsson

Stockholm 2021

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-196-1
Cover illustration: Lena Hellström Sparring, Din Tecknande Sparringpartner.

Making use of patient-reported outcome measures in health care: The case of EQ-5D in the Swedish national quality registries

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Olivia Ernstsson

The thesis will be defended in public at Karolinska Institutet, Samuelssonsalen,
Friday 18th of June 2021, at 9.00.

Principal Supervisor:
Associate professor Niklas Zethraeus
Karolinska Institutet
Department of Learning, Informatics,
Management and Ethics
Division of Medical Management Centre

Co-supervisor:
PhD Emelie Heintz
Karolinska Institutet
Department of Learning, Informatics,
Management and Ethics
Division of Medical Management Centre

Opponent:
Professor Job van Exel
Erasmus University Rotterdam
Erasmus School of Health Policy & Management
Department of Health Economics

Examination Board:
Professor Lena Wettergren
Uppsala University
Department of Public Health and Caring Sciences

Associate professor Johan Thor
Jönköping University
School of Health and Welfare
Jönköping Academy for Improvement of Health
and Welfare

Professor Mikael Svensson
University of Gothenburg
Institute of Medicine
School of Public Health and Community Medicine

POPULAR SCIENCE SUMMARY OF THE THESIS

Many of us have had the personal experience of responding to health questionnaires during health care visits. Such health questionnaires, called patient-reported outcome measures (PROMs), are being used increasingly in health care to capture the health status, problems, or symptoms experienced by patients. This information is meant to improve the delivery of health care to you, the individual patient, for example by informing the choice of treatment or evaluating the effects of treatments. Ideally, the information you provide when responding to PROMs in routine health care can be used to improve health and health care, for both yourself and others.

The aim of this thesis was to increase our understanding of the use of PROMs in routine health care. Though there are many potential benefits to using routinely collected PROMs, research is needed to better understand certain aspects of the methods used for measuring health and the ways in which data can be used to improve health and health care. The aim was explored by studying a specific case – that of EQ-5D in the Swedish national quality registries (NQRs). In the NQRs, health-related information is recorded for specific patient populations in routine health care. Most NQRs include PROMs data, and EQ-5D is the most frequently used measure. It is a short questionnaire consisting of five questions (regarding mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a scale from 0–100, on which respondents are asked to rate their current health. As a generic PROM, EQ-5D can be used to study health outcomes across different patient populations.

The thesis included four research studies. The first study provided an overview of how EQ-5D data were collected in 41 NQRs, and how the results were made available and used. The EQ-5D data collected in routine health care settings were used to assess outcomes after interventions, in quality indicators, for comparing outcomes across different health care providers, for quality improvement, and in consultations with patients. In the second study, thoughts expressed by patients with type 1 diabetes who participated in interviews revealed that different health aspects, time perspectives, and reference points were considered when they reported and valued their own health through different questionnaires, including EQ-5D. Data from two different NQRs were used in the third and fourth study, respectively. The third study, in which statistical methods were used to evaluate the ability of the EQ-5D instrument to capture aspects of health as intended, showed that EQ-5D performed well in measuring the self-reported health of patients with amputations of a lower limb. In the fourth study, the analyses of EQ-5D data showed that the health of patients who were treated with electroconvulsive therapy for major depression generally improved after treatment, and that there were no apparent differences between subgroups who received different pulse widths.

This thesis, based on both methodological and applied research, contributes to our understanding of the ways in which EQ-5D data routinely collected in health care may be used for different purposes, with the ultimate intention of improving the care and health of patients.

ABSTRACT

Introduction: Patient-reported outcome measures (PROMs) have been widely used in clinical trials and research, for example to monitor the health of specific populations or to evaluate treatment effects. In the Swedish national quality registries (NQRs), structured individual-level data for specific patient populations are collected in routine health care settings. In addition to disease-related information, most NQRs include PROMs data. The most common PROM in the NQRs is EQ-5D, which is a generic instrument that can be used for measuring and valuing health across different patient populations and disease areas. The EQ-5D questionnaire includes questions addressing five dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a visual analogue scale (EQ VAS). In addition, EQ-5D provides an indirect method for obtaining health state values and is therefore widely used for the purpose of economic evaluation in health care.

There has been increasing interest in the use of PROMs in routine health care. Ideally, routinely collected PROMs data could be used to inform decisions to improve the quality of care and, ultimately, the health of the population. In addition to the potential benefits in patient-clinician encounters, aggregate-level PROMs data could for example be used for assessing the effectiveness of treatments, for detecting variations between health care providers or regions, or as input for decision-making. Real-world cases in which PROMs have been widely implemented in routine health care are needed to better understand the actual use of routinely collected PROMs data.

Aim: The overall aim of the thesis was to increase knowledge on the use of PROMs collected in routine health care. Using the case of EQ-5D in the Swedish NQRs, this thesis addressed the overall aim by investigating current practices of routine collection and use of EQ-5D data, and by exploring measurement properties, thoughts behind patients’ responses, and applications of EQ-5D as an outcome measure.

Methods: All four studies included in this thesis examined aspects of making use of EQ-5D data, but differed in their designs, samples, data, and analyses. Study I provided an overview of how EQ-5D data were collected in the Swedish NQRs, and how the collected EQ-5D data were made available and used. Information for each registry was obtained from webpages and through personal communication with representatives from the registries. In Study II, twenty patients with type 1 diabetes participated in qualitative individual interviews. Participants were asked to describe their thoughts out loud while reporting and valuing their own health, using EQ-5D-5L, EQ VAS, and time trade-off (TTO). The interviews were analyzed using qualitative thematic analysis.

Study III and Study IV were based on data obtained from two NQRs: the Swedish Amputation and Prosthetics Registry and the Swedish National Quality Registry for ECT. In Study III, patients responded to either EQ-5D-3L or EQ-5D-5L six months after a major lower limb amputation (LLA). The assessment and comparison of the measurement properties for the two EQ-5D versions included analyses of feasibility, response patterns, informativity, and convergent and known-group validity. In Study IV, multiple linear regression analyses were used to examine the association between pulse width and health-related quality of life (HRQoL) after electroconvulsive therapy (ECT) in patients with unipolar or bipolar major depression.

Findings: In Study I, the overview of current practices of the collection, presentation, and use of EQ-5D data in the NQRs showed that EQ-5D was administered across various populations and interventions, and often in combination with disease-specific measures. EQ-5D data were most frequently collected in registries targeting patients with conditions related to the musculoskeletal system. In Study III, the assessment of measurement properties indicated that there were advantages to using EQ-5D-5L over EQ-5D-3L in patients with a major LLA, mainly due to improved informativity and validity of the descriptive system.

In Study II, thoughts expressed during the qualitative interviews revealed some variation in the aspects considered and the time perspectives and reference points used when assessing one's own current health using the EQ-5D-5L, EQ VAS, and TTO. Some participants expressed a desire to discuss their responses with the health care provider.

The overview showed several examples of ways in which routinely collected EQ-5D data were used, including assessment of interventions, health economic studies, benchmarking, in quality indicators, quality improvement, and in patient-clinician encounters. Still, 19 of the 41 NQRs reported that they were unaware whether the collected EQ-5D data were used for follow-up, quality improvement, or decision-making. In Study IV, which was an example of the use of EQ-5D data for assessing real-world outcomes associated with an intervention, the results showed no robust associations between pulse width and HRQoL after ECT. Nevertheless, relatively large mean improvements in HRQoL were observed for patients treated with ECT for unipolar or bipolar depression, regardless of the pulse width received.

Conclusions: The findings from this thesis contribute to the understanding of the use of PROMs data routinely collected in the Swedish NQRs, of considerations in the choice of EQ-5D version in specific populations, of ways in which EQ-5D data can be used to assess specific interventions, and of the interpretations and thoughts behind patient's responses when reporting and valuing their own health. The case of EQ-5D in the Swedish NQRs revealed several opportunities of making use of routinely collected PROMs data with the ultimate intention to improve health care and patient health. Still, there is potential to increase the use of EQ-5D data for follow-up, quality improvement, or decision-making. For future implementations of PROMs in routine health care, it may be useful to further explore how routinely collected PROMs data could be of most use to patients and other stakeholders, and to further explore prerequisites for making use of routinely collected PROMs data at different levels of the health care system.

Key words: patient-reported outcome measures, health-related quality of life, registries, routinely collected health data, EQ-5D, time trade-off

LIST OF SCIENTIFIC PAPERS

I. **Ernstsson O**, Janssen MF, Heintz E. Collection and use of EQ-5D for follow-up, decision-making, and quality improvement in health care – the case of the Swedish National Quality Registries. J Patient Rep Outcomes. 2020;16(4):78.

II. **Ernstsson O**, Burström K, Heintz E, Mølsted Alvesson H. Reporting and valuing one's own health: a think aloud study using EQ-5D-5L, EQ VAS and a time-trade off question among patients with a chronic condition. Health Qual Life Outcomes. 2020;18(1):388.

III. **Ernstsson O**, Hagberg K, Janssen MF, Bonsel G, Korkmaz S, Zethraeus N, Heintz E. Health-related quality of life in patients with lower limb amputation – An assessment of the measurement properties of EQ-5D-3L and EQ-5D-5L using data from the Swedish Amputation and Prosthetics Registry. (Manuscript).

IV. **Ernstsson O**, Heintz E, Nordenskjöld A, Johnson JA, Korkmaz S, Zethraeus N. Association between pulse width and health-related quality of life after electroconvulsive therapy in patients with unipolar or bipolar depression: an observational register-based study. (Manuscript).

RELATED PUBLICATIONS

i. **Ernstsson O**, Tinghög P, Alexanderson K, Hillert J, Burström K. The External Validity of Mapping MSIS-29 on EQ-5D among individuals with Multiple Sclerosis in Sweden. MDM Policy Pract. 2017;2(1): 2381468317692806.

ii. McKay KA, **Ernstsson O**, Manouchehrinia A, Olsson T, Hillert J. Determinants of quality of life in pediatric- and adult-onset multiple sclerosis. Neurology. 2020;94(9):e932-e941.

CONTENTS

1 INTRODUCTION....1
2 BACKGROUND....3
  2.1 Patient-reported outcome measures....3
    2.1.1 Use of patient-reported outcome measures in routine health care....4
    2.1.2 Methodological aspects of measuring and valuing health....7
  2.2 EQ-5D – a generic patient-reported outcome measure....9
    2.2.1 Measurement properties of EQ-5D....10
  2.3 The case of EQ-5D in the Swedish national quality registries....11
    2.3.1 Rationale....13
3 RESEARCH AIMS....15
  3.1 Overall aim of the thesis....15
  3.2 Aims of the specific studies....15
4 MATERIALS AND METHODS....17
  4.1 Settings and data sources....17
  4.2 Primary measures....18
  4.3 Overview of the studies....22
  4.4 Materials and methods used in the specific studies....23
    4.4.1 Study I....23
    4.4.2 Study II....24
    4.4.3 Study III....26
    4.4.4 Study IV....28
  4.5 Ethical considerations....30
    4.5.1 Ethical review....30
    4.5.2 Benefits and risks....30
5 RESULTS....33
  5.1 Study I – Overview of EQ-5D in the Swedish national quality registries....35
  5.2 Study II – Thoughts when reporting and valuing health....36
  5.3 Study III – Measurement properties of EQ-5D....38
  5.4 Study IV – Using EQ-5D to assess outcomes after treatment....40
    5.4.1 Results of additional analyses....41
6 DISCUSSION....43
  6.1 The case of EQ-5D in the Swedish national quality registries....43
  6.2 Prerequisities for making use of data: measurement and interpretability of results....44
  6.3 Implications for policy, practice, and research....46
  6.4 Methodological considerations....48
    6.4.1 The case of EQ-5D in the Swedish national quality registries....48
    6.4.2 Primary data collection: Study II....48
    6.4.3 Secondary data collection: Study III and Study IV....49
7 CONCLUSIONS....51
8 ACKNOWLEDGEMENTS....53

9 REFERENCES.....57
10 APPENDICES.....63
  10.1 Appendix A – Study I: Template for data collection.....63
  10.2 Appendix B – Study II: Interview guide.....65
  10.3 Appendix C – Study IV: Results of additional analyses.....66

LIST OF ABBREVIATIONS

ECT	Electroconvulsive therapy
HRQoL	Health-related quality of life
ICD	International Classification for Diseases
IPWRA	Inverse-probability-weighted regression adjustment
LISA	Longitudinal integrated database for health insurance and labour market studies
LLA	Lower limb amputation
NQR	National quality registry
PCHC	Paretian Classification of Health Change
PREM	Patient-reported experience measure
PRO	Patient-reported outcome
PROM	Patient-reported outcome measure
QALY	Quality-adjusted life year
Q-ECT	Swedish National Quality Registry for ECT
SALAR	Swedish Association of Local Authorities and Regions
SwedeAmp	Swedish Amputation and Prosthetics Registry
TTO	Time trade-off
VAS	Visual analogue scale

1 INTRODUCTION

Patient-reported outcome measures (PROMs) provide standardized measures of health, as assessed directly from a patient’s own perspective. Traditionally, PROMs have been integrated as outcome measures in clinical research for assessing health outcomes or treatment effects. There is growing interest in the implementation of PROMs in routine health care, where they would provide several opportunities of capturing real-world health outcomes and effectiveness, for assessing and developing the quality of health care.

In Sweden, there are more than 100 national quality registries (NQRs) that contain individual-level data collected from routine health care settings. The registries may be used for monitoring and improving the quality of care for specific target populations, such as patients with a certain diagnosis or patients undergoing certain interventions. The registries typically include data on both health care processes and outcomes. The majority of the registries include one or several disease-specific or generic PROMs, with EQ-5D being the most frequently used measure. However, routine collection of PROMs in health care is unlikely to contribute to improvements in health care on its own. So far, there is limited knowledge regarding the extent to which the routinely collected EQ-5D data are being used for purposes that intend to ultimately contribute to improved health care or health outcomes.

In this thesis, the use of EQ-5D in the Swedish NQRs was applied as a case to better understand the use of PROMs collected in routine health care. Specifically, the thesis encompassed four research studies which have examined different applications and prerequisites for making use of routinely collected EQ-5D data. Hereafter, the four studies will be referred to by their roman numerals (Studies I–IV).

In Study I, we examined current practices of how EQ-5D data were being collected in the Swedish NQRs, and how the collected EQ-5D data were made available and used for follow-up, quality improvement, or decision-making.

In Studies II and III, some essential prerequisites for making use of data were examined. In Study II, qualitative think-aloud interviews were conducted to better understand how patients think and reason when reporting and valuing their own health. In Study III, we examined and compared the measurement properties of EQ-5D-3L and EQ-5D-5L in patients with a major lower limb amputation.

In Study IV, routinely collected EQ-5D data were used to examine the association between a certain treatment parameter and health-related quality of life in patients who were treated with electroconvulsive therapy for unipolar or bipolar major depression.

2 BACKGROUND

The goal of a health care system is ultimately to improve the health of the population (1). The health care system, which in itself encompasses various systems and processes, may be separated into different levels, i.e., the micro-, meso- and macro-levels (2). The micro-level of the health care system concerns aspects directly related to the provision of health care to individual patients, including the patient-clinician encounter. The meso-level represents the part of the health care system addressing management within health care units, hospitals, or regions. Lastly, the macro-level cover authorities and national structures for cooperation and health policy. Given the fact that the resources available for health care are scarce (3), it is essential to prioritize activities that lead to improved health. To assess the performance of health care, a central question concerns the extent to which the health care system operates to maximize the health of the population (4, 5). Thus, the monitoring of health outcomes plays an important role at several levels of the health care system, i.e. for assessing the health of individual patients and for assessing the performance and quality of health care.

Since the health care system should operate to improve the health of individuals, it could be argued that the patient perspective on health and health care is central (4). Most processes in health care are carried out to achieve outcomes that are important for patients, such as survival, reducing symptoms or disability, and improving health-related quality of life (HRQoL) (6). Moreover, while certain measures are clinically relevant for understanding the status of a condition or disease, measures of patient-reported health status or HRQoL may contribute with a perspective on the extent to which a patient is affected by their disease (7).

Before moving on to the methods for measuring health and HRQoL, the definitions of these concepts should be addressed. One widely used definition of health, as specified in the Constitution of the World Health Organization (1948), is that “health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” (8). By contrast, there is no well-established definition of HRQoL. Still, investigations of the application of measures of HRQoL seems to reveal an agreement that HRQoL is a multidimensional concept (9), and thus several health-related aspects can influence a person’s HRQoL. In this thesis, the definition described by Fayers & Machin (2016) is used, namely that HRQoL can be seen as “the set of outcomes that contribute to a patient’s well-being or overall health, or a summary measure or scale that purports to describe a patient’s overall well-being or health” (p.4) (9).

2.1 PATIENT-REPORTED OUTCOME MEASURES

Among the many different health outcomes measures available, some are specifically constructed to capture health status as assessed from the patient perspective. A patient-reported outcome (PRO) can be defined as “any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else” (10). Standardized measures, PROMs, may be constructed to measure various aspects of health and HRQoL, such as physical or social functioning, symptoms, or abilities (9). In this thesis, *patient-reported outcome measures* (PROMs) and *patient-reported outcomes* (PROs) are used as separate concepts. While the concept of a PROM refers to any instrument developed to measure certain health dimensions or constructs, PRO refers to a broader concept entailing a variety of health-related outcomes reported from the patient perspective. The terms *instruments* and *measures* are used interchangeably.

There are a wide range of different PROMs available, developed to capture aspects relevant for different target populations and to be used for different purposes. Condition- or disease-specific measures are intended to capture aspects relevant for patients with a specific condition or disease, while generic measures should be broadly applicable and facilitate comparison across populations (11). Depending on the purpose and scope of measurements performed, several measures – both disease-specific and generic – may be used in combination (12). One example of a commonly used generic PROM is EQ-5D, which is a short questionnaire consisting of questions addressing five health dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) and a visual analogue scale (hereafter referred to as EQ VAS) (13, 14).

2.1.1 Use of patient-reported outcome measures in routine health care

Importantly for the motivation underlying this thesis, PROMs have been described as transitioning from being predominantly used as outcome measures within clinical research to more recently being used for planning and managing care, to assure quality of care, monitor the health of the population, and to inform policy-making (5). Still, compared to the use of PROMs in clinical trials and in research, the routine use of PROMs in clinical practice remains less developed, or at least less documented in the scientific literature (6, 7). In theory, the opportunities for making use of PROMs data collected in routine health care range from the direct benefit when used in the patient-clinician encounters, to the use of PROMs data at an aggregate level for quality improvement efforts and for informing decision-making in health care. Though the potential benefits of using PROMs addresses all levels of the health care system (15), the use of routinely collected PROMs data for assessing the quality of health care is at an early stage of development (6). The following sections will, without any intention of being exhaustive, provide a brief overview of essential aspects for making use of routinely collected PROMs data in health care. These aspects cover the intended and observed effects, and facilitators of and barriers to making use of PROMs data.

When searching for literature on this topic, most previous publications addressing the potential benefits of using PROMs in routine health care are published as discussions, commentaries, or editorial papers. Still, this literature describes various potential benefits from using data, at both the individual and the aggregate level. There are several examples of use of individual-level data in patient-clinician encounters at the micro-level of the health care system. PROMs data can be used for identifying and prioritizing health problems and/or symptoms, screening, shared decision-making, assessing treatment outcomes, and facilitating communication between patients and clinicians (6, 12, 16-20). In addition, PROMs can be seen as tools for promoting patient-centered care (16) and may play an important role for prioritizing the patient over the disease in clinical practice (7).

Moreover, routinely collected PROMs data can be analyzed and used at an aggregate level for improving health care. Aggregate PROMs data may be used in prognostic models for assessing the probabilities of different treatment outcomes for certain interventions (21), which could also be integrated in patient-clinician encounters. At the meso- and macro-levels of the health care system, data may be used for quality improvement initiatives, for assessing and comparing health care providers in terms of effectiveness and quality of care, as input to health technology assessments, value-based payments, and for assessing the performance of programs or health care delivery systems (6, 16-18, 20).

2.1.1.1 Intended and observed effects from using PROMs in routine health care

Although previous sections have listed a number of potential benefits of using PROMs routinely in health care, they do not reveal the intended mechanisms and theories explaining the ways in which PROMs could improve health care and health outcomes for patients. First of all, it should be noted that routine use of PROMs is a complex concept in itself. The implementation of PROMs cannot be uniform, as the choice of instrument, the design for data collection and feedback, and the use of data need to be purposefully adapted to fit the specific health care context and target population. In addition, the implementation of PROMs may be further complicated by the fact that different uses of data are intended to achieve different goals. As previously mentioned, the goal may be to improve the health of the individual (e.g., through screening or shared decision-making) or to improve the quality of health care (e.g., by assessing and comparing provider performance).

Several publications by Greenhalgh and colleagues have used theory-driven approaches to address the mechanisms behind the use of PROMs in health care. In 2009, Greenhalgh summarized the literature on the extent to which applications of PROMs in clinical practice could impact the process and outcomes of care (16). Possible applications were summarized in regard to two dimensions: the level of aggregation of data and whether data were used in patient consultations or not. For these applications to work as intended, an active process of interpreting and acting on the results was considered necessary. For example, the use of PROMs could lead to action if the results indicate the presence of a specific problem (screening), if necessary changes can be made if a treatment is not working as desired (monitoring), if it can be used as a tool for discussing areas that are of priority for the patient, or as a tool for making the patient more involved in decisions when weighing benefits and risks of different treatment options (16).

Moreover, based on policy documents, publications, and meetings with stakeholders, Greenhalgh and colleagues (2018) developed a theoretical framework explaining the ways in which feedback of performance data and PROMs data may stimulate provider behavior to improve care and outcomes for patients (22). In summary, feedback on aggregate-level PROMs data can stimulate quality improvement initiatives by providing support for patient's choice of provider, by making providers accountable to commissioners, regulators, and the general public, and by providing data for benchmarking that enables providers to compare their performance with others. Notably, several contextual factors were considered important, such as the timeliness of providing feedback on PROMs data, and the extent to which PROMs data were interpreted in the context of other performance data (22).

It has been suggested that PROMs can be used routinely in clinical practice for both evaluating and improving processes and outcomes of health care (23). By searching for studies evaluating the actual benefits of PROMs in clinical practice, some literature reviews have summarized and assessed the effects on patient care and/or health outcomes (19, 23). In summary, controlled trials demonstrate large variations in study designs, selected PROMs, patient populations, health care settings, modes of administration, education and involvement of health care professionals, the timing, recurrence, and content of feedback, and provision of recommendations or support on how to act on the results (19, 23). Variations were also observed in the effect measures used, such as changes in provider behaviors, health status, satisfaction, and – less frequently – communication, resource use, concordance, and patient behavior (23). No firm conclusions could be drawn from the identified studies, due to the heterogeneity and limitations in the study designs (19, 23). Still, some results indicated that routine use of PROMs in clinical practice might have positive effects on certain process variables, e.g. identifying health problems and patient-provider communication (19).

A large share of the previous research has been focused on use at the individual level, while less is known about the use of aggregate-level PROMs data for quality assurance or quality improvement, or as an indicator of health care performance (20, 22, 24). One previous initiative frequently described in the literature is the NHS PROMs Program. In 2009, the National Health Service (NHS) in England implemented a program of routine collection of one generic PROM (EQ-5D) and several condition-specific PROMs, before and after four elective surgical procedures: hernia repair, hip replacement, knee replacement, and varicose veins (25, 26). The intention behind the program was to enhance provider performance by publicly presenting PROMs data for different health care providers (26). However, the program seemed to have had limited impact on provider performance and patients' decisions regarding their choice of provider (26). It has been argued that collected data remain largely unused due to them not being perceived to fit the purpose, and due to challenges in accessing, understanding, and acting on the results (27). Furthermore, some criticism relate to the difficulties of making meaningful interpretations of data, for example due to a large proportion of missing data, or insufficient provision of feedback to clinicians and patients (21, 26).

2.1.1.2 Facilitators and barriers

Related to the promising opportunities associated with the use of routinely collected PROMs, several facilitators and barriers are presented in the literature. Advancements in technology has likely been a facilitator for the collection and monitoring of PROMs data in clinical practice, through increased use of electronic records and clinical registries (6, 7). The possibility of having standardized data collection across providers facilitates use and evidence synthesis of data, for purposes such as analyzing real-world outcomes, safety, effectiveness, and evaluating the performance of providers (21, 28).

In a systematic review of qualitative studies, some examples of identified barriers to routine data collection of PROMs were the associated workload, use of questionnaires that are not user-friendly, and lack of guidelines on how and when to collect, analyze, and interpret the data (29). Other barriers include lack of integration between use of data at the individual level and the aggregate level, that patients may be unsure of the purpose of completing a PROM assessment, survey fatigue, incompatible IT systems, inadequate reporting or feedback of PROMs data, and difficulties drawing conclusions due to large proportions of missing data (30). Moreover, factors contributing to the limited influence on clinical decision-making include that data are often fed back at one time point rather than following the decision process longitudinally, and that data are made available in a format inadequate for interpretation (31).

Although the potential impact on decision-making has been mentioned as one of the largest potential benefits of routinely collected data, concerns have been raised regarding tensions resulting from differing priorities at different levels of the health care system (17, 32). One challenge of motivating routine collection only for its use at a higher health care system level is the lack of accessible feedback at the micro level, i.e., to health care providers and patients (30). Furthermore, the purpose for which data are collected at the clinical level, i.e. providing the most appropriate care for an individual patient, may differ from the priorities at higher levels of the health care system. The risk for incorrect interpretation has been pointed out when making decisions at a system level based on data that have been collected within routine health care, for example when there are competing priorities between individual care and efficiency (17). At the aggregate level, additional challenges include managing the risk of selection bias (e.g., healthier populations), adequate case-mix adjustments, and identifying the most suitable timing of measurements in order to make meaningful interpretations (32).

To overcome the barriers to making use of PROMs data, there has been an emphasis on the importance of involving and engaging multiple stakeholders, including for example patients, clinicians, researchers, regulators, and policy makers (5, 30). Stakeholder engagement may be used for strengthening collaborations and partnerships, for ensuring that the selected outcomes correspond to the needs of the stakeholders, and for developing strategies to improve the use of PROMs for decision-making (5, 30).

2.1.2 Methodological aspects of measuring and valuing health

2.1.2.1 Measurement properties

In order to make use of the information from a PROM, an important consideration is whether the selected measure can provide valid and reliable results. The measurement properties of an instrument can be assessed through analyses of validity, reliability, sensitivity, and responsiveness (9). Validity represents the ability of an instrument to measure what it is intended to measure and its ability to measure variation in that variable. Analyses of validity could for example cover assessments of convergent, divergent, or known-group validity. Reliability is assessed by investigating whether measurements are reproducible and consistent. Furthermore, the sensitivity of an instrument concerns its ability to detect actual differences between individuals or groups, and the responsiveness of an instrument examines its ability to detect changes for an individual or group over time (9). In addition to the statistical methods used for assessing measurement properties, qualitative methods are valuable in several aspects, for example when constructing and testing a questionnaire (9).

Another consideration relevant for this thesis is analyses of the information captured by an instrument. By studying the distributions of responses, the proportions of respondents reporting no problems (i.e. ceiling) or extreme problems (i.e. floor) can be examined. In addition to the proportions of patients indicating the highest or lowest response categories, there are certain analyses (e.g. Shannon's indices) that facilitate assessment of the informativity across all possible response options (33).

2.1.2.2 Measuring and valuing health for health economic evaluation

For the purpose of health economic evaluations – in which two or more alternative treatments or interventions are compared in terms of their costs and health effects (3) – certain outcome measures are preferred. The preferred outcome measure for a cost-utility analysis is the quality-adjusted life year (QALY), which combines the value of a health state (also referred to as quality weight or utility weight) and the time an individual spends in that health state. The health state values are anchored in 0 (dead) and 1 (full health) (3). Some measures enable the calculation of negative values, which are thus considered to be worse than dead.

One way of eliciting health state values is through the use of certain PROMs that provide both a health state classification system and an indirect method for assigning values to different health states. Some examples of commonly used generic preference-based instruments, each covering different aspects of health, are EQ-5D (13, 14), SF-6D (34), and HUI (35). These PROMs are commonly referred to as preference-based measures (36) or multi-attribute health utility instruments (37), and should, at least in theory, facilitate health technology assessment by providing values that are broadly comparable across interventions, treatments, and conditions (36). For this category of PROMs, patients are asked to fill out a questionnaire, based on which the health state of a patient can be classified. Values are then calculated based on previously elicited health state valuations. These sets of values, which have been estimated in specific valuation studies, are based on any of the available health state valuation methods, for example standard gamble, time trade-off (TTO), rating scale or visual analogue scales (VAS) (38). Thus, PROMs that integrate preferences for different health states differ from other measures in that they provide different weights for different dimensions of health (9).

One of the most established methods for directly eliciting health state values is TTO, which was originally developed by Torrance, Thomas & Sackett (1972) (39). In the TTO method, the respondent is required to make a choice between living in a specified health state for a certain time period (e.g. ten years) followed by death or living in full health for a shorter time period. The method originally involved that the number of years in full health was varied until a point of indifference was reached, in which the respondent considered the two alternatives to be equal (39). Since individual interviews are relatively resource-demanding, alternative solutions for health state valuations, such as open-ended survey questions, have been explored (40). However, it should be noted that different designs and elicitation procedures for TTO have been shown to influence the health state values (40, 41).

In addition to *what* to value and *how* to value health, another consideration is the perspective of *who* should value health (42). This is a normative consideration with regard to whether health should be valued by individuals in the general population or in specific patient populations, and whether the respondents should value their currently experienced health states or hypothetical health states that are described to them (43, 44). There are several arguments for each positions, including that only patients themselves are able to judge what it is like to experience a specific health state, or that societal decisions should be made by the general public (38, 44).

Internationally, there are some variations in the methods recommended for reimbursement decisions (45). For example, the National Institute for Health and Care Excellence (NICE) requests the use of QALYs as an outcome measure (46). These recommendations specify that changes in HRQoL should be directly reported by patients and values should be based on public preferences using a choice-based valuation method. In Sweden, the Dental and Pharmaceutical Benefits Agency also recommends the use of QALYs, based on either a direct or an indirect valuation method, as outcome measure for health economic evaluations (47). By contrast, these guidelines specify that values should preferably be elicited from persons who are in the specific health state. Thus, from a Swedish perspective, research on valuations of experience-based health states are of particular relevance.

In summary, differences in health state values depending on the measure used will likely occur as a result of different methodological choices. At a general level, these choices concern the aspects of health covered by the different measures and health state classification systems, the valuation methods used, and the perspective applied.

2.2 EQ-5D – A GENERIC PATIENT-REPORTED OUTCOME MEASURE

In the late 1980s, EQ-5D was developed to be a concise, generic instrument with the purpose to measure, compare and value health status across disease areas (48). The questionnaire consists of two parts: a descriptive system covering five health dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and the EQ VAS (13, 14). There are currently two versions available for adults. The original version, EQ-5D-3L, has three severity levels on each of the dimensions covered by the descriptive system and the more recently developed version, EQ-5D-5L, has five severity levels (13, 14, 49). EQ-5D-5L was developed in 2009 to improve the sensitivity of the instrument and to reduce the ceiling effects (i.e. the proportion reporting no problems) observed in the original version (48, 49). There is a crosswalk algorithm that facilitates comparisons when different EQ-5D versions have been administered over time or in different populations (50).

One of the key purposes of constructing EQ-5D was to enable health economic evaluation, i.e., to develop an instrument of health status measurement that could also inform resource allocation decisions (48). Thus, the EQ-5D instrument provides an indirect method for valuing health states. An individual's responses to the questionnaire result in a descriptive health profile, which can be assigned a health state value (EQ-5D index) through the use of a value set. EQ-5D value sets are commonly country-specific and developed to reflect the preferences of the general population, and most value sets available are based on TTO or VAS valuations (51). The first initiative for developing an EQ-5D value set was conducted in the UK (52), and the result remains one of the most established and widely used value sets internationally (48). Value sets developed more recently have often been based on standardized valuation protocols (53).

Currently, there is no EQ-5D value set developed from the standardized valuation protocol available for Sweden. Two studies have been conducted to develop experience-based value sets for EQ-5D-3L (54) and EQ-5D-5L (55). The methods used for these two value sets differ from those in the valuation protocol in several ways. In the experience-based value sets, members of the general population have valued their own current health state. A survey question was used, meaning that respondents valued their health without assistance from an interviewer. Because an open-ended TTO question was used, respondents did not go through iterations before deciding on the number of years in full health they considered to be of equal value to living ten years in their current health state. Furthermore, the open-ended TTO question did not include a valuation procedure for health states considered worse than dead, and therefore, the scale for the index values does not include negative values.

The use of EQ-5D may serve several purposes. As EQ-5D provides an indirect method for valuation of health states that facilitates the calculation of QALYs, the instrument has been widely used for the purpose of economic evaluations of health care programs or interventions (51). In the early establishment of EQ-5D, the development of the instrument coincided with the introduction and more widely adopted use of health technology assessments, which may have contributed to the demand for and popularity of collecting EQ-5D data (48). Since then, EQ-5D has been used as an outcome measure in clinical trials, observational studies, and population health surveys, and more recently, there has been growing interest in investigating its usefulness for routine outcome measurement in health care (48).

2.2.1 Measurement properties of EQ-5D

The measurement properties of EQ-5D have been evaluated in several patient populations and conditions. In a literature review (2018) examining the validity and sensitivity of some of the most commonly applied generic preference-based measures, EQ-5D was the instrument for which the largest number of studies was identified (36). The studies identified covered conditions across 16 ICD classes (i.e., International Classification of Diseases) and showed that EQ-5D performed well across several conditions, such as diabetes and several types of cancer, but showed mixed performance for cardiovascular disease and visual disorders, and poor performance in hearing impairment, multiple sclerosis, personality disorder, schizophrenia, and dementia (36). Although some conditions related to the musculoskeletal system have been addressed, no previous study has examined the measurement properties of EQ-5D in patients with amputation of a lower limb.

The newest EQ-5D version, EQ-5D-5L, was developed to address some of the limitations observed in the three-level version. A recently published literature review (2020) examining studies on the validity, reliability, and responsiveness of EQ-5D-5L identified approximately 100 studies encompassing a range of general and patient populations (37). The conditions most frequently covered by these studies were related to the musculoskeletal system and cancer (8 studies each) and respiratory conditions (7 studies). The findings demonstrated relatively large variations in the proportion of patients reporting no problems (i.e., ceiling) in all dimensions, ranging from 2% to 36%.

Moreover, the measurement properties of EQ-5D-3L and EQ-5D-5L have been compared in more than 40 studies including a variety of populations. The findings from a literature review including 24 of these studies showed that both missing values and floor effects were generally low (<5%) for both versions (56). Generally, there were indications that EQ-5D-5L performs better in several aspects, especially in terms of the distributional characteristics, e.g., ceiling effects and informativity of the descriptive system (56, 57). When comparing across conditions, the lowest ceiling effects, for both versions, were observed in patients undergoing hip and knee replacement (58, 59), orthopedic, psychosomatic, or rheumatic rehabilitation (60), and patients with acute stroke (61). Notably, although the literature has increased further in recent years, none of the identified studies comparing EQ-5D-3L and EQ-5D-5L has had a specific focus on mental health conditions, such as major depression.

Some concerns have been raised regarding the extent to which EQ-5D is suitable for capturing aspects relevant for mental health conditions (62, 63). The literature addressing the measurement properties of EQ-5D has specifically focused on conditions such as depression, anxiety, bipolar disorder, personality disorder, schizophrenia, and dementia (36, 63). In a review of the literature (search until 2011), the overall conclusion from the identified studies was that most findings indicated satisfactory validity and responsiveness of EQ-5D-3L in patients with depression. By contrast, the limited evidence on the validity of EQ-5D-3L in bipolar disorder was mixed, and there were no studies addressing the responsiveness in this patient population (63).

2.3 THE CASE OF EQ-5D IN THE SWEDISH NATIONAL QUALITY REGISTRIES

In this thesis, the use of EQ-5D in the Swedish national quality registries (NQRs) was applied as a case to better understand the use of PROMs collected in routine health care. The NQRs contain structured individual-level information and are used for monitoring and improving the quality of care for specific patient populations (2, 64). The registries specifically target patients with a certain diagnosis, patients undergoing certain interventions or treatments, or patients who are at certain risk (2). The vision is that the NQRs should be used "in an integrated an active way for continuous learning, improvement, research and management to create the best possible health and care together with the individual" (65).

There are some variations in the structure, size, function, and data registered in the NQRs, as a consequence of how health care is organized for specific patient populations (2). Typically, individual-level information registered includes basic demographics (e.g., sex, age), process measurement and disease-related information (e.g., diagnoses, treatments, complications, rehabilitation), and outcomes of care (e.g., survival, clinically relevant outcomes, disease-specific and generic PROMs). Furthermore, some NQRs contain questions regarding the satisfaction and/or experience of care, i.e., patient-reported experience measures (PREMs).

Historically, the NQRs have usually been established on the initiative of one or several health care providers and the first registry was established as early as in the 1970s (64). Currently, there are more than 100 Swedish NQRs. Each NQR has a registry organization and a steering group involved in the management and development of that registry (2). The registries receive logistic and economic support from the Swedish government and the Swedish Association of Local Authorities and Regions (SALAR) (64). Although the organization of the NQRs is decentralized, there are procedures in place for quality assurance. All NQRs send reports and applications for funding on an annual basis and receive feedback on performance from an executive committee (12).

Some aspects of the data coverage should be addressed, as they are important for the interpretation of results in this thesis. First, the quality of data entered into the NQRs can be evaluated by assessing their completeness and coverage. Completeness may be defined as the proportion of all eligible patients in the target population who are included in the registry, and coverage as the number of health care units that are affiliated with and provide data to the NQR (64). Efforts to ensure high data quality involve automated checks for preventing incorrect data entries, reviewing outliers, and comparing data with the population registries and medical charts (64). Second, although the NQRs cover a wide range of conditions and interventions, it should be noted that most NQRs are focused on specialist care and that there are several health conditions for which health care is not followed through an NQR (64). Another consideration is that registries are focused on certain conditions or interventions, meaning that data on the provision of health care to patients with multiple conditions may be recorded in several registries.

In 2015, Nilsson and colleagues reviewed the inclusion and use of PROMs in the Swedish NQRs. Almost 90% of the registries included at least one PROM or PREM (12). Disease-specific PROMs were most common (more than 60% of the NQRs), followed by generic PROMs (more than 45% of the NQRs). Based on a review of generic measures included in the registries, the most frequently used instrument was EQ-5D (35 registers), followed by

SF-36/RAND-36 (10 registers). From this review, it became clear that the design for data collection and use of PROMs data varied substantially between NQRs. Some NQRs administered PROMs selectively at specific hospitals or clinics, while other NQRs had nationwide collection of PROMs. Furthermore, one of the key messages from this review was that data analyses and presentation of PROMs data were still at a basic level (12).

2.3.1 Rationale

There are several motives for studying the case of EQ-5D in the Swedish NQRs. Through structured nationwide data collection, the NQRs provide opportunities to study real-world practices involving use of PROMs collected in routine health care. As a result of the way in which the NQRs are structured and organized, PROMs data from the registries could be used for several, if not all, the previously mentioned purposes. In addition to discussing the results in patient-clinician encounters, the EQ-5D data may be used for different purposes at the meso- and macro-level of the health care system as well, e.g., for quality improvement, evaluation of real-world outcomes, health economics, and decision making.

Moreover, among the relatively large number of PROMs and other outcomes included in the NQRs, EQ-5D is the PROM most widely adopted across different patient populations and interventions. Several NQRs have administered EQ-5D, commonly in combination with other disease-specific measures, in for several years, which facilitates analyses of the extent to which data have been used for different purposes. Another benefit is the level of detail of clinically relevant information reported at the same time point, such as information on diagnoses and treatments, intermediate measures, and disease-specific PROs. Ideally, the nationwide registration should contribute to the registry population being representative for the patient population, which would enable more meaningful interpretations of PROMs data.

The case and the studies included in this thesis cover several aspects important for the understanding of the use of PROMs in routine health care. Study I provided an overview of current practices of collection, presentation, and use of EQ-5D data. Study II and Study III examined prerequisites for the implementation, interpretation, and use of EQ-5D data. These studies were conducted to explore the patient perspective of using EQ-5D-5L, EQ VAS, and TTO for reporting and valuing one’s own current health and to assess the measurement properties of EQ-5D-3L and EQ-5D-5L in a specific patient population. In Study IV, EQ-5D was applied as an outcome measure for analyzing real-world outcomes after a specific treatment. In Study III and IV, registry data were obtained from two NQRs that were established for the purposes of following and evaluating the health care and outcomes for patients receiving specific interventions: patients with lower limb amputations (LLAs) and patients treated with electroconvulsive therapy (ECT) for major depression.

3 RESEARCH AIMS

3.1 OVERALL AIM OF THE THESIS

The overall aim of the thesis was to increase knowledge on the use of patient-reported outcome measures (PROMs) collected in routine health care. Using the case of EQ-5D in the Swedish national quality registries (NQRs), this thesis addressed the overall aim by investigating current practices of routine collection and use of EQ-5D data, and by exploring measurement properties, thoughts behind patients' responses, and applications of EQ-5D as an outcome measure.

3.2 AIMS OF THE SPECIFIC STUDIES

Specific aims and/or objectives were formulated for each of the studies included in the thesis. The following aims were formulated for the specific studies (I–IV):

I. To increase knowledge on how EQ-5D data are collected within the Swedish NQRs, and how the data are made available and are being used in the Swedish health care system.
II. To increase knowledge on how individuals think and reason when reporting and valuing their own current health using the EQ-5D-5L descriptive system, EQ VAS, and an open-ended TTO question.
III. To assess the measurement properties of EQ-5D-3L and EQ-5D-5L in patients with a major lower limb amputation by comparatively examining the instruments in terms of feasibility, distributional characteristics, and validity.
IV. To examine the association between pulse width and health-related quality of life measured within one week after electroconvulsive therapy and at six-month follow-up in patients with unipolar or bipolar major depression.

4 MATERIALS AND METHODS

4.1 SETTINGS AND DATA SOURCES

In this thesis, the case concerned the Swedish NQRs in which EQ-5D data and other PROMs have been collected in Swedish routine health care settings. In Sweden, the 21 regions have the main responsibility for funding and providing primary and specialist health care to the population (66). Health care is primarily funded through taxes, and partly from patient fees and other private expenditure (66). There are numerous registries in Sweden that records data for the entire population, including registries for monitoring certain aspects of health care. However, these mandatory population registries rarely include detailed disease-specific data or PROs (12, 64). By contrast, the NQRs contain information on processes and outcomes for various patient populations.

Swedish NQRs. Three of the studies related directly to the Swedish NQRs, either in general or for specific registries among them. While the Swedish NQRs should encompass data from multiple health care providers (2), most registries rely primarily on data reported from specialized health care (64). Registration in the NQRs is voluntary, and there is some variation in the ways in which data are recorded in the registries (e.g., when, what, how, and by whom). As previously mentioned, there are variations in the contents and structures of the NQRs, as a consequence of the way in which health care is organized for certain patient populations and medical areas (2). Further, the coverage and completeness of data may vary both within registries (e.g., regional differences) and between registries.

In addition to the registries included in the overview of the collection and use of EQ-5D data in Swedish NQRs, two of the studies included in the thesis used data obtained from two specific NQRs – the Swedish Amputation and Prosthetics Registry (SwedeAmp) and the Swedish National Quality Register for ECT (Q-ECT).

SwedeAmp was established as an NQR in 2011 (67, 68). In SwedeAmp, patient health data are collected to enable assessment of LLAs and their consequences along the health care trajectory up to 24 months after surgery. The care processes associated with LLA involve several different public and private health care providers (68). With the intention to involve all key health care professions, SwedeAmp includes data recorded by surgeons, certified prosthetists and orthotists, rehabilitation therapists, and physiotherapists (67). Several PROs are used to examine patients’ situation, mobility, function, and HRQoL at 6, 12, and 24 months after amputation (67). At these three time points, EQ-5D is administered using a paper questionnaire either during a health care visit or in a telephone interview.

At the end of 2018, the registry included data concerning approximately 5,800 patients, 7,800 amputation surgeries, and 2,000 follow-up registrations (67). In 2019, data were reported to the registry from 13 of the 21 regions in Sweden (68). Descriptive statistics regarding the characteristics of the patient population showed that approximately 60% were male, the mean age at amputation was 74 years, and 89% of the registrations were for unilateral amputations (67).

Q-ECT, which was established as an NQR in 2011, is used for monitoring compliance to the Swedish clinical guidelines for ECT, for quality assurance and for research purposes (69, 70). Data from all 21 regions, for approximately 3,600 patients, were recorded in the Q-ECT in 2019 (70). Several PROMs are used to assess depressive symptoms, remission status, memory loss, and HRQoL. Since 2016, the registry administers follow-up assessments six months after ECT (70). EQ-5D-3L is administered using paper questionnaires at the following time points: before ECT, within one week after ECT, and six months after ECT.

In 2019, all 47 treating units in Sweden reported data to the Q-ECT (70). The completeness of patients included in the registry has improved over time, from 79% in 2012 to over 90% since 2014. In 2019, it was estimated that the registry included 93% of all eligible patients (70). Data from the registry reveal current treatment practices and use of ECT in Sweden. The predominant indications for ECT are major depression and affective disorders (69). In 2019, 60% of the patients were women and the mean age at treatment was 53 years (70).

Individual interviews. The interview study was conducted to examine the patient perspective of reporting and valuing one's own health and was thus not directly related to a specific NQR. In the interview study, participants were recruited from the Center for Diabetes (Stockholm, Sweden), which is a specialist clinic for adults with type 1 diabetes or type 2 diabetes that is difficult to treat. Patients with type 1 diabetes typically visit the Center for Diabetes on an annual basis and care is provided by a team with members from several professions, such as physicians, diabetes specialist nurses, dietitians, physical therapists, and podiatrists. Importantly for the case used in this thesis, the National Quality Registry for Diabetes does not administer EQ-5D to patients included in the registry.

4.2 PRIMARY MEASURES

EQ-5D. EQ-5D represents a set of instruments that may be used to describe and value health (71). The respondent is asked to self-report his/her current health in the two parts of EQ-5D: a descriptive system including five items which cover five health dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a vertical VAS (EQ VAS) with the end points of 0 ("the worst health state you can imagine") and 100 ("the best health state you can imagine") (13, 14). The EQ-5D descriptive system was conceptualized to measure deviations from full health (49). This is reflected by the response options, which correspond to different severity levels of problems experienced.

Currently, there are two EQ-5D versions available for adult populations: EQ-5D-3L and EQ-5D-5L (Table 1) (13, 14, 49). The versions cover the same five dimensions but differ in the number of response options. EQ-5D-3L has three severity levels for each dimension (no, some, extreme problems/unable) and can theoretically identify 243 health states (3^5). EQ-5D-5L has five severity levels (no, slight, moderate, severe, extreme/unable) and can thus identify 3,125 health states (5^5). In the development of EQ-5D-5L, some adjustments were made to the wording of the descriptive system, e.g., the response option representing extreme problems in the mobility dimension was changed from “confined to bed” in EQ-5D-3L to “unable to walk about” in EQ-5D-5L (49).

Table 1. Overview of the dimensions and severity levels covered by the EQ-5D-3L and EQ-5D-5L descriptive systems.

	EQ-5D-3L		**EQ-5D-5L**	
Dimension	Level	Description	Level	Description
MOBILITY	1	No problems in walking about	1	No problems in walking about
			2	Slight problems in walking about
	2	Some problems in walking about	3	Moderate problems in walking about
			4	Severe problems in walking about
	3	Confined to bed	5	Unable to walk about
SELF-CARE	1	No problems with self-care	1	No problems washing or dressing myself
			2	Slight problems washing or dressing myself
	2	Some problems washing or dressing myself	3	Moderate problems washing or dressing myself
			4	Severe problems washing or dressing myself
	3	Unable to wash or dress myself	5	Unable to wash or dress myself
USUAL ACTIVITIES (e.g., work, study, housework, family or leisure activities)	1	No problems with performing my usual activities	1	No problems doing my usual activities
			2	Slight problems doing my usual activities
	2	Some problems with performing my usual activities	3	Moderate problems doing my usual activities
			4	Severe problems doing my usual activities
	3	Unable to perform my usual activities	5	Unable to do my usual activities
PAIN/ DISCOMFORT	1	No pain or discomfort	1	No pain or discomfort
			2	Slight pain or discomfort
	2	Moderate pain or discomfort	3	Moderate pain or discomfort
			4	Severe pain or discomfort

	3	Extreme pain or discomfort	5	Extreme pain or discomfort
ANXIETY/ DEPRESSION	1	Not anxious or depressed	1	Not anxious or depressed
			2	Slightly anxious or depressed
	2	Moderately anxious or depressed	3	Moderately anxious or depressed
			4	Severely anxious or depressed
	3	Extremely anxious or depressed	5	Extremely anxious or depressed

For the Swedish versions of EQ-5D-3L and EQ-5D-5L, some additional adjustments of wording were made. The dimension for self-care was adjusted from “*hygien”* (“hygiene”) in EQ-5D-3L to “*personlig vård*” (“personal care/self-care”) in EQ-5D-5L, and the dimension for usual activities from “*huvudsakliga aktiviter*” (“main activities”) in EQ-5D-3L to “*vanliga aktiviteter”* (“usual activities”) in EQ-5D-5L. Furthermore, adjustments were made to level 1 for self-care, from “*Jag behöver ingen hjälp med min dagliga hygien, mat eller påklädning”* (“I do not need any help with my daily hygiene, food, or clothing”) in EQ-5D-3L to “*Jag har inga svårigheter med att tvätta eller klä mig”* (“I have no problems washing or dressing myself”) in EQ-5D-5L.

Responses to EQ-5D can be summarized in several ways. The EQ-5D health profile is a five-digit number that describes a patient’s responses on the severity level of each of the five health dimensions (72). For example, health profile 11111 represents no problems in all dimensions, while 33333 and 55555 represent extreme problems in all dimensions for EQ-5D-3L and EQ-5D-5L responses, respectively. The EQ-5D level sum score is the unweighted sum of the five digits. The possible scores range between 5 (best) and 15 (worst) for EQ-5D-3L and between 5 (best) and 25 (worst) for EQ-5D-5L (72) . Furthermore, each EQ-5D health profile can be assigned a value (sometimes referred to as a quality weight). The EQ-5D index value assigns different weights to specific dimensions and severity levels, which represents preferences for different health states (usually as stated by members of the general population) (72). In this thesis, EQ-5D index values were calculated using value sets based on TTO methods: the UK value set for EQ-5D-3L (52) and the crosswalk value set for the UK for EQ-5D-5L (50). Both value sets have a possible range from -0.594 to 1 (full health).

Another method for analyzing EQ-5D data is by presenting the Paretian Classification of Health Change (PCHC) (72). By comparing individual responses at two measurement points, the change in health can be categorized as either improved (if at least one dimension has improved and none of the others has worsened), worsened (if at least one dimension has worsened and none of the others has improved), no change (if all dimensions are unchanged), or mixed (if at least one dimension has improved and at least one dimension has worsened).

TTO. In this thesis, an open-ended TTO question was used as a direct method for valuing one’s own current health. This particular TTO version has been included in general population surveys in Sweden, and the results have been used in studies to develop experience-based value sets for EQ-5D-3L and EQ-5D-5L (54, 55).

The open-ended TTO question presented a horizontal line corresponding to a timeline of 0–10 years (i.e., each year was marked and labelled, and each half year was marked, but not labelled), and read: “Imagine that you are told that you have 10 years left to live. In connection with this, you are also told that you can choose to give up some life years to live for a shorter time period in full health. Indicate the number of years in full health that would be of equal value to 10 years in your current health state.” Below the line, there was an additional sentence that read: “(If you think that you currently have full health, you should mark 10 years.)”

4.3 OVERVIEW OF THE STUDIES

The thesis includes four separate sub-studies, in which different research design and methods were applied. An overview of the materials and methods used in the four studies is presented below in Table 2.

Table 2. Overview of the materials and methods used in the four sub-studies.

	Study I	**Study II**	**Study III**	**Study IV**
Study focus	Collection, presentation, and use of EQ-5D data in the Swedish NQRs	Thoughts when reporting and valuing one's own current health	Measurement properties of EQ-5D-3L and -5L in patients with a major lower limb amputation	The association between pulse width and HRQoL after electroconvulsive therapy
Study design	Overview of current practice	Qualitative interview study	Observational retrospective register-based study	Observational retrospective register-based study
Participants or sample	The Swedish NQRs with collection of EQ-5D data	Patients with type 1 diabetes	Patients with a major lower limb amputation	Patients treated with ECT for unipolar or bipolar depression
Data	Documents and personal communication with registry representatives	Semi-structured think-aloud interviews	National registry data from the Swedish Amputation and Prosthetics Registry	National registry data from the Swedish National Quality Registry for ECT, with linkages to other databases[1]
Primary measures	EQ-5D-3L; EQ-5D-5L	EQ-5D-5L; EQ VAS; time trade-off	EQ-5D-3L; EQ-5D-5L	EQ-5D-3L; EQ VAS
Analysis	Descriptive statistics and summary of examples	Qualitative thematic analysis	Statistical analysis of feasibility, distributional characteristics, and validity	Statistical analysis including multiple linear regression analyses

[1] Longitudinal integrated database for health insurance and labour market studies (LISA), Statistics Sweden; the National Patient Registry and the Swedish Prescribed Drug Registry, the National Board of Health and Welfare

4.4 MATERIALS AND METHODS USED IN THE SPECIFIC STUDIES

4.4.1 Study I

Design. Study I consisted of a descriptive overview based on information provided from documents and through personal communication. The purpose of extracting data was to summarize current and previous practices and examples of how EQ-5D data are collected, made available, and used at different levels of the health care system.

Sampling. All registries with a license for the use of EQ-5D at the point of data collection were included. Thirty-seven NQRs with an EQ-5D license were identified. Since the Swedish Neuro Registries had a common license for their ten sub-registries, a total of 46 registries were included in the overview.

Data collection. The data collection was guided by a template developed for the purpose of examining the research question addressed in Study I. The template included questions regarding the registry in general, the collection of PROM data, administration of PROMs, and use of PROM data (Appendix A). In addition to the information about EQ-5D specifically, the overview included information on other PROMs (information provided in the Supplementary materials to the published article).

The data collection took place between August 2018 and June 2019. First, information from documents (e.g., websites, annual reports) was added to the template for each NQR. Second, the information was either confirmed or complemented through personal communication with a representative from each NQR. The representative was usually involved in the registry organization, steering committee, or specifically in the work with PROMs. The registry representative was encouraged to provide examples for each reported use of EQ-5D data. Lastly, all representatives had the opportunity to confirm the information before publication, for instance with regard to the categorization of how EQ-5D data were collected, made available, and used for different purposes.

Outcomes and definitions. Registries were categorized into disease areas in accordance with the classification system developed and used by SALAR (65). In addition, the registries were categorized as “diagnosis registries” (i.e. including patients with a specific diagnosis) or “intervention registries” (i.e., including patients undergoing specific interventions).

The ways in which EQ-5D data were collected, presented, and used were categorized based on the template for data collection and the information provided from the registries. Target groups to whom data were made available included care givers, patients, and other decision-makers. Further, the categories reflected whether data were presented at an aggregate level (i.e., annual reports/websites; research publications; reports adapted specifically for clinics, units, or teams; reports targeting patients; other reports) or at an individual level (i.e., feedback directed at health care professionals entering data; feedback directed at patients). Use of data was categorized based on the use reported by the NQRs, including assessment of interventions, health economic studies, quality indicators, benchmarking, quality improvement, and individual consultations with patients. The category for individual patient consultations included use of EQ-5D data for screening, monitoring, decision aids, and for shared decision-making, as these were considered difficult to separate.

Certain criteria were used in the categorization of specific use of data. To be categorized as used in a quality indicator, EQ-5D data should be included in a quality indicator as defined by the registry or some other stakeholder. Further, it should represent a quantitative summary, clearly indicate good or poor quality, and be relevant for improvement (73). To be categorized as used for benchmarking purposes, the use of EQ-5D data had to involve a comparison of results between health care providers, with the purpose of performance assessment and to enable identification of improvement needs (74). To be categorized as used in quality improvement, EQ-5D data should be used actively in efforts to improve the quality of care.

Data analysis. The information reported from the registry representatives was condensed and tabulated after completion of the data collection. Results were summarized and described both in terms of descriptive statistics and by presenting examples of reported use of data for different purposes. The results regarding collection and use of EQ-5D data were presented by registry category and by registry type.

4.4.2 Study II

Design. Study II was as a qualitative interview study conducted to better understand individual’s thoughts when reporting and valuing their own current health. The interviews were guided by think-aloud interview technique and a semi-structured interview guide (75). Twenty individual interviews (28–90 min) were conducted in Stockholm, Sweden, during the time period February–July 2018.

Sampling and participants. Using the sampling strategy of purposeful sampling (76), patients with type 1 diabetes were recruited to participate in the study. It was expected that individuals who were experiencing health problems more likely would consider the trade-off between life years and health compared to individuals who were not experiencing any health problems. Sampling was performed among patients with type 1 diabetes, as diabetes is a chronic disease that impacts everyday life and may affect several dimensions of health.

The inclusion criteria were men and women between 18 and 70 years who had been diagnosed with type 1 diabetes for at least five years. In addition, for feasibility reasons, only people speaking Swedish well enough to participate in an interview were included. Study participants were recruited through an academic specialist clinic (Center for Diabetes), where patients were informed about the study during a regular health care visit. Eligibility was assessed by the four specialist nurses involved in the recruitment of participants for this study. The recruitment weeks (one week each in January, May, and June 2018) were selected based on the clinic’s schedule and the perceived need for additional interviews. Patients who expressed willingness to participate were later contacted by telephone for more information about the study and for the possibility to ask questions. Out of 77 patients who received information, 41 patients agreed to be contacted and 20 patients (13 men, 7 women) participated in the study.

Measures. The interviews were focused on three health assessments. Participants were asked to report or value their own health using the EQ-5D-5L descriptive system, EQ VAS, and an open-ended TTO question. Paper versions of the questionnaires (Swedish versions) were used.

Data collection. Three pilot interviews were conducted prior to the main data collection. The interviews followed a semi-structured interview guide (Appendix B) including introductory questions (e.g., the experience of the onset of symptoms and diagnosis, and experience of other diseases), followed by think-aloud exercises and probing questions for 1) the EQ-5D-5L descriptive system, 2) EQ VAS, and 3) the TTO question. Participants were given paper versions of the questionnaires and were asked to describe their thoughts when responding to the three assessments. At the end of their interview, each participant could add to or elaborate on their thoughts. Background characteristics for the participants were collected through a short survey filled out after the interview. No additional information regarding the participants' health status was collected. All 23 interviews (three pilot interviews and 20 interviews in the main data collection phase) were conducted by the thesis author (OE).

The concept of *information power*, developed by Malterud and colleagues, was used as guidance for determining the sample size (i.e., the number of interviews) (77). In this view, information power depends on various characteristics of the study design and the data collected: study aim (narrow or broad), sample specificity (dense or sparse), established theory (applied or not), quality of dialogue (strong or weak), and analysis strategy (case or cross-case). In addition to discussing these characteristics, from the 14th interview and onwards, the thesis author and the last author of the publication discussed the extent to which additional interviews provided new or contradictory findings.

Data analysis. The audio-recorded interviews were transcribed verbatim. The transcribed data were managed and coded manually using NVivo 12 PRO (QSR International Pty Ltd). The data were analyzed using a qualitative thematic analysis (78). First, the thesis author (OE) conducted the initial coding for each interview separately. The first two interviews were read and discussed among the members of the research team. Second, OE reviewed the initial coding by re-reading the findings addressing EQ-5D-5L descriptive system, EQ VAS, and TTO, separately. In cases of clear overlap, the codes addressing the different assessments were combined. When certain findings mainly related to one or two assessments, this was indicated in the presentation of results. The analysis was an iterative process developing from description of patterns in the data (i.e., codes) to interpretation of meaning (i.e., categories and themes). Lastly, the categorization and interpretation were discussed and reconciled between the research team members. The quotes presented in the article were translated from Swedish into English.

4.4.3 Study III

Design and data source. Study III was a retrospective register-based study. Individual-level data for patients who had major LLAs between the years 2010 and 2018 were obtained from the SwedeAmp registry (67). Data from the six-month follow-up, supplemented by some demographic and clinical information registered in connection with the surgery, were used for the assessment of measurement properties.

The study population consisted of two subsamples of patients who had their six-month follow-up during one of two time periods: 1) from 2011 until 2016 – patients who responded to EQ-5D-3L (descriptive system), or 2) from 2017 and onwards – patients who responded to EQ-5D-5L (descriptive system and EQ VAS).

Sampling and participants. All patients who were included in SwedeAmp and met the inclusion criteria were included in the study. The inclusion criteria were to have had: a unilateral major amputation, including transtibial amputation (TTA), transfemoral amputation (TFA), or knee disarticulation (KD); primary amputation or re-amputation as surgical procedure; a follow-up assessment at six months after surgery, including an EQ-5D-3L or EQ-5D-5L measurement. Patients with missing data regarding sex and/or age were excluded. Data extracted for this study concerned amputations conducted between 2010 and 2018 and follow-ups conducted between 2011 and 2019, at six months after surgery.

Measures. EQ-5D-3L and EQ-5D-5L were the main measurements of interest for this study. EQ-5D data were analyzed and presented in several ways: distribution of responses on the five dimensions and EQ VAS (which was only available for EQ-5D-5L); EQ-5D health profiles; EQ-5D level sum scores; and EQ-5D index values. Index values were calculated using the UK value set for EQ-5D-3L (52) and the cross-walk value set for the UK for EQ-5D-5L (50). The two value sets both range from -0.594 to 1 and were chosen for the purpose of facilitating comparison between the EQ-5D versions.

Several disease-specific measures were included. The Locomotor Capabilities Index (LCI-5) is a patient-reported assessment of the capability to perform locomotor activities while wearing a prosthesis (79). LCI-5 consists of 14 items that are graded on a 5-point ordinal scale: 0 (no); 1 (yes, if someone helps me), 2 (yes, if someone is near me), 3 (yes, alone, with ambulation aids); 4 (yes, alone, without ambulation aids). Two subscales are constructed to represent basic capabilities (7 items) and advanced capabilities (7 items). The prosthetic use score assesses the amount of time that a person normally wears their prosthesis during a week (80). It is calculated by multiplying the number of days per week (0–7) by the number of hours per day (six response options), resulting in a value between 0 (no use) and 100 (each day for more than 15 hours). The Swedish translations of these instruments were used.

The assessment also included single item questions addressing the patient’s overall situation as amputee (very good; good; neither good nor bad; bad; very bad), and the occurrence of phantom limb pain (none; yes, a little; yes, moderate; yes, a lot) or residual limb pain (none; yes, a little; yes, moderate; yes, a lot) during the preceding three months. The questions were administered in Swedish and were translated into English by the members of the research team for this publication.

Statistical analysis. The participant characteristics of the subsamples were compared as regards sex, age at amputation, underlying diagnosis leading to the amputation, and amputation level. Tests for examining potential statistically significant differences included the chi-squared test (nominal categorical variables), the Mann-Whitney U-test (ordinal categorical variables), and the independent t-test (continuous variables), at a significance level of 0.05. Distribution of responses to the other (disease-specific) outcomes were presented for the two subsamples, and between-group differences were analyzed using the Mann-Whitney U test.

The measurement properties were explored through analyses of feasibility, distributional characteristics, and validity. For each of these assessments, pre-defined hypotheses were tested. Feasibility was defined as the proportion of missing data on the EQ-5D, overall or internal missing by dimension. For the analysis of overall missing, we used a variable from SwedeAmp indicating respondents who were invited to respond to EQ-5D but did not do so. The hypothesis was that a smaller proportion of missing data would be observed for EQ-5D-5L responses than for EQ-5D-3L responses. All patients with complete data on the EQ-5D descriptive system were included in the following analyses.

The assessment of distributional characteristics included assessment of response patterns and informativity. Response patterns were assessed by examining distributions of responses on each health dimension, the proportion of patients reporting no problems, and the number of unique health profiles on EQ-5D-3L and EQ-5D-5L. The hypothesis was that a smaller proportion would report no problems on EQ-5D-5L than on EQ-5D-3L. In addition, the distribution of index values was assessed for EQ-5D-3L and EQ-5D-5L, where a distribution with no or few clusters was preferred. Informativity was defined as the ability of an instrument to capture all possible health profiles and to discriminate between persons with different health status (33). Informativity was assessed using Shannon's indices per EQ-5D dimension (33). There are two indices: Shannon's index (H′) represents the extent to which information is evenly distributed across the possible response options, whereas Shannon's Evenness index (J′) corrects for the number of possible response options (56). The possible range of Shannon's index (H′) is 0–1.58 for EQ-5D-3L and 0–2.32 for EQ-5D-5L. The possible range of Shannon's Evenness index (J′) is 0–1 for both versions. A higher index value indicates that more information is being captured. The hypothesis was that a higher Shannon's index (H′) and an equal or higher Shannon's Evenness index (J′) value would be observed for EQ-5D-5L, compared with those for EQ-5D-3L.

Construct validity was assessed using convergent and known-groups validity. Convergent validity assesses whether a person's response to one instrument is associated with that person's response to another instrument which, at least in theory, should measure the same or similar constructs (9). Separate analyses were conducted for the two EQ-5D versions. Convergent validity was assessed by examining the strength of association of the specific

EQ-5D version and other measures using Spearman's rank-order correlation. The following thresholds were used to interpret the correlation: absent (correlation coefficient rho < 0.2), weak ($0.2 \leq$ rho < 0.35), moderate ($0.35 \leq$ rho < 0.5), and strong (rho ≥ 0.5) (81). Certain relationships were hypothesized on beforehand: EQ-5D mobility and LCI-5; EQ-5D self-care and LCI-5; EQ-5D usual activities and LCI-5; EQ-5D mobility and the prosthetic use score; EQ-5D pain/discomfort and phantom limb pain; EQ-5D pain/discomfort and residual limb pain; EQ-5D level sum score and overall situation as amputee; EQ VAS (for EQ-5D-5L only) and overall situation as amputee. For these associations, the hypothesis was to observe equal or stronger correlations for EQ-5D-5L compared to the EQ-5D-3L.

Known-groups validity assesses the extent to which an instrument discriminates between groups with anticipated or known differences in the construct covered by the measure that is being evaluated (9). The hypothesis was that patients with a higher amputation level (TFA/KD) and patients with an amputation due to diabetes and/or vascular disease would have worse health. In addition, the two EQ-5D versions were compared in terms of relative efficiency (RE), a measure used to evaluate which instrument is most sensitive in detecting differences between known groups (9). First, independent t-tests were conducted to compare mean index values for groups defined by amputation level and by underlying diagnosis, separately for each subsample. Second, a calculation of the RE ratio between the two EQ-5D versions were conducted based on the squared t-statistics, using the largest t-statistic as the denominator. Thus, the RE ratio had a value between 0 and 1 (strongest).

The statistical analyses were performed using IBM SPSS Statistics 27.0 (Armonk, NY: IBM Corp).

4.4.4 Study IV

Design and data sources. Study IV was a retrospective observational register-based study, using individual-level data from the Q-ECT registry. In addition, linkages were made to the longitudinal integrated database for health insurance and labour market studies (LISA), the National Patient Registry, and the Swedish Prescribed Drug Registry. The study design and statistical analyses were chosen before data were obtained from the registries [for study protocol, see (82)].

Sampling and participants. Inclusion criteria were formulated to identify study participants in Q-ECT. The Swedish versions of the ICD-10 codes were used to identify certain indications for treatment. The inclusion criteria for this study were: adults (≥ 18 years); major unipolar depression (ICD-10 codes F32.1–F.32.3, F33.1–F33.3) or bipolar depression (ICD-10 codes F31.3–F31.5) as indication for treatment; having received index treatment with unilateral electrode placement at first ECT session; and lastly, having EQ-5D measurements (EQ-5D index or EQ VAS) at least before and within one week after ECT, and having information on pulse width at first ECT session. All patients in Q-ECT who met the inclusion criteria were included in the study. If a patient had several treatment series that fulfilled the inclusion criteria, only the first treatment was included.

Only index treatments were included, meaning that patients in general had received three ECT sessions a week until remission, or a physician judged that the maximum treatment benefit possible had been achieved within the specific ECT treatment series. All patients included in the study population had unilateral electrode placement (d'Elia) in their first ECT session. The rationale for this criterion was that the use of unilateral electrode placement is standard practice in the treatment of unipolar or bipolar depression in Sweden, and therefore is used in approximately 80% of the index treatments registered in Q-ECT (70).

The ECT treatment series for patients included in the study population were carried out in Sweden during 2011–2019. All 21 regions were represented in the obtained data. However, the majority of patients in the study population received treatment in one of the three largest regions, Stockholm (39%), Västra Götaland (9%), or Skåne (9%).

Explanatory and outcome variables. The study was conducted to explore the influence of one ECT treatment parameter, namely pulse width (expressed in milliseconds [ms]). The main explanatory variable was pulse width during the first ECT session. The pulse width (range 0.25–1.00 ms) was categorized into three subgroups: < 0.5 ms; = 0.5 ms; > 0.5 ms. The primary outcome was HRQoL (as assessed using EQ-5D-3L index and/or EQ VAS) within one week after ECT. The secondary outcome was HRQoL six months after ECT. EQ-5D-3L index values were calculated using the UK value set (52).

Statistical analysis. Demographic and clinical characteristics were summarized for the total sample and for subgroups receiving different pulse widths, to compare their characteristics. Between-group differences were tested using the chi-squared test, the chi-squared test for trend, or analysis of variance (ANOVA), depending on the data characteristics. Furthermore, descriptive statistics were used to compare the study population and the registry population for purposes of assessing the generalizability of the study findings.

For both EQ-5D-3L index and EQ VAS, the means of the differences between baseline and within one week after ECT were examined for the total sample using paired t-tests. The threshold for a minimally important difference in EQ-5D index was 0.082 (83). Changes in health, as assessed using the EQ-5D descriptive system, were analyzed using the PCHC (72).

The association between pulse width and HRQoL after ECT was examined using multiple linear regression analyses, with adjustment for baseline HRQoL (84). Since the Bruesch-Pagan test indicated presence of heteroscedasticity, ordinary least squares (OLS) regression with robust standard errors was used (85). All primary and secondary outcomes were examined as continuous variables and tested in separate regression models (and interpreted after Bonferroni correction). Covariates were included based on their expected association with the explanatory variable and outcome variables. Model 1 included pulse width and HRQoL at baseline. In addition, sex, age, and indication were added into Model 2. Lastly, concurrent treatment (with a larger anticipated proportion of missing data) was added into the final model, Model 3. For all covariates in Model 3, multicollinearity was examined by interpreting the variance inflation factor.

The statistical analyses were conducted with a significance level of 0.05. After the Bonferroni correction, the regression models were interpreted using a significance level of 0.0125. Statistical analyses were conducted using IBM SPSS Statistics 27.0 (Armonk, NY: IBM Corp) and STATA Statistical Software 15 (College Station, TX: StataCorp LCC).

Additional analyses. This thesis contains additional analyses for Study IV. Distributions of responses before and after ECT were summarized by EQ-5D dimension and EQ VAS score for the total sample. Furthermore, to mitigate the possible influence of confounding by indication and covariate imbalance between groups (86), the association between pulse width and HRQoL was further explored by conducting inverse-probability-weighted regression adjustment (IPWRA) for Model 3 (using the *teffects ipwra* command in STATA). First, a multinomial logit regression was applied for the treatment model, i.e., for estimating the probabilities of treatment assignment. Variables representing patient characteristics were included based on their anticipated impact on being assigned a certain pulse width at first ECT (i.e., sex, age, and indication). Second, a multiple linear regression model with robust standard error was applied for the outcome model, including all covariates in Model 3. Diagnostics, including overlap and covariate balance, were checked. This procedure was repeated for all four outcomes, i.e., EQ-5D index and EQ VAS both within one week after ECT and six months after ECT.

4.5 ETHICAL CONSIDERATIONS

4.5.1 Ethical review

Approval was granted from an ethical review board for the research studies including personal data: Study II (Regional Ethical Review Board in Stockholm, Ref. no: 2017/526-31; 2017/2123-32), Study III (Regional Ethical Review Board in Stockholm, Ref. no: 2018/1137-31/2), and Study IV (Regional Ethical Review Board in Uppsala, Ref. no: 2014/174; 2014/174/1; 2014/174/2; 2020-05154). Ethical review was not required for Study I, as it did not include any personal data or sensitive information.

4.5.2 Benefits and risks

The sub-studies differed in design, type of data, and involvement of participants. Thus, different potential benefits and risks were identified for each study.

Study I. The potential risks were considered limited, as no personal information was processed. The overview of how EQ-5D data were collected, made available, and used was associated with benefits both for the understanding of the context of the thesis and as a contribution to the relatively limited scientific literature on routine collection and use of PROMs in health care.

Study II. The study used qualitative interview data gathered with active participation from study participants. Participation in the study was not expected to lead to any direct benefits for the participants, except the possible perception that their experiences were shared. Ethical considerations in this study included voluntary participation, informed consent, confidentiality, and correct processing of personal data. In the first recruitment step, participants were given written information about the study. In addition to information about the purpose and procedure of the research study, the written information specified that: participation was voluntary and participants had the right to withdraw at any time; the study had been approved by an ethical review board; personal data would be processed and stored by Karolinska Institutet; the health care provided would not be affected by their decision to participate or not. The written information also specified compensation for participation. Participants had the opportunity to ask questions in the second stage of recruitment, when contacted by telephone by the thesis author. The nurses who initiated the recruitment were not informed about who participated in the study. All study participants gave informed consent prior to their interview.

Another identified risk was related to the sensitive topic, i.e. that questions could provoke thoughts and reflections both during and after the interview. The choices regarding the study sample and some inclusion criteria (i.e. age and years since diagnosis) were made to avoid having interviews coincide with a new diagnosis or treatment. Participants were informed through the written study information that the interview would include questions about health, disease, life, and death. Lastly, all participants received contact details to a counsellor at the end of the interview.

Studies III–IV. The register-based studies were not expected to lead to any direct benefits for the study participants. Still, making use of collected data may contribute to indirect benefits for patients and health care professionals who invest time and effort in recording health data. Register data could also be considered to be used well for when included in methodological research (as in Study III), since data collection solely for that purpose is resource-demanding and carries a risk of limited benefits for the patients and others involved. Precautions for managing the identified potential risks of the register-based studies included secure processing of the pseudonymized personal data.

5 RESULTS

The findings from each of the four sub-studies are summarized in this section. All four sub-studies are relevant to the case used in this thesis: the collection and use of EQ-5D in the Swedish NQRs. The study-specific research questions were examined for a number of patient populations, including a variety of patient populations and interventions covered by the NQRs included in the overview, and the specific study samples. The demographic characteristics for each of the study samples are summarized in Table 3. No such information is presented for Study I, as it entailed information about the 41 registries that included EQ-5D data at the time of data collection.

Table 3. Overview of demographic characteristics, by study samples, for Studies II–IV.

	Study II	**Study III**	**Study IV**
Patient population	Patients with type 1 diabetes	Patients with a major lower limb amputation	Patients with unipolar or bipolar depression
Total, n	20	685	5,046
Women, n (%)	7 (35)	221 (32)	3,008 (60)
Men, n (%)	13 (65)	464 (68)	2,038 (40)
Age, mean (SD)	46 (15)	72 (14)	53 (18)
18–39 years, n (%)	8 (40)	17 (3)	1,317 (26)
40–59 years, n (%)	6 (30)	90 (13)	1,703 (34)
60–79 years, n (%)	6 (30)	366 (53)	1,665 (33)
80–100 years, n (%)	0 (0)	212 (31)	361 (7)

The following samples and subsamples were included in the studies: Study II included 20 patients with type 1 diabetes; Study III included 685 patients with a major LLA; Study IV included 5,046 patients who were treated with ECT for unipolar or bipolar depression. EQ-5D results were available for all study participants in Study II–IV (Table 4). In Study II, EQ-5D index and EQ VAS scores were in the ranges 0.50–1 and 30–95, respectively, at the time of the interviews. For the two subsamples in Study III, the mean EQ-5D-3L and EQ-5D-5L index values were 0.55 and 0.53 at the six-month follow-up after a major LLA. In Study IV, the mean EQ-5D-3L index values increased from 0.31 before ECT to 0.71 one week after ECT, and the equivalent increase in mean EQ VAS scores was from 26 to 61.

Table 4. EQ-5D results in Studies II–IV, presented by subsamples of respondents or by point of measurement.

	Study II [1]	**Study III** [2]		**Study IV** [3]	
Condition	Type 1 diabetes	Diabetes and/or vascular disease; Other diagnoses		Unipolar or bipolar depression	
Measurement and intervention/treatment	n/a	Six months after major lower limb amputation		Before and after electroconvulsive therapy	
Subgroups: Subsample or measurement	n/a	EQ-5D-3L sample	EQ-5D-5L sample	Before ECT	One week after ECT
Number of observations, n	20	425	260	5,046	5,046
EQ-5D version	EQ-5D-5L	EQ-5D-3L	EQ-5D-5L	EQ-5D-3L	EQ-5D-3L
Ceiling (profile 11111), n (%)	9 (45.0)	28 (6.6)	15 (5.8)	38 (0.8)	1,078 (21.6)
Floor (profile 33333 or 55555), n (%)	0 (0.0)	0 (0.0)	0 (0.0)	13 (0.3)	<5 (<0.1)
EQ-5D index, min–max	0.50–1	-0.35–1	-0.40–1	-0.594–1	-0.594–1
EQ-5D index, mean (SD)	0.87 (0.15)	0.55 (0.31)	0.53 (0.31)	0.31 (0.29)	0.71 (0.28)
EQ VAS, min–max	30–95	n/a	0–100	0–100	0–100
EQ VAS, mean (SD)	75 (16.09)	n/a	67 (20.52)	26 (18.91)	61 (23.34)

[1] Study II: Mean EQ VAS scores calculated for 17 respondents.

[2] Study III: Diabetes and/or vascular disease include: diabetes with or without vascular disease; atherosclerosis without diabetes; other peripheral vascular diseases. Other diagnoses include: trauma; tumor; infection unrelated to diabetes or vascular disease; acquired deformity; congenital deformity; other; unknown.

[3] Study IV: EQ-5D index values, n=4,990; EQ VAS scores, n=4,914.

5.1 STUDY I – OVERVIEW OF EQ-5D IN THE SWEDISH NATIONAL QUALITY REGISTRIES

Forty-one of the 46 NQRs with an EQ-5D license administered EQ-5D to measure patient-reported health in their registry population. Ten of the registries were sub-registries of the Swedish Neuro Registries. The 41 registries targeted patient populations with a variety of diagnoses and/or interventions. Most frequently, EQ-5D was measured and collected in registries targeting conditions related to the musculoskeletal system (n=11), the nervous system (n=9), the circulatory system (n=7), and less frequently in registries targeting cancer (n=3), psychiatry (n=3), pediatrics (n=2), infection (n=2), other areas (n=2), the endocrine organs (n=1), or the stomach and intestines (n=1).

Registries used different modes of administration and times of measurement in the collection of EQ-5D data. Twenty-four registries administered EQ-5D-3L and 16 administered EQ-5D-5L (Table 5). Examples of different times of measurement included before and after intervention, continuous measurement at specified time-points, or once per patient, and examples of different modes of administration included paper questionnaires or web questionnaires. All but one of the 41 registries also administered other PROs, such as disease-specific PROMs or questions regarding general health, symptoms, function, use of assistive devices, or working ability.

Table 5. Numbers of registries reporting collection of EQ-5D-3L or EQ-5D-5L (2018), by category and type.

	Total, n	**Intervention registries**, n		**Diagnosis registries**, n	
		EQ-5D-3L	EQ-5D-5L	EQ-5D-3L	EQ-5D-5L
Musculoskeletal system	11	4	4	2	1
Nervous system [1]	9	2	0	4	3
Circulatory system	7	1	1	4	1
Cancer	3	0	0	1	2
Psychiatry	3	2	0	1	0
Infection	2	0	0	0	2
Other areas	2	0	0	2	0
Endocrine organs	1	0	0	1	0
Pediatrics [2]	1	0	0	0	1
Stomach and intestines	1	0	0	0	1
All registry categories	40	9	5	15	11

[1] Of which six were sub-registries within the Swedish Neuro Registries

[2] In addition, one registry collected EQ-5D-Y, the child-friendly EQ-5D version, only.

EQ-5D data from the NQRs were summarized and presented at either an individual or an aggregate level. The most frequently reported channel for presenting results was through the registries' annual reports or websites (n=29). Other examples of how aggregate data were made available included reports of aggregate data directed at clinics, units, or professional teams (n=20), and scientific publications (n=18). Several registries made individual-level data available to health care professionals (n=17) and/or patients (n=12).

The overview shows that EQ-5D data collected in the Swedish NQRs were being used in quality indicators, for assessment of interventions, health economic studies, quality improvement, benchmarking, and/or in individual patient consultations (Figure 1). Twenty-two registries reported use of EQ-5D data, while 19 registries reported that data were not used, or that they were unaware of whether data were used. Use of EQ-5D data was reported by 12 of the 26 diagnosis registries (46%) and 10 of the 15 (67%) intervention registries.

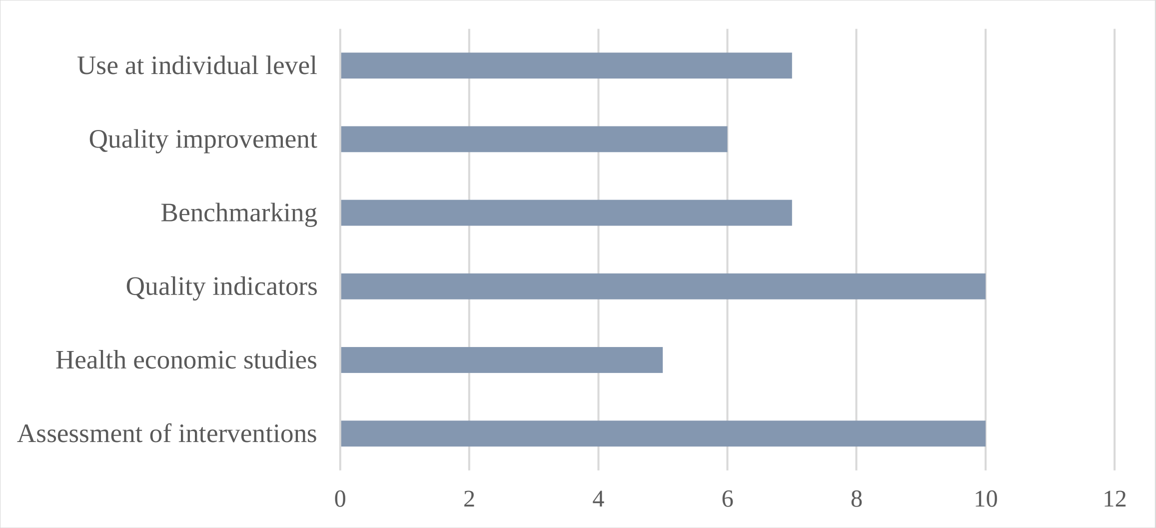

Figure 1. Numbers and distribution of 22 registries reporting use of EQ-5D data for different purposes. Note that each registry reporting use of EQ-5D data could report several categories of use.

More specifically, in regard to the use of EQ-5D data in quality indicators and for benchmarking, registries commonly reported that data from these analyses were publicly available in the registries' own reports or online tools, and on national platforms. One example of such a platform is “Health care in numbers” (“Vården i siffror”), which is an online tool for quality improvement initiatives, with continuous updates concerning indicators of quality and efficiency in Swedish health care (87).

5.2 STUDY II – THOUGHTS WHEN REPORTING AND VALUING HEALTH

The thematic analysis of how individuals think and reason when reporting their own current health resulted in two themes (i–ii) based on four categories (a–d). When reporting and valuing one's own current health using EQ-5D-5L, EQ VAS, and TTO, the thought processes of individuals interpreting and completing these assessments involved (i) personalizing questions and considering what aspects to include in the response, and (ii) using reference points and comparators to enable assessment of their own health. The categories described how interview participants approached the assessments by (a) contextualizing and interpreting instructions, (b) relating the questions to their own health, (c) using different recall periods and time perspectives, and (d) using personal, interpersonal, or normative comparators.

In summarizing the findings, some patterns should be highlighted. First, participants reflected on the purpose and usefulness of results as part of the process of deciding what aspects to consider in their response. Participants personalized the questions when interpreting the questionnaires, for example by linking certain aspects of their own health to the presumed purpose of the questionnaire. Participants actively reasoned regarding what aspects to include or not in their assessment. This process raised concerns about the subjectivity of the questions and, by extension, the interpretability and usefulness of the results. Some participants described wanting to explain their responses, either through follow-up questions or in dialogue with their health care provider.

> *"If someone notices it and say 'why are you 75 instead of 100?' Then you can talk about it. But it requires that someone looks at that damn questionnaire that no one looks at, when you fill it out in the reception and like, check, I have filled out this form a hundred times and no one cares. Eh. So that's part of it.. So, like, you need an interview to capture it. Or that someone who sits there really understands – oh what does this mean?"* [Participant 3]

Second, different time perspectives and recall periods were applied for the three assessments. When responding to the EQ-5D-5L descriptive system and EQ VAS, participants generally considered their health either at the present time or retrospectively over a longer period of time. When responding to the TTO, participants described thoughts of their future. These thoughts included expectations of their future health status, their subjective life expectancy, and activities and goals that they desired to fulfil. Fifteen participants did not trade any life years to live for a shorter period in full health. Reasons for this decision included that one's experienced problems were considered manageable and wanting to have as much time as possible to support and/or spend time with children, family, and friends.

> *"Yeah, but the thing about death and that I only have ten years left to live, okay, then I start counting how old my children will be... And all that... How long will I get to be with my husband (...) And how can I maximize that?"* [Participant 16]

Third, different interpretations and reasoning were expressed in regard to the terms "best imaginable health" in EQ VAS and "full health" in the TTO question. Different reference points were applied in the EQ VAS assessment (e.g., an ideal person or personal circumstances). Still, the concept of full health in the TTO question was considered difficult to define, imagine, and relate to. Two participants decided not to answer the question since the scenario of full health was considered unrealistic.

> *”You never get that opportunity... To live in full health, that is not an option that exists in reality. Like, what you can do is to live as good as you can with the time you have left, I think”* [Participant 18]

Lastly, some findings were more clearly linked to one or another of the assessments. Especially in regard to the descriptive system, some participants expressed a desire to discuss aspects associated with health rather than aspects associated with problems or disease. In addition, one of the most prominent examples was that the impact on others was more clearly integrated in the response to the TTO.

> *“But here it is definitely- I want to have as much time as possible with, eh.. My partner, my family (…) So then it is no longer just based on me, but here is it also about… Giving time to others. While the two first are definitely how I experience myself. Or mostly focused on that I would say (…) But this was definitely something I kept in mind, that others should be given time as well”* [Participant 11]

5.3 STUDY III – MEASUREMENT PROPERTIES OF EQ-5D

In the study sample, which included 685 patients, more than two thirds were men (68%). Patients had either a TTA (below-knee amputation) (76%), a TFA (above-knee amputation) (18%), or a KD (knee disarticulation) (7%). The majority of patients had diabetes and/or vascular disease as the underlying diagnosis (83%) and the mean age at amputation was 72 years (range 20–100).

Two subsamples were used for the assessment and comparison of measurement properties: EQ-5D-3L respondents (n=425) and EQ-5D-5L respondents (n=260). The subsamples did not differ significantly in the distributions of sex, age, amputation procedure, level of amputation, or underlying diagnosis (p>0.05). Regarding PROMs and PROs, no significant between-group differences were observed regarding the overall situation as amputee, LCI-5, or the prosthetic use score (p>0.05). However, a larger proportion of the EQ-5D-5L respondents reported phantom limb pain (p<0.001) and residual limb pain (p<0.05).

Measurement properties of EQ-5D-3L and EQ-5D-5L. The study sample for assessing the distributional characteristics and validity included 685 patients with complete EQ-5D data at follow-up six months after amputation. An additional 15 patients (73% men, mean age 73.9 years) were included only for the analysis of feasibility (i.e., analysis of missing data) of EQ-5D-3L and EQ-5D-5L. The feasibility of EQ-5D-3L and EQ-5D-5L, as assessed based on the proportion of complete responses to the descriptive system, was 98% for both versions. Ten patients in the EQ-5D-3L subsample and 5 patients in the EQ-5D-5L subsample had no, or an incomplete, EQ-5D measurement.

Regarding the distributional characteristics of responses to the EQ-5D-3L and EQ-5D-5L, patients most frequently reported problems related to mobility (78% and 75%), pain/discomfort (70% and 75%), and usual activities (61% and 74%). Less frequent reporting of problems was observed for anxiety/depression (44% and 52%) and self-care (38% and 48%). Severe or extreme problems with mobility were reported by 23% on EQ-5D-5L and only 7% on EQ-5D-3L. The proportion of responses indicating no problems in any of the dimensions (i.e. the ceiling) was 7% for EQ-5D-3L and 6% for EQ-5D-5L. No problems were most frequently reported in the self-care dimension for both EQ-5D-3L (62%) and EQ-5D-5L respondents (52%) (Table 6). The mean EQ-5D index was 0.55 (SD 0.31) for EQ-5D-3L respondents and 0.53 (SD 0.31) for EQ-5D-5L respondents.

Table 6. Numbers and proportions of patients reporting no problems (ceiling) in each of the health dimensions six months after a major lower limb amputation, by subsamples of EQ-5D-3L and EQ-5D-5L respondents.

	EQ-5D-3L sample	**EQ-5D-5L sample**
	n=425	n=260
Mobility, n (%)	95 (22.4)	64 (24.6)
Self-care, n (%)	262 (61.6)	136 (52.3)
Usual activities, n (%)	168 (39.5)	65 (25.0)
Pain/discomfort, n (%)	129 (30.4)	69 (26.5)
Anxiety/depression, n (%)	240 (56.5)	124 (47.7)

Furthermore, EQ-5D-5L showed higher Shannon’s index values (H′) and Shannon’s Evenness index values (J′) than EQ-5D-3L across all dimensions. The largest difference between the versions was observed in the mobility dimension: Shannon’s index (H′) was 1.09 for EQ-5D-3L and 2.21 for EQ-5D-5L, and Shannon’s Evenness index (J′) was 0.69 for EQ-5D-5L and 0.95 for EQ-5D-5L. The mobility dimension was of particular interest due to the change from “confined to bed” in EQ-5D-3L to “unable to walk about” in EQ-5D-5L. When analyzing the EQ-5D-3L dimensions only, the lowest informativity was observed for the mobility dimension, in which 71% of the respondents used severity level 2 (moderate problems).

The statistical analysis of convergent validity showed correlations between the two EQ-5D versions and other measures that were in line with the hypotheses. For both versions, the dimensions of mobility and self-care showed moderate to strong correlations (rho ≥ 0.42) with LCI-5 scores. By comparison, the correlations between usual activities and LCI-5 scores were slightly weaker, i.e., rho -0.31 to -0.32 for EQ-5D-3L and -0.34 to -0.36 for EQ-5D-5L ($p<0.01$). For both EQ-5D versions, moderate correlations were observed between the mobility dimension and the prosthetic use score and between the pain/discomfort dimension and the questions related to phantom and residual limb pain. A moderate correlation was also found between the anxiety/depression dimension and the overall situation as amputee, which was not covered by the pre-specified hypotheses. Although only available for EQ-5D-5L, the EQ VAS showed a moderate correlation with the overall situation as amputee (rho 0.46).

The interpretation of findings related to known-groups validity was also based on hypotheses. On average, worse HRQoL was observed among women, patients of older age, and patients with a higher amputation level. However, none of the EQ-5D versions indicated differences in HRQoL between subgroups by underlying diagnosis. Only EQ-5D-5L could discriminate between subgroups defined by different amputation levels (RE ratio 0.538). A summary of the findings is presented in Table 7.

Table 7. Summary of the findings related to feasibility, distributional characteristics, and validity of EQ-5D-3L and EQ-5D-5L in patients with a major lower limb amputation.

	EQ-5D-3L	**EQ-5D-5L**	**Summary of findings**
Feasibility	/	/	The proportions of complete responses on the descriptive system were similar (97.7% for EQ-5D-3L versus 98.1% for EQ-5D-5L)
Ceiling (i.e., proportion reporting no problems in all five dimensions)	/	/	The proportions were similar (6.6% for EQ-5D-3L and 5.8% for EQ-5D-5L)
Informativity	–	+	The informativity was higher for EQ-5D-5L than for EQ-5D-3L across all dimensions
Convergent validity	–	+	The correlations between EQ-5D and other disease-specific measures were generally stronger for EQ-5D-5L than for EQ-5D-3L
Known-groups validity	–	+	Only EQ-5D-5L was able to discriminate between groups defined by amputation level; Neither the EQ-5D-3L nor EQ-5D-5L index values were able to discriminate between groups defined by underlying diagnosis

Note: + indicates higher or better results; – indicates lower or worse results, / indicates similar results.
In this table, the results for two separate subsamples (EQ-5D-3L and EQ-5D-5L respondents) are summarized and compared, based on response patterns and to what extent the results are in line with the study hypotheses.

5.4 STUDY IV – USING EQ-5D TO ASSESS OUTCOMES AFTER TREATMENT

In Study IV, the study sample included 5,046 patients with unipolar (82%) or bipolar depression (18%). Similar to the Q-ECT registry population, the study sample included approximately 60% women and 40% men, and the mean age was 53 years (SD 18.2). EQ-5D index and EQ VAS measurements before and within one week after ECT were extracted for 4,990 and 4,914 patients, respectively (4,858 patients completed both). A follow-up measurement at six months after ECT was available for a subsample (EQ-5D index, n=730; EQ VAS, n=851).

Patients received a pulse width of either <0.5 ms (15%), 0.5 ms (72%), or >0.5 ms (13%) at their first ECT session. There were statistically significant differences between the subgroups in terms of the distribution of sex, age, and the presence or absence of psychotic features (p<0.05). The subsample receiving <0.5 ms pulse width was on average younger and included a larger proportion of women, compared to the other two subsamples (p<0.0001). In addition, the proportion of patients with psychotic features was smaller among those receiving <0.5 ms (12%) compared with the other two subsamples (18–19%) (p<0.0001). Furthermore, there were statistically significant differences in the proportion of patients with concurrent use of medication, e.g., antidepressants (p=0.032), antipsychotics (p<0.0001), benzodiazepines (p<0.0001), and antiepileptics (p=0.020).

Before ECT, the mean EQ-5D index value was 0.31 (SD 0.29) and the mean EQ VAS score was 26 (SD 18.91) for the total sample. Within one week after ECT, the corresponding numbers were 0.71 (SD 0.28) for EQ-5D index and 61 (SD 23.32) for EQ VAS. Per subgroup, the mean improvements between the measurements were 0.38 (<0.5 ms), 0.40 (0.5 ms), and 0.36 (>0.5 ms) for EQ-5D index, and 33 (<0.5 ms) and 35 (0.5 ms and >0.5 ms) for EQ VAS.

Association between pulse width and HRQoL after ECT. The regression analyses showed no statistically significant associations between pulse width and HRQoL within one week after ECT, after adjusting for HRQoL at baseline, sex, age, depression diagnosis, psychotic features, and concurrent medications (p>0.05). After adjusting for these characteristics, the regression analysis for the secondary outcomes six months after ECT showed significantly lower EQ-5D index values (β -0.089) for the subsample receiving 0.5 ms, compared with those receiving <0.5 ms (p=0.011). The result of the corresponding analysis for EQ VAS six months after ECT was not statistically significant (p>0.05).

5.4.1 Results of additional analyses

Association between pulse width and HRQoL after ECT. As an additional analysis to further examine the association between pulse width and HRQoL after ECT, IPWRA was explored. For Model 3, these analysis showed non-significant associations (for pulse width 0.5 ms and >0.5 ms compared to the reference <0.5 ms) for all primary and secondary outcomes, including EQ-5D index at six months (β -0.047; 95% CI -0.112–0.019) (Appendix C).

Changes in HRQoL from before ECT to one week after ECT. The response distributions for the total sample showed overall improvements across all health dimensions within one week after ECT (Figures 2–3), in particular in the dimensions for usual activities and anxiety/depression. Overall, improvements were observed also for EQ VAS (Figure 4). Still, approximately one out of ten patients reported extreme problems with anxiety/depression (13%) and/or usual activities (10%), and/or EQ VAS scores ≤ 25 (10%) one week after ECT.

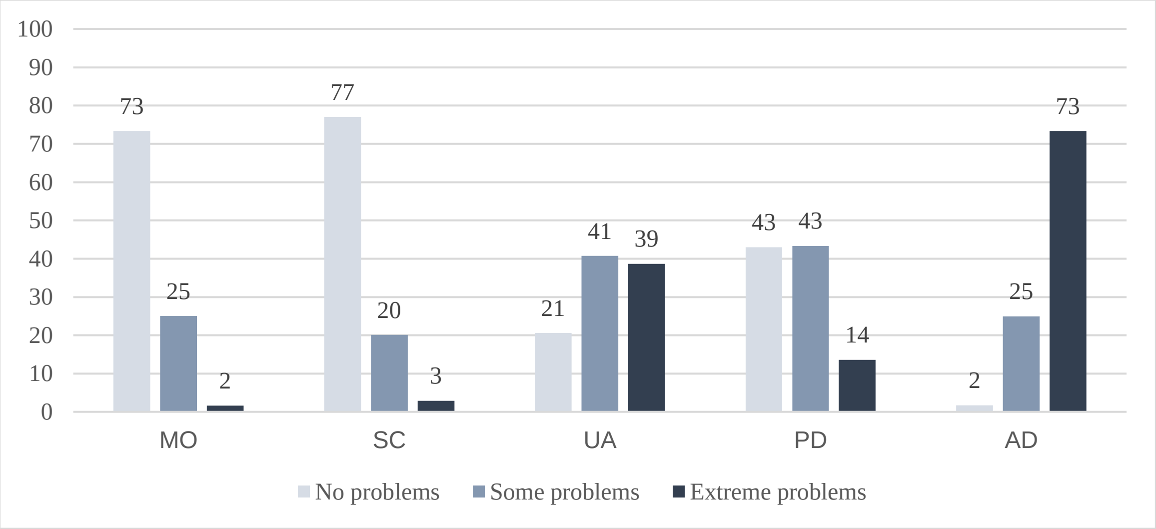

Figure 2. Distributions of responses (%) to the EQ-5D descriptive system before ECT (n=4,990).
Note: MO = mobility, SC = self-care, UA = usual activities, PD = pain/discomfort, AD = anxiety/depression.

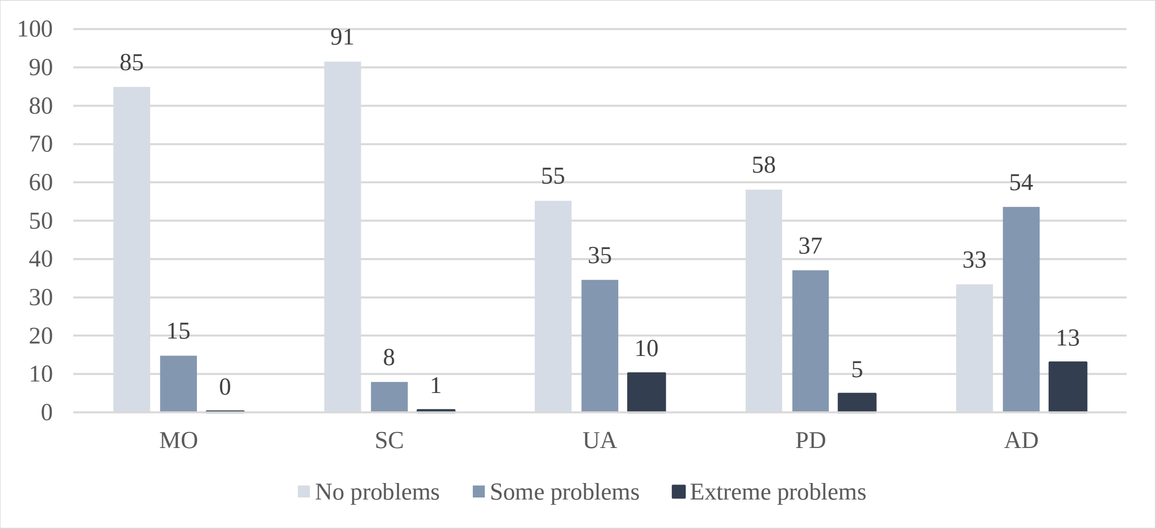

Figure 3. Distributions of responses (%) to the EQ-5D descriptive system one week after ECT (n=4,990).
Note: MO = mobility, SC = self-care, UA = usual activities, PD = pain/discomfort, AD = anxiety/depression.

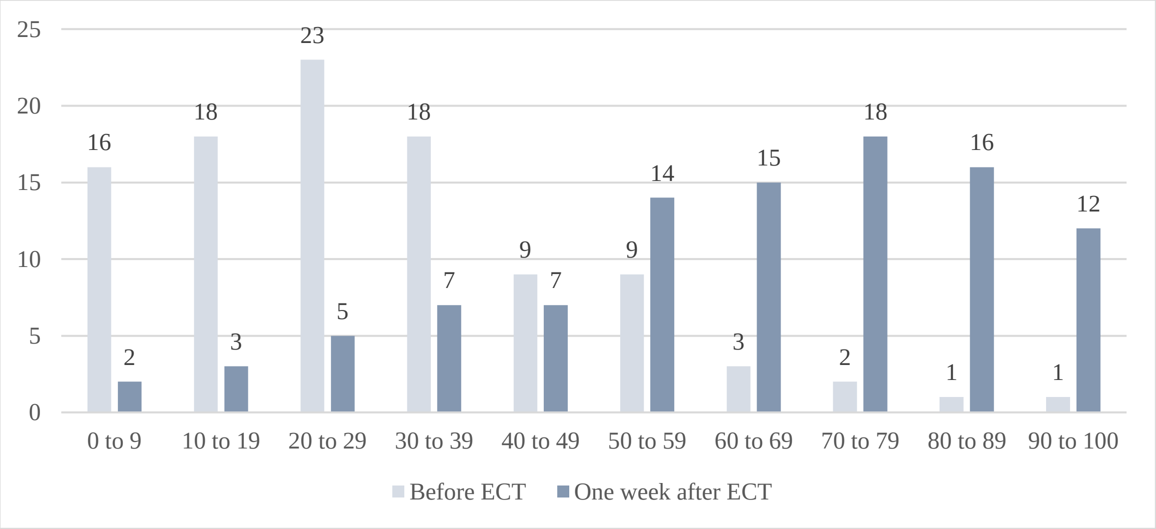

Figure 4. Distributions of EQ VAS responses (%) before ECT (n=4,914) and one week after ECT (n=4,914).
Note: EQ VAS ranges between 0 (worst imaginable health state) and 100 (best imaginable health state).

6 DISCUSSION

This thesis aimed to increase knowledge on the use of PROMs routinely collected in health care by examining a specific case: the use of EQ-5D routinely collected in the Swedish NQRs. The results revealed that EQ-5D data were collected in 41 Swedish NQRs, covering a variety of patient populations and interventions. Most registries had administered EQ-5D for several years, and two had collected EQ-5D data for 20 years. EQ-5D was most frequently administered in registries targeting conditions related to the musculoskeletal system. For patients with amputation of a lower limb, the results demonstrated several advantages of the measurement properties of EQ-5D-5L compared to the EQ-5D-3L. The use of EQ-5D data covered several purposes, including assessment of interventions, health economic studies, quality indicators, benchmarking, quality improvement, and use at the individual patient level. When applying EQ-5D as an outcome measure for examining HRQoL after ECT in patients with unipolar or bipolar depression, no robust association was found between a certain treatment parameter (pulse width) and patients’ HRQoL after ECT. Lastly, the thesis also presented findings regarding how patients think and reason when reporting and valuing their own current health using EQ-5D-5L, EQ VAS, and TTO. Some variation was revealed in the individual interpretation of the questionnaires, for example with regard to what aspects to include in the assessment and use of different time perspectives and comparators.

6.1 THE CASE OF EQ-5D IN THE SWEDISH NATIONAL QUALITY REGISTRIES

The results from this thesis showed that EQ-5D was administered for a variety of conditions and interventions in Swedish health care, often in combination with other PROs. In the early development of EQ-5D, the instrument was intended to be used alongside other measures, as well as for facilitating resource allocation decisions (48). Based on the studies included in this thesis, some benefits of using EQ-5D in combination with other measures may be highlighted. As could be expected, most measures for patients with LLA focused specifically on mobility and physical functioning, while most measures for patients with major depression specifically focused on depressive symptoms and memory. In these two cases, the generic EQ-5D might contribute to a more holistic perspective on the impact on a patient’s HRQoL, by combining both effects and adverse effects related to physical and mental aspects in one measure. Still, though the wide implementation of EQ-5D across populations facilitates between-group comparison, the usefulness of EQ-5D for specific purposes will likely vary depending on the patient population and the medical area. For example, the timing of measurement may vary depending on whether the purpose of a registry is to follow a defined population or a specific intervention.

The case of EQ-5D in the Swedish NQRs provided an opportunity to study the use of PROMs across several different levels of the health care system, including both use at the individual patient level in clinical practice and aggregate-level analyses. Although the NQRs have an existing infrastructure for nationwide data collection in various populations and for various conditions, the thesis showed that some challenges in making use of data remain. Nineteen registries reported that EQ-5D data were not being used for follow-up, quality improvement, or decision-making. Furthermore, the thesis demonstrated few examples of how routinely collected EQ-5D data were used for decision-making at higher levels of the health care system. One important learning outcome from this case is that routine data collection alone is an insufficient strategy for realizing the potential benefits of PROMs data.

While reflecting on the results from this case, an essential consideration for implementing a PROM in routine health care is the extent to which a selected measure can enable for meaningful and actionable interpretations. One of the main potential benefits of implementing a generic preference-based measure, as opposed to other PROMs, lies in its many areas of use for analysis at an aggregate level and across conditions. The thesis revealed examples of EQ-5D data being included in individual patient-clinician encounters and in quality indicators, assessments of interventions, health economic studies, quality improvement, and benchmarking. However, some aspects of the use of PROMs in routine health care may have been missed in this case due to the specific focus on one PROM. Since several PROMs are often needed to capture aspects that are relevant for patient health, one could also expect that several PROMs and/or other outcome measures would be needed to achieve all potential benefits, from micro- to meso- and macro-level analyses. Thus, the purpose of and intended mechanisms for making use of data could guide the choice of measure and design in future implementations of PROMs in routine health care.

6.2 PREREQUISITIES FOR MAKING USE OF DATA: MEASUREMENT AND INTERPRETABILITY OF RESULTS

The extent to which useful interpretations can be made depend on the ability of the selected outcome measure to provide valid and reliable results (9). Ideally, a PROM should have satisfactory measurement properties and cover aspects that are relevant for the specific patient group of interest. As shown in Study I, EQ-5D is commonly implemented in areas related to the musculoskeletal system, and several registries, including the SwedeAmp registry, have shifted from administering EQ-5D-3L to administering EQ-5D-5L. In patients with an LLA, satisfactory measurement properties were demonstrated for both EQ-5D-3L and EQ-5D-5L, with indications of better performance for EQ-5D-5L, especially in terms of improved informativity and validity. Thus, these results indicated that the change from EQ-5D-3L from EQ-5D-5L have improved the measurement of patients’ HRQoL following major LLA in the registry. At the same time, the performance of the instrument used for comparison should also be considered. Several disease-specific measures were used for comparison, including both established PROMs and single-item questions developed specifically for SwedeAmp. The disease-specific LCI-5 has been evaluated in terms of validity, reliability, responsiveness, and sensitivity in patients participating in prosthetic training after a unilateral LLA (79, 88), and the measurement properties of the prosthetic use score and the overall situation as an amputee were assessed during their development (80).

In Study IV, EQ-5D was applied as an outcome measure for examining the association between pulse width and HRQoL after ECT. Although concerns have been raised regarding the use of EQ-5D for measuring other conditions related to mental health (62), satisfactory measurement properties have been demonstrated in patients with major depression (63). A related concern is the extent to which EQ-5D captures aspects relevant for patients. A previously conducted interview study presented a variety of aspects important for the HRQoL of patients with various mental health conditions, including well-being and ill-being; physical health; self-perception; hope and hopelessness; autonomy, control, and choice; relationships and belonging; and activity (89). In the design of Study IV, EQ VAS was included as an additional primary outcome measure to allow aspects beyond the five health dimensions to be captured. Still, both the outcomes based on direct assessments from the patient (EQ VAS) and the outcomes based on stated preferences (EQ-5D index) led to the same conclusions.

In addition to the measurement properties, the methods and perspectives applied for health state valuation are central to the interpretation of EQ-5D index values. These values provide information valuable for the purpose of health economic evaluations. Still, there are several other ways in which EQ-5D data may be presented (72), which could provide more detailed information regarding the changes in specific dimensions of health. For example, in patients undergoing ECT for unipolar or bipolar major depression, the results one week after treatment demonstrated relatively large improvements in EQ-5D index (mean difference 0.40), EQ VAS (mean difference 34), and when combining the dimensions in a PCHC analysis (73% improved).

Importantly for the interpretation of EQ-5D results, the thesis revealed novel findings regarding patients' thoughts when reporting and valuing their current health using EQ-5D-5L, EQ VAS, and TTO. One interesting finding was the shift in time perspective when responding to different questionnaires: from considering current or past health when responding to EQ-5D and EQ VAS, to considering the future when responding to the TTO question. However, it is challenging to draw conclusions based on this single study regarding whether this is a general finding for the TTO valuation method or a specific finding for patients with a chronic condition. If expectations on future health are clearly integrated into valuations of all currently experienced health states, the characteristics of the participants and the conditions (e.g., the expected progress of disease) may greatly impact the health state values. By contrast, a previous study exploring the influence of different responder characteristics found no association between the expectations of future health and TTO valuations (90).

Overall, it is challenging to draw conclusions regarding the extent to which the findings from Study II are transferable to other TTO designs. The perspective of valuations (i.e. currently experienced health or hypothetical health states) is only one of several sources of variation between studies. Examples of other methodological choices include the time frame, the phrasing of the question(s), the elicitation procedure, the mode of administration, and the use of visual aids (40, 41, 91). Furthermore, it should be noted that the open-ended question neither had iterations to assist the participant to find a point of indifference, nor presented the possibility to value health states as being worse than dead. Still, some similarities with previous findings regarding valuations of described health states should be noted. For example, a common finding was the use of both own experiences and imagined health states for reference (92). Furthermore, the finding that having children and family impacted on the willingness to trade life years has been observed in previous studies as well (90, 92).

Moreover, an important consideration in this discussion, perhaps in particular for the use of TTO as a direct method for valuing currently experienced health, is that several interview participants questioned what purpose this question could serve in health care. In previous literature, some concerns have been raised regarding ethical considerations of experience-based valuations, for example if it is possible to justify that patients in severe health conditions are asked to imagine being dead or in full health (93). Interestingly, our findings showed that “full health” provoked more thoughts and feelings than “dead”. One of the two participants who refused to answer the question said that the question awakened only feelings of hopelessness. In addition to that the TTO method should be accepted by the respondents, another consideration is to what extent this question can be broadly used across different patient populations.

6.3 IMPLICATIONS FOR POLICY, PRACTICE, AND RESEARCH

The findings from this thesis can be discussed further in terms of their implications for policy, practice, and research. The overview of EQ-5D in the NQRs represents an important contribution to the relatively scarce literature showing real-world examples of how PROMs are implemented in routine health care and what purposes they address. Furthermore, this case highlights the need for improving the use of routinely collected PROMs data and may encourage collaborations and/or initiatives for making use of data based on the examples provided in the overview of EQ-5D in the Swedish NQRs.

The findings from the interview study highlight the importance of using both quantitative and qualitative assessments for understanding the responses to PROMs and for capturing the patient perspective. One learning outcome from the interview study, which may be transferable to future implementations of PROMs, was the desire expressed by patients to better understand the purpose for which PROMs data were collected. Overall, the findings from the qualitative interviews may contribute to a better understanding of what underlies the responses to both the EQ-5D-5L descriptive system, EQ VAS, and to same extent also the EQ-5D index, which is based on the TTO valuation method. Concerning the direct and indirect methods for health state valuation, the variations in interpretations (e.g. time perspectives and reference points) call into question their interchangeable use in health economic models and in decision-making.

Several implications may be discussed in relation to the findings from the two register-based studies. First, the study examining the measurement properties of the two EQ-5D versions indicated that there are advantages to using EQ-5D-5L over EQ-5D-3L specifically for patients with a major amputation of a lower limb. As a consequence of improved performance in terms of informativity and validity, the change from EQ-5D-3L to EQ-5D-5L in SwedeAmp has likely contributed to more accurate and detailed information about the severity of problems experienced by patients. This finding may also be transferable to other patient populations experiencing similar problems related to the musculoskeletal system. Second, relatively large improvements in HRQoL were observed for patients undergoing ECT for unipolar or bipolar major depression, yet there was no indication of a robust statistical relationship between pulse width at first ECT session and HRQoL after ECT. Still, there may be consequences not captured by the selected outcome measure or by the study design. Other treatment parameters could be further explored to optimize the treatment effects for patients treated with ECT.

Some findings from this thesis may also be used to inform future initiatives for implementing EQ-5D in routine health care. Although EQ-5D is widely adopted in health economic evaluations, the studies and examples provided in this thesis revealed several areas of use beyond that purpose. The added value of introducing EQ-5D or other generic measures in the NQRs may be to capture the wider perspective of the impact on patients' HRQoL, as shown in Study III and Study IV. Moreover, there are several opportunities associated with EQ-5D data already collected. A specific suggestion for further research is to explore the measurement properties of different PROMs by taking advantage of the simultaneous administration of several measures. As previously mentioned, satisfactory measurement properties are key considerations for making meaningful interpretations and making use of PROMs data. In addition, the relatively large samples of patients included in the registries enable other analyses of routinely collected data, including reference data (e.g. for populating health economic models) for defined patient populations or associated with specific treatments, side effects, or by clinically relevant subgroups.

Making use of PROMs routinely collected in health care is a research area still under development, and several challenges remain. By design, this thesis has revealed if, and to what extent, routinely collected EQ-5D data have been used. Nevertheless, it provides limited guidance regarding how, when, and why the implementation of EQ-5D and other PROMs works as intended. Based on the findings from this thesis, several recommendations for future research can be made. Most importantly, more research is needed to better understand motives and requirements for making use of data. For example, future studies could use the case of EQ-5D in the Swedish NQRs to examine opportunities for and barriers to making use of data to improve health care, and ultimately, health outcomes. Further, although Study II contributed to increased understanding of patients' thoughts when responding to the questionnaires, future studies could further explore ways in which the use of routinely collected EQ-5D data can be useful to patients.

6.4 METHODOLOGICAL CONSIDERATIONS

In addition to the study limitations addressed in the manuscripts, a few additional methodological considerations associated with the case and the specific studies should be addressed.

6.4.1 The case of EQ-5D in the Swedish national quality registries

There are several strengths to the case adopted for examining the use of PROMs collected in routine health care. The NQRs enable structured data collection in health care settings across Sweden, and EQ-5D is one of few PROMs widely implemented in several registries. The implementation of PROMs in clinical practice is not consistent across groups and settings. Rather than studying a specific group or setting, this case enabled an assessment of the ways in which routine collection and use of PROMs were implemented across several patient populations and conditions. In addition, there are benefits to the use of real-world outcome data from the registries for answering specific research questions important to the overall aim of contributing to the knowledge about the use of PROMs in routine health care.

Nevertheless, the case is also associated with some limitations. Although there are several benefits to using EQ-5D (e.g, for comparison across groups, assessment of interventions, and health economic evaluations), other PROMs or other outcome measures may be more suitable for certain purposes (e.g. monitoring health or treatment effects for the individual patient). It is likely that registries have included different PROMs for different purposes, for example using disease-specific measures for capturing relevant changes in the health of individuals, and generic measures for meso- and macro-level analyses and comparisons.

Study I was essential for understanding the current practices in the collection, presentation, and use of EQ-5D data in the Swedish NQRs. It should be noted that the template used for data collection was developed based on previous experiences and discussions among the authors and collaborators. Thus, the results regarding different presentations and use of data may have been influenced by the examples provided in the template. Further, as the template did not provide any definitions, the reporting was based on the definitions used by the respondents. In case of future studies, it may be beneficial to apply a theoretical framework as a basis for developing examples and to have clear definitions of the categories representing different types of use of data in the correspondence with the registries. Furthermore, a restriction to having the registry organizations as the primary sources of information is that the findings regarding the use of data for follow-up, quality improvement, and decision-making may not be exhaustive. For example, the health care providers may have used EQ-5D data for purposes that the registry representatives were unaware of.

6.4.2 Primary data collection: Study II

Several methodological considerations relate to the sampling and data collection in the qualitative interview study. It should be noted that the patient population chosen for the interviews was not included in any of the registries with ongoing collection and use of

EQ-5D data. If the aim of this study had been to explore the patient perspective of making use of EQ-5D data, recruitment of patients included in one of the NQRs with ongoing EQ-5D collection could have been a more suitable strategy. Nevertheless, for the research question explored in this study, the chosen approach may have better reflected patients' initial reactions to the questionnaires.

As mentioned in the publication, a challenge with the chosen interview technique is that it depends on the interviewee's ability to articulate his/her thoughts. In addition, the interview participants were not presented with details regarding a specific purpose or setting in which the health assessments could be distributed. Speculatively, the finding demonstrating that participants reflected on the purpose of measurement might have been less clear if they had received more detailed information prior to or during the interviews.

Lastly, several actions were taken to enable assessments of the credibility of study findings (94). For example, the publication included descriptions of the study context, the sampling and recruitment strategy, background characteristics of the study participants, quotes from the interviews, and an overview and examples of the data analysis. In addition, some details about myself as a researcher were provided. Still, it should be noted that the interviewer plays an active role in the research process and perspectives, for example by having certain characteristics, experiences, and preconceptions of the specific research question (95). As a tool for reflection, a personal logbook was used for taking notes regarding ideas and impressions after each interview, and regarding the need for additional sampling.

6.4.3 Secondary data collection: Study III and Study IV

Although there are many potential benefits from assessing real-world data from registries, it is associated with certain limitations. The variability of observational data reflects real-world practices, which is desirable when describing current processes and outcomes, and for making decisions related to health care. However, this variability of the data, for example in terms of time intervals, missing data, and changes in treatments and risk factors, creates certain challenges for research or evaluative purposes (28). The information available from the NQRs made it possible to examine and adjust for certain differences in demographic and clinical characteristics. Furthermore, the reason for exploring IPWRA in Study IV was to adjust for the possible imbalances between the three treatment groups.

Still, some analyses based on registry data may be challenging. In general, the exact reasons for missing data in registries are unknown. Especially in Study III, we discussed whether analyses of feasibility could be performed using registry data. In the case of SwedeAmp, the registry had included a variable indicating the reason for not responding to EQ-5D, which enabled analyses of feasibility. Furthermore, a challenge related to the use of a retrospective study design was that the measurement properties of the EQ-5D versions could only be examined in separate subsamples as the two versions had been administered during separate time periods. Given this design, the research team have had discussions regarding the extent to which the analyses reflected differences in measurements or differences between subsamples. At very least, the use of several measures and multiple tests was considered to increase the likelihood of identifying advantages and disadvantages of the two EQ-5D versions. In addition, the analyses of convergent validity were less influenced by the use of separate subsamples, as the correlation between two measures was examined at the same time-points for the same individuals.

Although the use of registries has the advantage of covering samples that are more representative of the target population, the inclusion criteria of having responded to EQ-5D may have contributed to sampling of a certain subgroup. For example, the patients included in the study samples may have had better health than those who were unable to respond to PROMs at baseline or at follow-up. Furthermore, although the variables are common for all units that contribute with data to the registry, registration is not mandatory. There are several sources of both measured and unmeasured variation, for example in terms of the exact time for follow-up (reflecting variations in patient's actual health care visits) and possibly also in the instructions and assistance provided when reporting one's own health using PROMs.

Lastly, some methodological choices concerning the EQ-5D instruments should be addressed. Both the descriptive system and the value sets play a role in the sensitivity and discriminatory power of EQ-5D (57). The Swedish version of EQ-5D and value sets for UK were applied in all studies. The value sets were chosen to facilitate comparison between EQ-5D versions (Study III) and to facilitate international comparison of results (Study IV).

7 CONCLUSIONS

Through the use of a specific case, the studies included in this thesis have examined several aspects important in the use of PROMs collected in routine health care. The thesis has presented findings that contribute to the understanding of the use of PROMs data routinely collected in the Swedish NQRs, considerations in the choice of EQ-5D version in specific populations, and ways in which EQ-5D data may be used to assess specific interventions. The findings demonstrated that EQ-5D data, often in combination with disease-specific measures, were collected across a variety of patient populations and interventions covered by the Swedish NQRs. EQ-5D was most frequently administered in registries targeting conditions related to the musculoskeletal system. For patients with a major amputation of a lower limb, the results indicated that there were several advantages to the use of EQ-5D-5L over the use of EQ-5D-3L. Furthermore, the thesis revealed real-world outcomes after ECT in patients with unipolar or bipolar depression. Routinely collected EQ-5D data were analyzed with the intention to explore ways of optimizing the treatment effect based on a specific treatment parameter, yet the results showed no robust associations between pulse width and HRQoL after ECT.

Moreover, the thesis has revealed findings related to the interpretations and thoughts behind patients' responses to EQ-5D-5L, EQ VAS, and TTO. The findings from the interviews revealed some variation in the use of time perspectives, comparators and reference points, and the aspects considered when completing the three assessments. Furthermore, some participants expressed a desire to explain the reasoning behind their responses in a dialogue with their health care provider. This qualitative assessment contributes to the understanding of EQ-5D results and the use of EQ-5D data, particularly in patients with type 1 diabetes or other chronic conditions.

The case of EQ-5D data in the Swedish NQRs has demonstrated several examples of ways in which routinely collected PROMs data may be used, with the ultimate intention to improve health care and patient health. While several examples of purposes for which EQ-5D data have been used were found, including assessment of interventions, health economic studies, quality indicators, benchmarking, quality improvement, and use at the individual patient level, a relatively large share of the NQRs reported that collected data were not being used for any of these purposes. Importantly, these findings highlight the fact that routine data collection alone is insufficient for achieving the potential benefits of using PROMs in health care. For future implementations of PROMs in routine health care, it may be useful to further explore how routinely collected PROMs data could be of most use to patients and other stakeholders, and to further explore prerequisites for making use of routinely collected PROMs data at different levels of the health care system.

8 ACKNOWLEDGEMENTS

The years I have spent as a doctoral student have involved more experiences than I can count. I feel grateful to everyone who has shared these experiences with me, who has supported me in different ways, and who has made this journey possible. Thank you!

First and foremost, I would like to express my gratitude to everyone who participated in the studies included in this thesis. To all the study participants who made this research possible – thank you! *Jag vill rikta ett stort tack till alla studiedeltagare – tack för er medverkan!*

To my main supervisor, Niklas Zethraeus. Thank you for encouraging me to lead my own learning process while still always being just a phone call away. You always seek to find solutions, which has helped me move forward many times. Thank you for believing in me and for the support you have given me when I needed it the most.

To my supervisor, Emelie Heintz. I don’t think this thesis would have been possible without your incredible support. Thank you for your leadership, encouragement, enthusiasm, and your genuine interest in the work that we do. We have shared many fun and challenging experiences together, and I have learnt so much from working with you. I am truly grateful for our close collaboration and our friendship.

To Evalill Nilsson, for the many ways you have supported me during my time as a doctoral student. Thank you for always being kind and attentive, for your genuine interest, and for always sharing your knowledge, time, and ideas so generously. You inspire me!

To my mentor Charlotta Dahlborg, thank you for our fun conversations, and your valuable advice and encouragement. I have appreciated getting a break from my research project to reflect on the bigger picture of my personal and professional development – thank you!

To Bas Janssen, Helle Mølsted Alvesson, Kristina Burström, Kerstin Hagberg, Gouke Bonsel, Seher Korkmaz, Axel Nordenskjöld, and Jeff Johnson for our collaborations in the studies included in this thesis. Thank you for generously sharing your expertise and for our many valuable discussions. Helle, thank you for patiently guiding me through a new research field and for the support and encouragement – I have learnt so much from working with you.

To all my amazing colleagues at QRC Stockholm. I would like to express my gratitude to Seher Korkmaz, Susanna Lagersten, Inger Rising, Lillemor Bergström, Staffan Lindblad – thank you for supporting me in different ways – I am inspired by the leadership you have shown. I would also like to thank the members of the EQ-5D project group – Kerstin Ankargård, Hans Lindqvist, Gunilla Jakobsson Ekman, Thomas Emilsson, Lena Söderqvist, Helena Tjärnberg – I have learnt so much and really enjoyed working with you!

To Jeff Johnson and Fatima Al Sayah, for giving me the wonderful experience of working with you and the APERSU team in Edmonton, Canada, for a few months. A special thanks to Sherry Lydynuik, Nathan McClure, Lilla Roy, Xuejing (Jennifer) Jin, Donna Pressick – I am grateful for getting to know you and for the fun memories I have from this time.

To Mimmi Åström and Vibeke Sparring, for our friendship, och för vår klubb. I am grateful for all the fun conversations we have had, and that you have never avoided the difficult ones. Mimmi, I find it hard to explain in words how important your friendship has been to me. From day one, you have been a true friend in both good and bad times. Thank you for laughing and crying with me. Vibeke, for always being there and for your fantastic sense of humor (in any situation, really). Your friendship and support have been important to me, both as a doctoral student and as Olivia.

To Emma Granström, for our friendship, conversations about life, our writing retreats, and for sharing this whole experience. I always feel better after we have talked.

To Filip Gedin and Max Kleijberg, I appreciate your friendship, support, and ability to laugh in every situation – all this has been important to me. Thank you for making me laugh, for our conversations, for the preparations for plan B, and for the endless sharing of PhD memes.

To Matilda Hagman, for being a kind, fun, and caring friend. Everything is more fun when you are around – some memories from the conversations we have had still make me laugh.

To Fanny Goude, I appreciate the honest and supportive conversations we have had when writing our kappas. We made it!

To all current and former fellow doctoral students – I am grateful for the small talk and conversations in the coffee corner, for sharing experiences, after-works and fun memories – Mimmi Åström, Emma Granström, Filip Gedin, Max Kleijberg, Fanny Goude, Marie Dahlberg, Cecilia Dahlgren, Sofia Sveréus, Sofi Varg, Anne Leppänen, David Ebbevi, Charlotte Klinga, Agnes Elmberger, Linda Sturesson Stabel, Sara Ingvarsson, George Keel, Mairi Savage, Rafiq Muhammad, Therese Johansson, Shuang Hao, Sara Korlén, Emma Söreskog, Chantelle Murley, Fitsum Sebsibe Teni. To the members of *Kappa-gruppen*, thank you for creating a friendly and supportive environment.

To all my colleagues at LIME for creating a friendly environment – a special thanks to Carl Johan Sundberg, Therese Wahlström, Clas Rehnberg, Henna Hasson, Peter Lindgren, Ulrika Schüldt Håård, Carolina Wannheden, Sara Tolf, Monica Nyström, Per Palmgren, Liisa Olsson, Maria Appelgren, Ludvig Andersson, Erik Attoff, Ludwig Berglund, and to Daniel Olsson, Matilda Hagman, Mesfin Kassaye Tessma, and Eva Hagel for being willing to discuss statistics whenever I have knocked on your door.

To my friends and family ♥

Mina fina vänner utspridda i landet och i världen – Charlotte Pihl, Lina Wahlström, Johanna Hillgren, Lova Graaf, Nathan McClure, Denise Stolt.

Min svärfamilj – Laszlo, Kerstin & Benny, Ronny, Cidde & Nina, David & Therese, Lin & Niklas, och alla kids! Ett särskilt tack till Kerstin för glada hejarop och choklad som kommit med posten. Ett särskilt tack också till Ronny, för all den omtanke du har visat.

Min familj ♥ Jag älskar er

Mina älskade storebröder, Axel och Fredrik. Jag är glad för det särskilda band vi delar, som jag upplever är lika starkt oavsett hur lång tid det går mellan gångerna vi ses. Och likaså Johannes, jag blir alltid lika glad av att träffa dig. Mamma, tack för att du har visat omtanke och stöttat mig, särskilt när jag har kämpat som mest med avhandlingen.

Pappa & Annie, för stödet och tryggheten ni ger mig, och för att ni är himla härliga. Jag är så glad över de fina minnena jag har från vår resa i Kanada. Det har varit ovärderligt att få sitta i er lilla röda stuga i Röjeråsen och skriva klart avhandlingen – tack! Annie, jag är så glad över att ha dig i mitt liv. Pappa, tack för att du alltid finns där, för att du alltid vill förstå, och för att du alltid har uppmuntrat mig att göra mina egna val.

♥ Thomas, för all kärlek, glädje, och stöd under våra år tillsammans. Du är min bästa vän och min största förebild i ditt sätt att tänka om, tänka nytt, och i ditt driv att fortsätta utvecklas. Jag är så tacksam över att du alltid tror att jag klarar av saker – stora som små. Och för att du alltid, alltid hänger på när vi behöver en danspaus.

9 REFERENCES

1. Black N, Campbell F, Gürol-Urganci I. Understanding health services. Second edition. London, UK: McGraw-Hill Education: Open University Press; 2017.
2. Jacobsson Ekman G, Lindahl B, Nordin AE (eds). National Quality Registries in Swedish Health Care. Stockholm: Karolinska University Press; 2016.
3. Drummond M. Methods for the economic evaluation of health care programmes. Fourth edition. Oxford, UK; New York, NY, USA: Oxford University Press; 2015.
4. Appleby J, Devlin NJ, Parkin DW. Using patient reported outcomes to improve health care. Chichester, West Sussex, UK ; Hoboken, NJ, USA: John Wiley & Sons Inc.; 2016.
5. Bartlett SJ, Ahmed S. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 1: introduction. J Clin Epidemiol. 2017;89:114-8.
6. Ciani O, Federici CB. Value Lies in the Eye of the Patients: The Why, What, and How of Patient-reported Outcomes Measures. Clin Ther. 2020;42(1):25-33.
7. Higginson IJ, Carr AJ. Measuring quality of life - Using quality of life measures in the clinical setting. BMJ. 2001;322(7297):1297-300.
8. WHO. Constitution of the World Health Organization 1948 [cited 2021-04-07]. Available from: https://www.who.int/about/who-we-are/constitution.
9. Fayers PM, Machin D. Quality of life : the assessment, analysis, and reporting of patient-reported outcomes. Third edition. Chichester, West Sussex, UK ; Hoboken, NJ, USA: John Wiley & Sons Inc.; 2016.
10. U.S. Department of Health and Human Services Food and Drug Administration (FDA). Guidance for Industry: Patient-reported outcome measures: Use in medical product development to support labeling claims 2009 [cited 2021-04-07]. Available from: https://www.fda.gov/media/77832/download.
11. Kane RL, Radosevich DM. Conducting health outcomes research. Sudbury, Mass, USA: Jones and Bartlett Publishers; 2011.
12. Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the Swedish National Quality Registers. J Intern Med. 2016;279(2):141-53.
13. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.
14. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.
15. Krawczyk M, Sawatzky R, Schick-Makaroff K, Stajduhar K, Ohlen J, Reimer-Kirkham S, et al. Micro-Meso-Macro Practice Tensions in Using Patient-Reported Outcome and Experience Measures in Hospital Palliative Care. Qual Health Res. 2019;29(4):510-21.
16. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res. 2009;18(1):115-23.
17. Kwon JY, Thorne S, Sawatzky R. Interpretation and use of patient-reported outcome measures through a philosophical lens. Qual Life Res. 2019;28(3):629-36.
18. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ. 2015;350:g7818.

19. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res. 2008;17(2):179-93.

20. Van der Wees PJ, Nijhuis-Van der Sanden MWG, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the Use of Patient-Reported Outcomes for Both Clinical Practice and Performance Measurement: Views of Experts from 3 Countries. Milbank Q. 2014;92(4):754-75.

21. Calvert M, Thwaites R, Kyte D, Devlin N. Putting patient-reported outcomes on the 'Big Data Road Map'. J R Soc Med. 2015;108(8):299-303.

22. Greenhalgh J, Dalkin S, Gibbons E, Wright J, Valderas JM, Meads D, et al. How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. J Health Serv Res Po. 2018;23(1):57-65.

23. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract. 2006;12(5):559-68.

24. Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. Qual Life Res. 2013;22(9):2265-78.

25. Devlin N, Appleby J. Getting the most out of PROMs - Putting health outcomes at the heart of NHS decision-making. London, UK: The King's Fund; 2010.

26. Kyte D, Cockwell P, Lencioni M, Skrybant M, Hildebrand MV, Price G, et al. Reflections on the national patient-reported outcome measures (PROMs) programme: Where do we go from here? J R Soc Med. 2016;109(12):441-5.

27. McKenna SP, Heaney A, Wilburn J. Measurement of patient-reported outcomes. 2: Are current measures failing us? J Med Econ. 2018:1.

28. Katkade VB, Sanders KN, Zou KH. Real world data: an opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. J Multidiscip Healthc. 2018;11:295-304.

29. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf. 2014;23(6):508-18.

30. Calvert M, Kyte D, Price G, Valderas JM, Hjollund NH. Maximising the impact of patient reported outcome assessment for patients and society. BMJ. 2019;364.

31. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? Soc Sci Med. 2005;60(4):833-43.

32. Greenhalgh J, Pawson R, Wright J, Black N, Valderas JM, Meads D, et al. Functionality and feedback: a protocol for a realist synthesis of the collation, interpretation and utilisation of PROMs data to improve patient care. BMJ Open. 2014;4(7):e005601.

33. Janssen MF, Birnie E, Bonsel GJ. Evaluating the discriminatory power of EQ-5D, HUI2 and HUI3 in a US general population survey using Shannon's indices. Qual Life Res. 2007;16(5):895-904.

34. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002;21(2):271-92.

35. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002;40(2):113-28.
36. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. Eur J Health Econ. 2018;19(4):557-70.
37. Feng YS, Kohlmann T, Janssen MF, Buchholz I. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. Qual Life Res. 2020.
38. Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Second edition. Oxford, UK: Oxford University Press; 2017.
39. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv Res. 1972;7(2):118-33.
40. Attema AE, Brouwer WBF. In search of a preferred preference elicitation method: A test of the internal consistency of choice and matching tasks. J Econ Psychol. 2013;39:126-40.
41. Hao S, Heintz E, Helgesson G, Langenskiold S, Chen J, Burstrom K. Influence of elicitation procedure and phrasing on health state valuations in experience-based time trade-off tasks among diabetes patients in China. Qual Life Res. 2020;29(1):289-301.
42. Dolan P. Developing methods that really do value the 'Q' in the QALY. Health Econ Policy L. 2008;3(1):69-77.
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(6):881-91.
44. 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.
45. EunetHta Joint Action WPS, Heintz E, Gerber-Grote A, Ghabri S, Hamers FF, Rupel VP, et al. Is There a European View on Health Economic Evaluations? Results from a Synopsis of Methodological Guidelines Used in the EUnetHTA Partner Countries. Pharmacoeconomics. 2016;34(1):59-76.
46. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013 [cited 2021-04-30]. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#measuring-and-valuing-health-effects.
47. The Dental and Pharmaceutical Benefits Agency. Tandvårds- och läkemedelsförmånsverkets allmänna råd (TLVAR 2017:1) 2017 [cited 2021-04-30]. Available from: https://www.tlv.se/om-oss/om-tlv/regelverk/allmanna-rad.html.
48. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. Appl Health Econ Health Policy. 2017;15(2):127-37.
49. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
50. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15(5):708-15.

51. EuroQol Research Foundation. EQ-5D-3L User Guide 2018 [cited 2020-10-25]. Available from: https://euroqol.org/publications/user-guides/.

52. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.

53. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goni JM. Overview, Update, and Lessons Learned From the International EQ-5D-5L Valuation Work: Version 2 of the EQ-5D-5L Valuation Protocol. Value Health. 2019;22(1):23-30.

54. Burstrom K, Sun S, Gerdtham UG, Henriksson M, Johannesson M, Levin LA, et al. Swedish experience-based value sets for EQ-5D health states. Qual Life Res. 2014;23(2):431-42.

55. Burstrom K, Teni FS, Gerdtham UG, Leidl R, Helgesson G, Rolfson O, et al. Experience-Based Swedish TTO and VAS Value Sets for EQ-5D-5L Health States. Pharmacoeconomics. 2020.

56. Buchholz I, Janssen MF, Kohlmann T, Feng YS. A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. Pharmacoeconomics. 2018;36(6):645-61.

57. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L Better Than EQ-5D-3L? A Head-to-Head Comparison of Descriptive Systems and Value Sets from Seven Countries. Pharmacoeconomics. 2018;36(6):675-97.

58. Greene ME, Rader KA, Garellick G, Malchau H, Freiberg AA, Rolfson O. The EQ-5D-5L Improves on the EQ-5D-3L for Health-related Quality-of-life Assessment in Patients Undergoing Total Hip Arthroplasty. Clin Orthop Relat Res. 2015;473(11):3383-90.

59. Conner-Spady BL, Marshall DA, Bohm E, Dunbar MJ, Loucks L, Al Khudairy A, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. Qual Life Res. 2015;24(7):1775-84.

60. Buchholz I, Thielker K, Feng YS, Kupatz P, Kohlmann T. Measuring changes in health over time using the EQ-5D 3L and 5L: a head-to-head comparison of measurement properties and sensitivity to change in a German inpatient rehabilitation sample. Qual Life Res. 2015;24(4):829-35.

61. Golicki D, Niewada M, Buczek J, Karlinska A, Kobayashi A, Janssen MF, et al. Validity of EQ-5D-5L in stroke. Qual Life Res. 2015;24(4):845-50.

62. Brazier J. Is the EQ-5D fit for purpose in mental health? Br J Psychiatry. 2010;197(5):348-9.

63. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess. 2014;18(34).

64. Emilsson L, Lindahl B, Koster M, Lambe M, Ludvigsson JF. Review of 103 Swedish Healthcare Quality Registries. J Intern Med. 2015;277(1):94-136.

65. National Quality Registries; Swedish Association of Local Authorities and Regions. About the National Quality Registries 2021 [cited 2021-04-30]. Available from: https://kvalitetsregister.se/englishpages/aboutqualityregistries.2422.html.

66. Anell A, Glenngård A, Merkur S. Sweden: Health system review. Health Systems in Transition; 2012.

67. Kamrad I, Soderberg B, Orneholm H, Hagberg K. SwedeAmp-the Swedish Amputation and Prosthetics Registry: 8-year data on 5762 patients with lower limb amputation show sex differences in amputation level and in patient-reported outcome. Acta Orthop. 2020:1-7.

68. The Swedish Amputation and Prosthetics Registry. Årsrapport 2019 - Amputations- & protesregister för nedre extremiteten (SwedeAmp) [In English: Annual report 2019 - The Swedish Amputations and Prosthetics Registry (SwedeAmp)] [cited 2021-04-30]. Available from: https://swedeamp.com/wp-content/uploads/2020/08/Arsrapport-2019_FINAL_0820.pdf.

69. Nordanskog P, Hulten M, Landen M, Lundberg J, von Knorring L, Nordenskjold A. Electroconvulsive Therapy in Sweden 2013: Data From the National Quality Register for ECT. J ECT. 2015;31(4):263-7.

70. Elvin T, Nordenskjöld A. Kvalitetsregister ECT Årsrapport 2019 [In English: The Swedish National Quality Registry for ECT Annual Report 2019] [cited 2020-10-25]. Available from: https://registercentrum.blob.core.windows.net/ect/r/-rsrapport2019_ECT-utan_PARdocx-rklzVIXM2U.pdf.

71. EuroQol Research Foundation. About EQ-5D [cited 2021-04-29]. Available from: https://euroqol.org/eq-5d-instruments/.

72. Devlin N, Parkin D, Janssen B. Methods for Analysing and Reporting EQ-5D Data: Springer; 2020 [cited 2021-04-29]. Available from: https://link.springer.com/book/10.1007%2F978-3-030-47622-9.

73. The National Board of Health and Welfare. Handbok för utveckling av indikatorer - för god vård och omsorg [In English: Handbook for developing indicators - for good care and nursing] Stockholm: The National Board of Health and Welfare; 2017 [cited 2021-05-12]. Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/handbocker/2020-8-6877.pdf.

74. Thonon F, Watson J, Saghatchian M. Benchmarking facilities providing care: An international overview of initiatives. SAGE Open Med. 2015;3:2050312115601692.

75. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. J Adv Nurs. 2003;42(1):57-63.

76. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health. 2000;23(4):334-40.

77. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. Qual Health Res. 2016;26(13):1753-60.

78. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3:77-101.

79. Franchignoni F, Orlandini D, Ferriero G, Moscato TA. Reliability, validity, and responsiveness of the locomotor capabilities index in adults with lower-limb amputation undergoing prosthetic training. Arch Phys Med Rehabil. 2004;85(5):743-8.

80. Hagberg K, Branemark R, Hagg O. Questionnaire for Persons with a Transfemoral Amputation (Q-TFA): initial validity and reliability of a new outcome measure. J Rehabil Res Dev. 2004;41(5):695-706.

81. Juniper E, Gordon H, Roman J. How to develop and validate a new health-related quality of life instrument. In: Spilker B, editor. Quality of Life and Pharmacoeconomics in Clinical Trials. Second edition. Philadelphia, PA, USA ; New York, NY, USA: Lippincott Williams & Wilkins; 1996.

82. Ernstsson O, Zethraeus N, Heintz E, Nordenskjöld A, Korkmaz S, Johnson J. Association between pulse width and health-related quality of life following electroconvulsive therapy (Study Protocol): Open Science Framework; 2020 [cited 2020-12-22]. Available from: https://osf.io/vxyag/.

83. Luo N, Johnson J, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. Med Care. 2010;48(4):365-71.

84. Vickers AJ, Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. BMJ. 2001;323(7321):1123-4.

85. Pullenayegum EM, Tarride JE, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of Health Utility Data When Some Subjects Attain the Upper Bound of 1: Are Tobit and CLAD Models Appropriate? Value Health. 2010;13(4):487-94.

86. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivar Behav Res. 2011;46(3):399-424.

87. The Swedish Association of Local Authorities and Regions (SALAR). Om Vården i Siffror [In English: About Health care in Numbers] 2021 [cited 2021-05-12]. Available from: https://vardenisiffror.se/om-varden-i-siffror.

88. Franchignoni F, Traballesi M, Monticone M, Giordano A, Brunelli S, Ferriero G. Sensitivity to change and minimal clinically important difference of the Locomotor Capabilities Index-5 in people with lower limb amputation undergoing prosthetic training. Ann Phys Rehabil Med. 2019;62(3):137-41.

89. Connell J, O'Cathain A, Brazier J. Measuring quality of life in mental health: are we asking the right questions? Soc Sci Med. 2014;120:12-20.

90. van Nooten FE, van Exel NJ, Koolman X, Brouwer WB. "Married with children" the influence of significant others in TTO exercises. Health Qual Life Outcomes. 2015;13:94.

91. Arnesen T, Trommald M. Are QALYs based on time trade-off comparable?--A systematic review of TTO methodologies. Health Econ. 2005;14(1):39-53.

92. Karimi M, Brazier J, Paisley S. How do individuals value health states? A qualitative investigation. Social Science & Medicine. 2017;172:80-8.

93. Brazier JE, Rowen D, Lloyd A, Karimi M. Future Directions in Valuing Benefits for Estimating QALYs: Is Time Up for the EQ-5D? Value Health. 2019;22(1):62-8.

94. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-12.

95. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet. 2001;358(9280):483-8.

10 APPENDICES

10.1 APPENDIX A – STUDY I: TEMPLATE FOR DATA COLLECTION

Name of registry: ***Contact person:*** ***Date for the collection of information*:**	
General information	**Information for the registry**
What patients are included in the registry?	
How is it decided that they will be included in the registry (at a health care visit, at diagnosis, etc.)?	
When is the last registration made in the registry for each patient (24 months after intervention, when the patient dies, etc.)?	
What year did the registry start collecting data?	
How many patients are registered in total, and per year?	
What is the estimated coverage? How many of all possible clinics in Sweden participate in the collection? How many of all eligible patients are given the opportunity to respond to the questionnaires? What is the response frequency?	
What background variables are registered for each patient (e.g., sex, age, geographical area, ethnicity, socioeconomics, marital status, education)?	
What background information is registered regarding the patients' disease?	
What diagnostic tests and treatments are registered for the patients?	
Collection of PROMs data	
For what patients are you collecting PROM, and specifically EQ-5D data?	

Which version of EQ-5D are you using? - Which version (e.g., EQ-5D-3L, EQ-5D-5L)? - Is EQ VAS included?	
What value set are you using to calculate the EQ-5D index?	
What other PROM instruments are you using?	
How many EQ-5D measurements do you have in total? For how many patients have you registered EQ-5D data, per year?	
When did you start collecting EQ-5D data (year, month)?	
Are you measuring once or several times per patient? When are data collected (e.g., pre-post treatment, at all visits, at follow-up at 3, 6, 9 months, or when patients do their own registration at home)?	
Administration of PROMs	
How are data collected (e.g. through a paper questionnaire, web survey, or face-to-face interviews)? If interview, who is doing the interview?	
Are patients responding themselves or are other respondents acting as proxies?	
Use of PROMs for quality improvement and decision-making	
How are PROM data made available to decision makers (patients, clinicians, directors, politicians)? - Individual feedback to the patient (e.g., instantaneous or at follow-up)? - Individual feedback for shared decision making, or decision support to the patient and clinicians? - Aggregated feedback in clinical setting to the patient? - Aggregated feedback in clinical setting to the health care professionals? - Aggregated feedback, other (e.g., webpage or annual report)? - Research (in what publications?) Registry-based randomized controlled trials (R-RCT)?	

How are PROMs data currently analyzed and used? • Quality improvement (examples)? • Guidelines and/or recommendations? • Benchmarking? • Evaluation of treatment/intervention effects? • Health economic evaluations?	

10.2 APPENDIX B – STUDY II: INTERVIEW GUIDE

The semi-structured interview guide was originally in Swedish.

Introductory questions

Can you tell me how long you have had diabetes?
    How did you first notice?
    How long did it take to find out what it was?
Do you have any other illness?

Think aloud: EQ-5D

I'm interested in knowing how you think when you fill out a questionnaire about your health today. Can you please describe your thoughts out loud while you respond to this questionnaire?

Examples of probing questions. Can you tell me more about your thoughts regarding, e.g.:

- *The different response options in the questionnaire*
- *Dimensions that seem more or less important/relevant*
- *The questions combining several aspects, i.e., pain/discomfort, anxiety/depression*

Think aloud: EQ VAS

Can you please describe your thoughts out loud while you respond to this questionnaire?

Examples of probing questions. Can you tell me more about your thoughts regarding, e.g.:

- *Best/worst imaginable health state*
- *What it would take to choose a higher/lower score*
- *What it would take to choose 100 on the scale*

Think aloud: Time trade-off

Can you please describe your thoughts out loud while you respond to this questionnaire?

Examples of probing questions. Can you tell me more about your thoughts regarding, e.g.:

- *How you come up with the number of years in full health that you think is of equal value to living 10 years in your current health state*
- *Full health*
- *Reasons for not wanting to give up any life years*
- *Imagining ten years left to live*

Closing questions

If we had done the same interview a week or a month ago, how would you have answered these questions?

I have asked all my questions. Is there anything you would like to add or elaborate on?

10.3 APPENDIX C – STUDY IV: RESULTS OF ADDITIONAL ANALYSES

See next page.

Appendix C. Results from the inverse-probability-weighted regression adjustment, for examining the association between pulse width (ms) and HRQoL after ECT.

	EQ-5D index			**EQ VAS**		
	Observations, n		IPWRA[a]	Observations, n		IPWRA[a]
	Raw	Weighted	Model 3[b], β (95% CI)	Raw	Weighted	Model 3[b], β (95% CI)
One week after ECT	3,056	3,056		2,994	2,994	
<0.5 ms (POmean)	403	1,022	0.693 (0.663–0.722)	393	1,000	60.188 (57.719–62.657)
0.5 vs <0.5 (ATE)	2,238	1,018	0.013 (-0.019–0.044)	2,194	998	1.140 (-1.487–3.768)
>0.5 vs <0.5 (ATE)	415	1,016	0.009 (-0.031–0.050)	407	996	1.260 (-2.075–4.596)
Six months after ECT	602	602		675	675	
<0.5 ms (POmean)	75	200	0.673 (0.614–0.731)	93	223	61.355 (56.688–66.022)
0.5 vs <0.5 (ATE)	409	201	-0.047 (-0.112–0.019)	464	226	-0.566 (-5.729–4.598)
>0.5 vs <0.5 (ATE)	118	202	-0.004 (-0.082–0.074)	118	226	0.568 (-5.660–6.796)

Note: ATE = average treatment effect; POmean = Potential-outcome means.

[a] Inverse-probability-weighted regression adjustment. Treatment model: multinomial regression estimating the probabilities of treatment assignment by patient characteristics: sex, age, indication (unipolar/bipolar; with/without psychotic features)

[b] Multiple linear regression with robust standard error, examining the relationship between pulse width at first ECT session and HRQoL after ECT. Outcome model adjusted for the IPW and for potential confounding variables in Model 3: HRQoL at baseline, age, sex, unipolar/bipolar depression, with/without psychotic features, and concurrent medications (antidepressants, antipsychotics, benzodiazepines, lithium, antiepileptics)