Western Graduate&PostdoctoralStudies

Western University
Scholarship@Western

Electronic Thesis and Dissertation Repository

8-26-2022 8:00 AM

Validity of Quality of Life Measurement in Economic Evaluations in Total Joint Replacement: Agreement between Western Ontario and McMasters' Osteoarthritis Index (WOMAC) and the EuroQOL 5D (EQ-5D) Utility scores

Sanju Valamparampil, *The University of Western Ontario*

Supervisor: Marsh, Jacquelyn, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Health and Rehabilitation Sciences
© Sanju Valamparampil 2022

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Other Medicine and Health Sciences Commons

Recommended Citation

Valamparampil, Sanju, "Validity of Quality of Life Measurement in Economic Evaluations in Total Joint Replacement: Agreement between Western Ontario and McMasters' Osteoarthritis Index (WOMAC) and the EuroQOL 5D (EQ-5D) Utility scores" (2022). *Electronic Thesis and Dissertation Repository*. 8856.
https://ir.lib.uwo.ca/etd/8856

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

Abstract

Health economic evaluations are commonly conducted through a cost-utility analysis, where health benefits are measured using utility scores. A common utility measure is the European Quality of Life (EQ-5D). Osteoarthritis (OA) research studies commonly use disease-specific quality of life tools such as the Western Ontario and McMaster's Universities Osteoarthritis Index (WOMAC) to derive utility scores, but the validity of this method is unknown. This research aims to determine the agreement between utility scores derived from WOMAC and the EQ-5D surveys among patients who have undergone Total Joint Replacement (TJR). To estimate the agreement, we calculated an intraclass correlation coefficient (ICC) and its 95% confidence interval (CI) and produced Bland Altman plots. Our results indicate good agreement between the two scores, as seen with the ICC value of 0.85, 95% CI (0.82 - 0.87).

Keywords: Total joint replacement; cost-utility analysis; mapping; EQ-5D-5L; WOMAC; Osteoarthritis

Summary for Lay audience

Osteoarthritis is a common chronic condition that affects the function of the joints. With a prevalence rate of 14.2% and projections indicating that the number of Canadians diagnosed with OA is set to increase, the economic burden posed by OA is also set to increase. Canadians rely on the publicly funded healthcare system to meet their healthcare needs. One of the significant issues with publicly funded healthcare programs is that the needs of the public outweigh the available healthcare resources. Therefore, decision-makers need to identify which healthcare needs will provide the greatest benefit to the largest possible majority of the public; since each decision comes with a cost and benefit. One of the methods often employed within healthcare to identify the costs and benefits associated with each of these decisions is the use of health economic evaluations.

With an increasingly aging population and the projected number of patients with OA, the resources required to treat these individuals will increase. As a result, more economic evaluations are being conducted within osteoarthritis research. An essential metric within these evaluations are utility scores. A utility score is a numerical value representing a patient's preference for each health state, with a value of 1 representing perfect health and 0 representing the worst imaginable health state from the patient's perspective. These values can be derived either directly using utility instruments or by converting disease-specific quality of life surveys into utility scores. However, the validity of converting disease-specific scores into utility scores following total joint replacement is unknown.

Therefore, this thesis aims to compare the utility scores between the European Quality of Life (EQ-5D-5L) and the Western Ontario and McMasters Osteoarthritis Index (WOMAC) among patients who have undergone total joint replacement surgery of the hip and or knee joints.

Co-Authorship Statement

The study was conceptualized and designed by my supervisor Dr. Jacquelyn Marsh, with support from my advisory committee members Dr. Edward Vasarhelyi and Dr. Lyndsay Somerville. I was responsible for writing and running analysis on R. Dr. Marsh and Dr. Somerville revised the manuscript assisted in interpreting the results and provided valuable feedback for the final thesis submission.

Acknowledgement

I would also like to thank my family for supporting me throughout the Masters program.

I want to thank my supervisor Dr. Jacquelyn Marsh for supporting me throughout the thesis writing process and helping me navigate through the Masters' program over the last 2 years and for providing me with the opportunity to learn at Western.

I would also like to acknowledge and thank Dr. Lyndsay Somerville, for providing us with feedback on our thesis and for her help in obtaining the data for this thesis.

I would like to acknowledge the immense help provided by Dr. Jacquelyn Marsh, Dr. Lyndsay Somerville, Dr. Edward Vasarheyli for supporting me as my committee members.

Finally I would like to thank all my course instructors for all their support.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

BMI	Body Mass Index
CI	Confidence Interval
CUA	Cost Utility Analysis
EQ-5D-5L	EuroQol- 5D-5L
GLM	General Linear Model
OA	Osteoarthritis
PBM	Preference Based Measure
PROM	Patient report outcome measure
QALYs	Quality adjusted life years
SD	Standard deviation
SEM	Standard Error of Measurement
SG	Standard Gambit
TTO	Time Trade Off
THR	Total Hip Replacement
TJR	Total Joint Replacement
TKR	Total Knee Replacement
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

Chapter 1

1.1Introduction

Osteoarthritis (OA) is a disease that affects the entire joint, involving the cartilage, joint lining, ligaments, and bone. It is characterized by breakdown of the cartilage, bony changes of the joints, deterioration of tendons and ligaments, and inflammation of the synovium or joint lining. (Kraus, et al., 2015;OARSI, 2016). Birthwhistle and colleagues (2015) has estimated the prevalence of OA in Canada stands at 14.2%, with women being more effected by the disease than men (15.6% vs 12.4%). As of 2010, OA-related care accounts for 2.9 billion dollars, and projections indicate that the number of individuals diagnosed with OA is set to increase in the coming years (Sharif et al., 2015; Stats Canada, 2019). With a rapidly aging population, combined with increased constraints on healthcare budgets, there is a clear need for the optimal allocation of these scarce healthcare dollars. An effective way to assess the value of health care programs is through health economic evaluations. Utility measures provide information regarding an individual's preference for health states and can be obtained through two methods. The first is to elicit preferences for health states directly through choice-based exercises such as the Time Trade-Off (TTO) and Standard Gamble (SG) techniques; however, these are time-consuming and burdensome for the patient and require specialized training for administration. Therefore, a common approach is using existing health-related quality of life questionnaires. Prescaled indexes are self-reported questionnaires that provide a utility value for each possible health state. Another method that is increasingly being used is to convert scores from HRQOL instruments into utility scores using algorithms; however, the validity of this method has mixed results and agreement between utility scores from a preference-based utility measure and a converted disease-specific measure among patients following total joint replacement is unknown.

Therefore, the primary objective of this thesis was to assess the agreement between the EQ-5D utility scores and utility scores generated by converting WOMAC scores among patients who have undergone hip and knee replacement.

1.2 Literature review

1.2.1 Osteoarthritis: detection and treatments

Osteoarthritis (OA) is a disorder of the whole joint that is the result of stress and changes to weight bearing joints of the body resulting from the breakdown of cartilage within the joints, inflammation, bone remodelling and loss of function for the joint (OARSI, 2016). It is a form of arthritis that commonly affects more women than men (Hawker, 2019). OA results in increased pain and reduced function of the joints, with the hips, knee, and spine being most affected due to the wear and tear of the cartilage (Xie et al., 2010). Cartilage refers to a soft and flexible connective tissue that sits between bones and helps mitigate friction between the bones in a joint. However, through the continuous use of joints throughout life, cartilages become thinner, resulting in greater friction between the joints, contributing to greater pain and a decrease in function of the joint.

Factors that contribute to the increased prevalence of OA within the Canadian population include age; sex; obesity; genetic predisposition; socio-economic status; high blood pressure; high bone mineral density; physically demanding jobs; injuries to the joint; weakness in muscles surrounding the joint; deformities in the joints; malalignment of the joints and participation in high – performance sports (Allen et al., 2022).

In Canada, 26.8% of adults (18 and over) are obese, and with a rapidly aging population that is projected to account for 23% of the population by 2030, the prevalence of OA is expected to increase. Furthermore, as more adults get diagnosed with OA and as non-operative forms of treatments lose their efficacy in reducing pain and maintaining function for the joints, the demand for total joint replacement surgeries (hip and knee) will increase.

1.2.1.1 Diagnosis

Globally, OA affects 240 million people over the age of 60 (Allen, et al., 2022). Typically, they present with symptoms such as joint pain, limited functions related to the joints and familial history. Physicians can use medical imaging technologies like X-ray machines to diagnose the magnitude of osteoarthritis's effect on the joints. A common radiographic grading technique used to classify patients with OA based on severity was first introduced by Kellgren and Lawrence in 1957, the KL classification for Osteoarthritis (Kellgren & Lawrence, 1957). The KL classification system is divided into grades 0 – 4 (0-None, 1 – doubtful, 2- minimal, 3 – Moderate and 4 – Severe). With each increasing grade, the space between the joints on the x-rays becomes narrower until there is no space between the joints (Grade 4 – severe Osteoarthritis). The visualized space on the x-rays represents the presence of cartilage and the lack of osteophytes.

1.2.1.2 Types of non-operative treatments

Treatments for OA vary depending on the level of severity of the disease. Common forms of treatment for those diagnosed with grade 2 or minimal OA on the KL scale include physical therapy, exercise regimens, injections, and pharmacological interventions such as non-Steroidal Anti-Inflammatory Drugs (NSAIDs). Meta-analyses have shown small to moderate effect sizes in improving pain management for patients with OA where physical therapy and NSAIDs are used in conjunction (Yu & Hunter,2015).

Clinical practice guidelines for non-operative management of OA suggest that exercises aimed at maintaining the functionality of the joints, educational programs for self-management of OA and programs aimed at changing dietary behaviour to reduce obesity are shown to be highly effective. Additionally, pharmaceutical options such as NSAIDs, topical treatments, and opioids have also shown to be effective means for non-operative management of OA (AAOS, 2021; OARSI, 2019). Similar evidence has been presented by Kolanski and colleagues (2020), where the strongest recommendation for non-operative management of OA were exercises such as walking, neuro-muscular training and aquatic exercises; weight loss; Tai Chi; Yoga; the use of canes; Tibo-femoral knee braces; and Patellofemoral knee braces (Kolasinski et al., 2020).

1.2.2 Hip and Knee OA

Globally as of 2022, 240 million patients have been diagnosed with symptomatic OA. Multiple cohort studies have indicated that the prevalence and incidence rates for knee OA remain the highest, followed by hip and hand OA. The prevalence rates have been estimated as high as 14% in the US for hip OA and 7% for knee OA (Allen et al., 2022). Additionally, according to the Canadian Joint Replacement Registry (CJRR), OA accounts for 72.5% and 99.4% of primary hip and knee replacements across Canada. (CIHI, 2020).

1.2.2.1 Total joint replacement (TJR)

Total joint replacement is recommended for individuals with severe OA (grade 4 on the KL scale) or once non-operative interventions have been exhausted. TJR is a surgical procedure that replaces the joints within the body with metal or plastic components. The prosthetic joints aim to reduce pain, improve function, and restore mobility for patients to provide a greater quality of life. In Canada, a total of 63,496 hip replacements and 75,073 knee replacements were performed between 2019-2020, with women accounting for 58% of all surgeries (Canadian Institute for Health Information. Hip and Knee Replacements in Canada: CJRR Annual Statistics Summary, 2018–2019. Ottawa, ON: CIHI; 2020.)**.**

1.3 Economic Burden

In 2010 OA-related care accounted for $2.9 billion in Canada and is expected to increase to $ 7.9 billion by 2031; (Sharif et al., 2015). The projected costs for hospitalizations are $2.9 billion; $ 1.2 billion for outpatient services; $1.2 billion for alternative treatments and out-of-pocket costs; $ 1 billion for drugs; $ 0.7 billion for rehabilitation; and $0.6 billion for treating side effects caused by drugs (Sharif et al., 2015). Additionally, the inpatient cost of performing TKR and THR in Canada as of 2020 was $ 10,500 (Canadian Institute for Health Information. Hip and Knee Replacements in Canada: CJRR Annual Statistics Summary, 2018–2019. Ottawa, ON: CIHI; 2020.). The CJRR 2019-2020 report stated that $1.4 billion was spent on these surgeries. Similarly, between 2015 to 2020, the national average for the percentage of change was 19.3%, indicating the rapid increase in the number of these surgeries being performed within that timeframe.

The economic burden posed by the increased prevalence of OA was examined by Tarride and colleagues (2012) for the province of Ontario by evaluating the results of Ontarians who completed the Canadian Community Health Survey (CCHS). Their study indicated that patients diagnosed with OA had higher hospitalization costs than those without; $2233 vs $1033, respectively (Tarride et al., 2012). The average cost for primary knee replacements in Canada was $9,083 and $9,591 for primary hip replacements for OA; these figures include physician costs (CIHI, 2021). It should be noted that their estimates are only for those diagnosed with OA and do not account for primary replacement of the hip due to hip fractures or any revision surgeries after TJR.

In a systematic review of the economic burden of OA, the direct costs ranged from $ 1442 to $ 21,335 annually in the United States (Xie et al., 2016). Xie and colleagues also noted that the increased burden of cost varied depending on the level of access individuals had to healthcare, but the costs remained high. In contrast, the quality of life for patients remained low (Xie et al., 2016). Similarly, the estimated productivity costs from work loss associated with patients diagnosed with OA in Canada between 2010 to 2031 are $ 12 billion to $ 17.5 billion (Sharif et al., 2017). These estimates provide a clear view of the economic and social burden that OA is expected to bring about with its increasing prevalence and highlights the need for economic evaluations to identify the best strategies to help address the issue.

1.3.1 Economic Evaluation

Health economic evaluations are a research methodology aimed at estimating the additional costs of treatment relative to the benefits achieved compared to the standard of practice. Health economic evaluations allow decision-makers to understand how worthwhile a new treatment might be for patients. With the increasing scarcity of healthcare dollars, the need for more cost-effective options has become more pronounced within the Canadian healthcare system. Multiple forms of health economic evaluations exist depending on the outcome of interest measured. The two main types of economic evaluation in healthcare are cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). CEA is a type of economic evaluation where the measurement of cost is in healthcare dollars while the measurement of effect is in natural units such as the number of life years gained, falls prevented, adverse events, quality of life, weight loss, or any outcome of interest, depending on the clinical area. With CUA, on the other hand, the measurement of cost is in healthcare dollars, while the measurement of effect is in quality-adjusted life years (QALYs) (Drummond et al.,2015). CUA uses utility scores, a preference-based quality of life measure, to calculate quality-adjusted life years.

1.4 Utility Scores

Utility refers to the patient's preference for each health state. Utility scores range between 0 and 1; 0 represents a health state equivalent to death for a patient, while 1 represents a perfect health state and utility values below 0 or negative utility values represent health states that are considered worse than death by patients.

Utility scores are an integral part of CUA for health economic studies as they are used to calculate quality-adjusted life years (QALY)s. QALYs are a summary value that accounts for both the quality and quantity of life that patients spend in a particular health state. CUA reports the additional cost per QALY gained when comparing a new treatment to the current standard of care. QALYs are a useful outcome within health economic evaluations as they allow for the comparison of the cost-effectiveness of various treatments across studies and clinical areas (Whitehead and Ali, 2019).

1.5 Patient-reported Outcome Measures

Patient-reported outcomes measures (PROMS) represent measurements obtained from patients about specific outcomes of interest through self-administered surveys. These measures provide a quick overview of a patient's health from their perspective and can be used to track the progress of the disease by administering the survey multiple times to determine the relative change in the severity of the disease over time or to evaluate the effectiveness of an intervention, such as through pre and post-treatment surveys. PROMs can be divided into; disease-specific, which relate to a specific health condition or region-specific, which assess specific areas of the body; and generic PROMs, which evaluate the overall quality of life.

1.5.1 Disease-Specific Measures

Disease-specific outcome measures are questionnaires that relate to a particular disease or health condition and evaluate the patient's health status in domains specific to the condition of interest. Disease-specific measures provide researchers and clinicians with the ability to understand how far a disease has progressed from a patient’s perspective through domains that are geared towards assessing certain aspects of the patient’s physical health that have been known to be most affected by a disease. While these domains and their descriptions provide a glimpse of how the disease has affected a patient’s daily activity of living, it does not, however, provide reliable information on the overall wellbeing of the individual, nor can the findings from such surveys be applied to a border disease population or a general population. In addition, disease-specific measures are highly sensitive and specific compared to generic measures since they were made to detect those changes that could only be found within a specific patient population. Within the field of hip and knee OA, one of the more established surveys commonly used in clinical research is the Western Ontario and McMaster's Osteoarthritis Index (WOMAC).

1.5.2 WOMAC

The WOMAC was first developed in the 1980s and has become a staple to measure status after osteoarthritis treatment relating to the hip and knee. The questionnaire includes three domains: pain, stiffness, and physical functioning. The three domains with 24 questions are primarily used to assess the impact of OA on a patient's daily life activities (ADL), functionality, gait, overall health, and QOL. Each question is rated on a scale between 0 and 4 (0 – none, 1 - slight, 2 – moderate, 3 – severe, and 4 – extreme), resulting in scores for each domain which are then added together to compute a final total WOMAC score.

A systematic review by McConnel and colleagues (2001) looked at studies that utilized the WOMAC within a population diagnosed with OA of the hip or knee and indicated that WOMAC had good reliability, as the survey displayed high internal consistency with Cronbach Alpha values being above 0.7 across multiple studies. Their review also provided evidence that the survey had good construct and known groups validity. Additionally, their review indicated that the survey was responsive, but the physical domain had the largest effect size, while the pain and stiffness subscales had significantly smaller effect sizes; indicating that the physical sub-scale could easily detect change that was clinically significant within a smaller sample population as opposed to the other subscales (McConnel et al., 2001). Similarly, the physical function subscale has been shown to have very good reliability, good validity, good sensitivity to change and good responsiveness when compared to similar surveys such as the Knee Injury and Osteoarthritis Outcome Score (KOOS) that specifically assess the physical functioning of the hip or knee for OA patients (White & Master, 2016).

Similarly, a systematic review by Lundgren-Nilsson and colleagues (2018) summarized the psychometric properties of PROMs used in OA by assessing their validity, reliability, and feasibility based on the outcome measures in the Rheumatology (OMERACT) filter. Their summary pointed out that the WOMAC had strong validity and reliability with moderate levels of feasibility or ease of use for patients (Lundgren-Nilsson et al., 2018).

1.5.3 Generic surveys

Generic measures are a type of PROM that ask participants to rate their health status or quality of life without references to specific diseases and provide users with a holistic view of health. Unlike disease-specific measures, generic surveys are designed to assess the overall wellbeing of individuals and seek to identify areas of health that are commonly affected in patients regardless of the disease. This means that the results of these surveys are highly generalizable to a broader patient population compared to disease-specific surveys. The survey's ability to be used across multiple disease populations makes these surveys ideal for health economic evaluations that seek to evaluate the costs and benefits of different treatments across various populations.

1.6 Preference-based Measures

Preferences can be measured using two methods; direct elicitation from participants or through the use of pre-scored health status indexes. Direct elicitation typically refers to methods where patients are provided alternative options and are asked to decide on which option they prefer. Direct elicitation methods include visual analog scale (VAS), standard gamble (SG), and time trade-off (TTO) methods. The VAS is a method where patients are asked to rate their health on a scale varying from 0 to 100, 0 to 10 or by simply placing a line on a horizontal line where one end represents the worse possible health and the other end represents the best possible health state. In the case of rating health through numerical values, the higher values indicate better health and lower values indicate low levels of health (Drummond et al., 2015). The SG is another method of direct elicitation where participants are provided two options. Alternative one is taking a treatment with the possibility of going back to complete health and living for a specified number of years or dying immediately. While alternative two results in the patient remaining in the current health state for a specified number of years. The point where the patient preference for alternatives 1 and 2 becomes the same represents the utility value for that health state. Finally, the TTO is another elicitation method where patients are provided two alternatives, like with the SG, however, the amount of time in a health state of perfect health (1) vs at a lower level of health (i) is varied; as the name implies, this method seeks to identify the amount of time patients are willing to give up to live in perfect health or the amount of health patients are willing to give up to survive for a longer period of time. These methods produce utility values for individuals when they complete these exercises through face-to-face interviews (Drummond et al., 2015), which can be burdensome and time-consuming. To overcome this limitation, researchers often use pre-scored surveys which correspond to health state values that have been derived from the general population. This method is called multi-attribute utility theory. Based on this method, many multi-attribute health status classification systems have been devised, such as the Health Utilities Index Mark 2 (HUI2), HUI3, the EQ-5D, and the SF-6D. The EQ-5D, in particular, has been extensively used in studies to determine treatments' effect on a patient's quality of life and has been used extensively in health economic evaluations (Dakin et al., 2018).

1.6.1 EQ-5D-5L

One of the most common preference-based measures is the EuroQol-5d (EQ-5D); a multi-attribute health classification (generic) survey developed in the European Union (EU), designed to examine the QOL through five key areas or attributes; mobility; self-care; activities of daily living (ADL); pain and anxiety. The original EQ-5D model consisted of three levels (EQ-5D-3L), with each attribute providing three response levels; no problem; some problem; unable to perform, extreme pain, or extreme anxiety. Research had indicated that the 3L system had ceiling effects and that the survey could not detect changes in certain populations. As a result, a modified version of the survey was published in 2009, consisting of five levels for each attribute: the EQ-5D-5L survey. This survey proved to be more capable of detecting changes and has been validated across multiple countries such as the UK, Canada, South Korea, Japan, China, and Spain, to name a few (Janssen et al., 2018).

The widespread use of the survey is due to its good psychometric properties and ease of administering it. The EQ-5D-5L consists of 5 questions, one for each attribute and five levels or response options for each attribute, such as no problems, slight problems, moderate problems, severe problems and unable to perform an activity or extreme pain and anxiety. When participants complete the survey, a 5-digit code called the health index value is produced, representing a specific health state. Since there are five attributes and five levels, the survey can identify 3125 (5^5) health states, compared to the 245 (5^3) health states that could be identified in the 3L version of the survey. Each health index value produced has a corresponding utility value that is based on a value set derived using one of the direct elicitation methods mentioned before for each country. The utility scores from all value sets will fall between 0 and 1; 0 represents a health state equivalent to death for participants, while 1 represents a perfect health state. In certain value sets, some of the derived utility values can be less than 0, and these represent health states that are worse than death for participants. As Bilbao (2018) pointed out, the EQ-5D-5L had a Cronbach's alpha of 0.86, a good correlation with WOMAC pain subscales with 0.688 and

0.782 for pain and function subscales, and a large difference in know-groups validity between the WOMAC and EQ-5D-5L and finally the EQ-5D-5L displayed high effect sizes (ES), and standardized response means (SRM), with similar values being found by Jin and colleagues along with good Guyatt Response Index (GRI) value, with similar results being found by Jin and colleagues (2019).

1.7 Orthopedic studies and cost-utility studies

A systematic review by Nwachukwu and colleagues (2015) indicated that there had been an increase in the number of CUA studies published over the past decade comparing total joint arthroplasty to various forms of similar surgical treatments and study designs. Their study identified 676 articles published between 1999 to 2014, of which 23 were determined to be of high quality compared to a previous study that reviewed articles from 1975 to 2001, which only identified 37 studies, of which only 11 qualified as CUA (Nwachukwu et al., 2015). Similar examples include work by Lau and colleagues (2021), who conducted a systematic review of the use of HRQOL and CUA within critical care. The study identified 80 CUA studies out of 8,926 possible studies and highlighted that difference exists in CUA methodology, questioning the validity of comparing QALYS. They also noted that 70% of HRQOL/QALYS were extrapolated from another source within studies that had multiple time points (Lau et al., 2021).

Similarly, Wisloff and colleagues (2014) conducted a systematic review of CUA published up to 2010. Their study aimed to identify the various methods of quantifying QALY gains across published CUA and determine the transparency of reporting QALYS and their reported size in gains (Wisloff et al., 2015). Of the 370 studies identified, the EQ-5D was the most common instrument for measuring QALYs gained, while the TTO was the most common valuation method used, followed by VAS, SG and PTO (Person Trade-off), while 42 studies had used mapping algorithms and 35 of those mapping studies had not specified their valuation method. Finally, a systematic review by Primeau and colleagues (2021) assessed the quality of published health economic evaluations within orthopedic populations and identified 93 health economic evaluations (Primeau et al., 2021). Their study indicated that the health economic evaluations conducted within Orthopedic sports medicine were of high quality, but nearly half of the studies did not perform full economic evaluations. In addition, their study pointed out that the majority of highly cited health economic evaluations were not of high quality in comparison to those published within Orthopedic sports medicine, but they have encouraged more full economic evaluations to be published to better understand effective ways of resource allocation (Primeau et al., 2021). These examples indicate the increased prevalence of health economic studies within medicine and health, its increased focus on CUA and its growing recognition across multiple research areas.

Orthopedic studies have recently begun leveraging preference-based surveys to assess the effectiveness of various therapies by comparing the QALYs gained from each treatment. However, in instances where health preference data is not collected explicitly and there is a need for health utility data, researchers have turned towards mapping strategies that allow the results from disease-specific surveys that assess clinical outcomes to be converted into health utility scores. Orthopedic studies have a particular use for these methods since many previous orthopedic studies did not use generic surveys to obtain health utility data; however, they have access to large databases that contain data from disease-specific surveys. For example, Dakin and colleagues utilized mapping strategies to estimate EQ-5D utilities from Oxford Knee scores from a patient population that underwent total knee replacements (Dakin et al., 2013). Similarly, multiple mapping algorithms have been published for converting disease-specific survey scores into health utility scores, with research on the validity of these methods (Dakin et al., 2020).

1.8 Mapping Algorithms

Mapping refers to the use of an algorithm to convert scores from disease-specific surveys into utility scores. Mapping is most useful for studies, clinical trials and assessment programs that have not administered any generic utility surveys during the progression of the study and would now like to add a health economic component to their study; most commonly a cost-utility study. In instances where health utility data has not been collected, but disease-specific surveys have been administered, mapping algorithms can convert those disease-specific scores into generic utility scores to be used in cost-utility studies. Mapping is known as cross walking since it allows researchers to predict health preferences from disease-specific or non-preference-based surveys. A recent systematic review by Dakin and colleagues (2018) published a systematic review that identified various mapping algorithms from patient-reported outcomes to the EQ-5D-3L or the EQ-5D-5L surveys.

Additionally, Dakin and colleagues (2020) also published a database for the Health Economics Research Center (HERC), which contains the various mapping algorithms published across various disease categories and patient groups. The database and systematic review provide health economic researchers with multiple sets of algorithms to use in CUA and highlights the crucial role of mapping within medical research. From their database, three algorithms were identified for converting WOMAC scores into utility scores. Such as those by Barton and colleagues (2008), Xie and colleagues (2010) and Wailoo and colleagues (2014). The algorithms have been most popular among clinicians and researchers and have been shown to have good external validity by Kaidaliri and Englund (2016). However, these algorithms were designed to convert WOMAC scores into utility scores for the EQ-5D-3L survey and not the EQ-5D-5L survey. Kaidalri and Englund noted that the linear regressions models used within the three algorithms by Barton, Xie and Wailoo would underpredict and overpredict depending on the varying health states and that these algorithms had a systematic bias, consequently cautioning against their use (Kaidaliri & Englund, 2016).

Recently, Bilbao and colleagues (2020) have published new algorithms based on multiple statistical methods such as general linear modelling, beta regression models, and Tobit models to convert the WOMAC scores into the EQ-5D-5L scores. These models were proven to be able to better predict utility values from WOMAC than those developed previously for the EQ-5D-3L surveys; through higher adjusted R^2 values (Bilbao et al., 2020). Previous studies by Xie (2010) and Barton (2008) focused on predicting scores from the EQ-5D-3L models and indicated that predicted utility scores could predict utility scores well but did not provide ICC values. Additionally, Barton's study revealed that the QALY gains from the predicted scores were significantly different in some cases resulting in inaccurate conclusions for CUA. Finally, Bilbao and Colleagues (2020) produced regression models which could be used in predicting the utility scores from WOMAC scores for a Spanish population using the EQ-5D-5L index values for the Spanish population. Their study produced multiple regressions (General Linear Model (GLM), Tobit model and Beta regression models), which have not yet been externally validated. Their results indicated that their preferred model was the GLM and the beta regression model, with domains of the WOMAC being used as covariates within the models.

1.9 Purpose: The study's primary objective is to determine the agreement between utility scores derived from WOMAC and EQ-5D-5L questionnaires on patients who have undergone TJR surgery on the hip or knee. A secondary objective is to compare agreement among subgroups divided by the length of time post-operative: 1 year; 2 to 5 years; 5 to 10 years, and over 10 years.

Chapter 2

2 Methods

2.1 Study Design

Our study employed a retrospective cohort design among patients who had undergone total hip and knee replacement surgery between June 2006 to December 2021. The study was approved by the University of Western Ontario's Health Science Research Ethics Board for Research involving human subjects and Lawson's approval.

2.2 Eligibility Criteria

We included individuals who had undergone total joint replacement surgery of the hip or knee at University Hospital between June 2006 and December 2021. Patients had to be at least 1-year post-operative and over the age of 18 years.

Participants were excluded if they were under 18; had undergone revision surgery, or did not have a complete WOMAC and EQ-5D-5L survey at a minimum of 1-year post-operative.

2.2.1 Sample size

There were 500 eligible patients in the database. This sample size provided sufficient power to provide estimates of agreement between the two scores (test-retest reliability = 0.90) with a prespecified level of precision (0.10). (Bonnet, 2002) for the total sample and for the subgroups of patients based on their post-operative time points: 1) 1 year; 2) 2 to 5 years; 3) 5-10 years; 4)>10 years. Additionally, to obtain good predictive accuracy the GLM mapping algorithm provided by Bilbao, a minimum of 400 patients were required, therefore, 500 participants would also provide sufficient power to run the GLM (Bilabo, et al. 2020).

2.2.2 Data collection

An institutional clinical database was used to identify patients who underwent a primary total hip or knee arthroplasty between 2006 to 2021. A retrospective review was performed to collect eligible patient data through the same database. All patients who undergo TJR at University Hospital are followed with patient-reported outcome measures, including the WOMAC and the EQ-5D-5L, at multiple time points such as the 3-month and 1-year post-operative time points and are further followed up every 1-2 years following surgery. The results from the questionnaires are entered into the clinical institutional database, and a score is calculated with standardized scoring algorithms. For the purposes of our study, we collected information regarding the participants' sex; age; date of most recent follow-up visit; type of primary joint replacement surgery (hip or knee), and their most recent WOMAC and EQ-5D-5L from their latest available follow-up visit.

The data set was further divided into subgroups based on how long it has been since their surgery. Subgroup 1 consisted of patients who were at the 1-year post-operation time point. Subgroup 2 consisted of patients who were 2 to 5 years post-operative. Subgroup 3 consisted of patients between 5 to 10 years post-operative and Subgroup 4 consisted of patients over 10 years post-operative.

2.3 Outcome measures

We collected demographic data, including age, sex, type of surgery (hip or knee), height (cm), weight (kg), BMI (kg/m^2), side of primary surgery (left or right), and follow-up visit time point.

The outcome measures used in this study are the EQ-5D-5L and the WOMAC survey. The EQ-5D-5L is a generic survey that assesses the overall quality of life in patients, while the WOMAC is a disease-specific survey that is aimed at evaluating the pain, stiffness and function of the hip or knee joints within OA patients. Both have been shown to have good measurement properties among patients with OA and who have undergone TJR (Lundgren-Nilsson et al., 2018; Bilbao et al., 2020)). The scores for each of the questions for the WOMAC surveys and the scores for each of the domains (pain, stiffness and physical function) were collected. In our study, the WOMAC scores for each of the domains were converted to a scale of 100, along with the WOMAC total score being out of 100. The higher scores indicated better health, while lower values indicated worse levels of health. The responses for the EQ-5D-5L surveys were combined into 5-digit health profiles which were used to identify the health utility index values from the Canadian EQ-5D-5L TTO value set (Xie et al., 2016). The participants were divided into subgroups based on the number of years post-op. Finally, we also collected information on the primary diagnosis factor for the surgery.

2.4 Mapping Algorithm

Based on the mapping algorithms developed by Bilbao and colleagues (2020), we created a General Linear Model (GLM) mapping algorithm to covert the WOMAC scores into a utility score.

The General Linear Model (GLM) utilized the following variables as covariates within the model to convert the WOMAC scores into utility scores: $Pain^2 + Pain^3$ + Function + Pain x Function. Each of the covariates for both the models required the original WOMAC values to be raised by powers and then be divided by 100 or 10,000 or required the values of the subscales to be multiplied and divided by 100, such as for the Pain X function covariate in the GLM models.

2.5 Plan for Analysis

Demographics for the study sample were summarized using means and Standard Deviations (SD) or frequencies and proportions where applicable. In addition, we calculated the distribution of the EQ-5D-5L utility scores and the predicted utility scores from the WOMAC, along with their mean and SD values. All data were analyzed using R version 4.0.4.

2.5.1 Agreement

To identify the level of agreement between the converted WOMAC utility and EQ5D utility values, we calculated the Intraclass Correlation Coefficient (ICC) (two-way mixed model with measures of consistency) and its 95% confidence interval. We considered ICC values below 0.5 to indicate poor reliability, 0.5 – 0.75 indicate moderate reliability,0.75 and 0.90 indicate good reliability and values > 0.90 indicate excellent reliability (Koo and Li, 2016).

We also calculated the standard error of measurement (SEM) and its 95% CI. The ICC provides information about the total variance (between and within-subject variability and random error), whereas the S.E.M. expresses individual measurement error only, without the influence of variance among patients.

2.5.2 Validity

To assess the validity of the converted WOMAC utility score, we performed a linear regression to determine the ability of patients' converted utility scores on the WOMAC questionnaire to predict the EQ-5D utility scores. To verify if our data met the assumptions of linear regression (linearity; homoscedasticity; independence of observations, and normality of data), we constructed residual vs fitted plots, quantile-quantile (qq) plots, scale-location plots, and residuals vs leverage plots. The first assumption of linearity refers to the idea that the relationship between the independent and dependent variables (Y and X) is linear. The second assumption, homoscedasticity, is centred around the equal distribution of errors or residual values, which is the difference between the actual and predicted utility scores in this case. The third assumption is the independence of observations which relates to the idea that the Y and X variables do not share any relations. Finally, the assumption of normality requires that residuals are normally distributed.

2.5.3 Mean difference of utility scores

We conducted a paired t-test to identify if there were any significant differences between the means of each utility score. Given our large sample size, we did not require the assumption of normality to be met.

2.5.4 Bland Altman Plot

To visually display the agreement between the WOMAC and EQ-5D-5L scores, we created Bland Altman plots which illustrate the magnitude of the difference between the two utility measures and show the distribution of the difference values over the entire range of the utility score. The closer the data points fall on the central line, the better the agreement between the two sets of measurements. We also displayed 95% confidence intervals around the mean difference. The graph also has two limits of agreement set a 1.96 SD away from the mean, which means that points that fall outside the confidence bands can have a significant amount of difference between the two means and indicates not only the presence of outliers in our data but also the presence of residual errors, these points can be considered to be extreme outliers. Residual errors are the difference between the observed and predicted utility values.

2.6 Outliers

Finally, we assessed our data for outliers using percentiles and boxplots. The percentiles were used to identify outliers that existed outside of 2.75% and 97.5% of the data. Boxplots for both the predicted and observed utility scores were made to graphically see the outliers present in our data. This was complemented by histograms displaying the distribution of the predicted and observed scores for the data

Chapter 3 Results

3.0 Results

A total of 500 participants were eligible for the study and were included in the analysis. The mean age of the patients was 64.4 (SD = 8.9), with 311 participants being female and 189 male (Table 1).

The mean EQ-5D-5L index value was 0.79 (0.18) (range -0.1 to 0.95). The mean total WOMAC score was 79.05(19.9), and the mean converted WOMAC utility score was 0.79(0.14). The mean values were similar for all subgroups (Table 2).

Characteristics (mean, SD)	Total sample population N = 500	Subgroup 1(1 year) N = 45	Subgroup 2 (2- 5 years) N = 287	Subgroup 3 (5 – 10 years) N = 116	Subgroup 4 (> 10 years) N = 52
Age(years)	64.39 (SD = 8.88)	64.79 (10.67)	65.13 (8.4)	64.83 (9.09)	59.04 (7.64)
Sex (Female), n(%)	311(62.22%)	29 (64%)	173 (60.3%)	74 (63.7%	35 (67.3%)
Height, cm	167.48 (10.21)	166.82 (10.99)	167.85 (9.95)	166.8 (10.62)	167.46 (10.22)
Weight, kg	94.19 (22.38)	93.83 (24.18)	94.30 (22.17)	92.05(22.99)	98.63 (20.4)
BMI, kg/, m2	33.54 (7.28)	33.46 (6.82)	33.45 (7.33)	32.94(7.02)	35.39 (7.86)
Primary THA (n)	183	16	104	44	19
Primary TKA	317	29	183	72	33
Operative Side (n %)					
Left	239	26	124	64	25
Right	261	19	163	52	27

Table 3.1 Demographics Table - Sample population demographics and characteristics by the total sample size and by subgroups

SD = standard deviation

QoL Scores (mean, SD)	Total sample population N = 500	Subgroup 1(< 1 year) N = 45	Subgroup 2 (2-5 years) N = 287	Subgroup 3 (5 – 10 years) N = 117	Subgroup 4 (> 10 years) N = 52
WOMAC					
Pain score	81.23 (20.84)	82.22 (16.74)	82.75 (21.68)	80.43 (18.57)	73.75 (22.81)
Function score	79.56(21.02)	81.34 (18.82)	81.5 (21.73)	78.4 (18.44)	69.91 (21.93)
Stiffness score	73.95 (21.86)	73.33(22.23)	75.65 (22.45)	72.41 (20.56)	68.51 (20.49)
Total score	79 (19.91)	80.03(17.1)	80.8 (20.74)	77.77 (17.58)	71.23 (20.73)
EQ-5D-5L utility index	0.79 (0.18)	0.78 (0.2)	0.80 (0.19)	0.77 (0.17)	0.76(0.17)
Converted WOMAC utility values	0.79(0.14)	0.81(0.12)	0.80(0.15)	0.79(0.12)	0.73(0.15)

Table 3.2 Summary table for predicted and observed utility score. Summarizes the domain scores for the total sample and subgroups along with the observed us and predicted utility values from the mapping algorithm.

SD = standard deviation.

3.1 Mapping Algorithm

Using the mapping algorithm, we derived the following equation to covert the WOMAC to a utility score:

$$Y = 0.129401 + 0.010297\,(Pain^2/100) – 0.005955(Pain^3/10000) + 0.007689(\text{Function scores}) - 0.004352(\text{Pain x Function}/100).$$

The residual standard error was identified as 0.1192, indicating that when the model predicts utility scores, it does so with an error of 0.1192. The model had a Multiple R^2 value of 0.58 and an adjusted R^2 value of 0.58, indicating a good fit for the model.

3.2 Agreement

3.2.1 Agreement between utility values

The mean scores for the predicted utility index value were 0.79 (0.14), with a minimum value of 0.15, a maximum value of 0.91 and a range of 0.75. We found good agreement between the EQ-5D-5L utility scores and the converted WOMAC scores (ICC = 0.85, 95% confidence interval =0.82 to 0.87). The Standard error of measurement (SEM) for the predicted utility scores for the total sample was calculated to be 0.054 (0.14 *$\sqrt{1-0.85}$),. indicating that the amount of error surrounding the predicted value is 0.054, which is low, providing further evidence of good agreement between the predicted and observed utility values.

Similarly, the ICC values based on the GLM for subgroups 1 – 4 have been summarized in table3.3

Subgroups	ICC values (C,2)	95 % Confidence interval (CI)
Subgroup 1 (N = 45)	0.777	0.592<ICC< 0.877
Subgroup 2 (N = 287)	0.879	0.848<ICC<0.904
Subgroup 3 (N = 116)	0.798	0.709<ICC<0.86
Subgroup 4 (N = 52)	0.811	0.67<ICC<0.891

Table 3.3 Intraclass Correlation (ICC) for subgroups. C stands for consistency, and 2 represents two surveys. CI = Confidence interval

The mean ICC values for each of the subgroups indicate that there is good agreement between the two methods. The subgroups with a lower number of participants had wider confidence intervals, while those with a larger sample size had a narrower 95% CI, as seen for subgroups 2 and 3. Despite the wide confidence intervals, the 95% CI for the ICC values for all subgroups remained within the moderate to good and good to excellent reliability.

3.3 Validity

To identify whether the assumptions for linear regression were met, normality plots for the data and linear regression analysis to compare the predicted and observed utility scores were conducted.

3.3.1 Normality plots for data

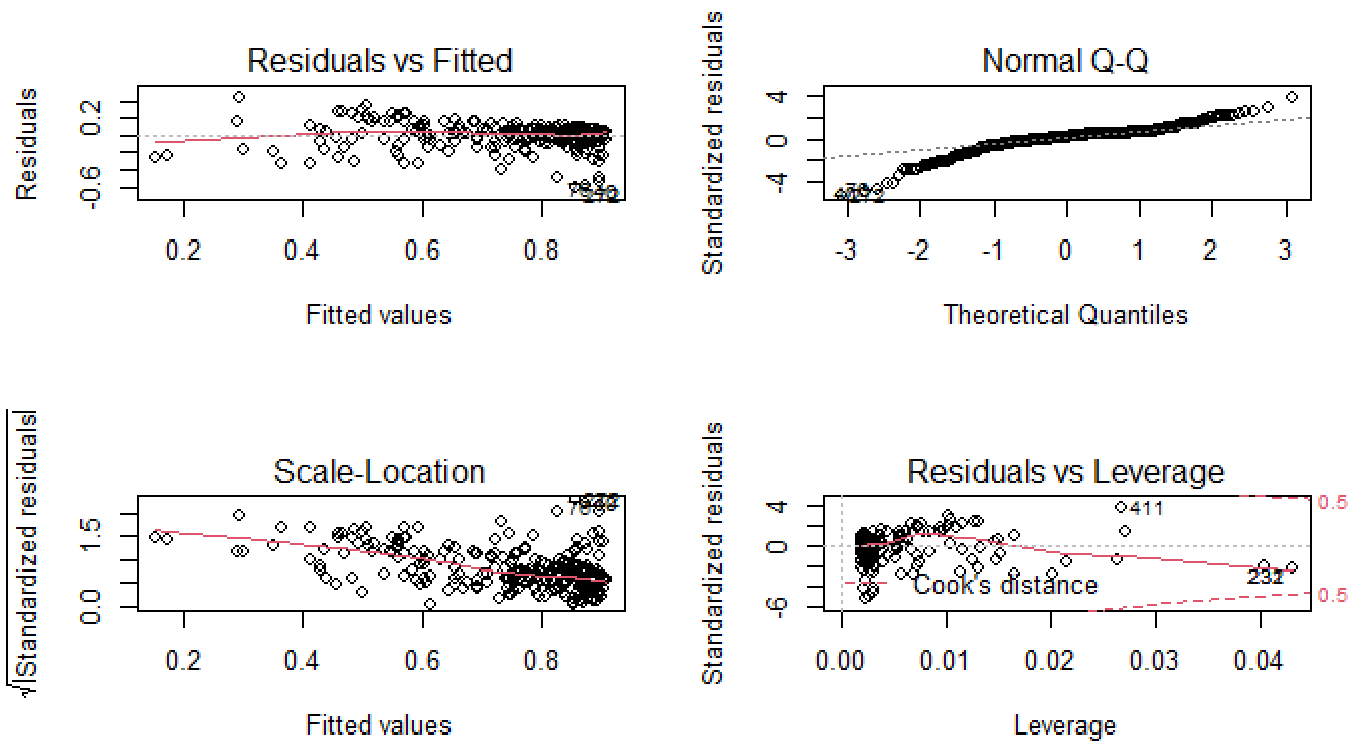

Figure 3.1. Normality plots for the linear model. Comparing observed utility scores to the predicted utility scores from the data. The top left graph displays the residuals vs fitted graph, where the fitted values (x-axis) are compared to the residual's values (y-axis) or differences between the predicted and observed utility values. The graph on the top right is a normal Q-Q plot comparing the theoretical quantiles to standardized residuals. The graph on the bottom left is a scale-location graph, comparing the fitted values to the square-rooted standardized residuals. The graph on the bottom right is a residual vs leverage graph, comparing the leverage of individual points to the standardized residuals; the Cooks Distance of note is 0.5.

The residuals vs fitted graph indicate that there is a large density of values towards the edge of the graph, meaning that the residuals are fairly lower at the higher utility scores than those at the lower scores where the points are more sparse. The red lines for the graph are relatively flat, with the tail end at the lower end of the fitted values being below the zero line. Ideally, the redline and the dotted line would remain parallel or on top of each other, indicating that no residuals are present in the data. However, since this is not the case, it indicates that the data is not linear. Similarly, since there is a cluster of points at the end of the graph, the assumptions of homoscedasticity have been violated, indicating that there is an unequal distribution of error variances in the data.

The Quantile-Quantile (QQplot) compares the quantiles and the standardized residuals. The graphs for the model show that the data is not following the 45-degree line perfectly, as it would have had the distribution of the data been normal. But since the data points on the lower halves of the QQ plot strays from the 45-degree line while the upper half remains on the dotted line, this indicates negative skewness in the distribution of points, along with the presence of a heavy tail.

For the scale-location plots for the linear model, the lack of homoscedastic data indicates heteroskedasticity or the unequal distribution of variances with the data. This is seen with the red line slowly declining towards the edge of the graph. Had the data been homoscedastic, the red line would have remained fairly flat, indicating that the variance distribution is equal.

Finally, the residuals vs leverage graphs for both models suggest the presence of influential points that can skew the data. The red lines on both graphs show that the lines travel upward and then downward, indicating the presence of both positive and negative residuals, respectively. Additionally, since the outliers on both graphs are identified with numbers, and since the dotted red lines are present on both graphs, it indicates that these outliers have significant leverage indicating that their residuals are large enough to change the slope of the redline.

Figure 3.3 displays a histogram of the distribution of the utility index scores from the EQ-5D-5L surveys, and the skewness towards the right indicates that the data is not distributed normally, with a large proportion of utility index scores being found on the right, indicating that on average the patients believe that they are in the best possible health state that they can be in. The lowest utility index value was -0.1023, and the highest value was 0.9490, with skewness of -2.05 and a kurtosis factor of 4.87. This also provides more evidence on the abnormal distribution of our data. Figure 3.4 shows the distribution of the WOMAC total scores within the sample, with the largest concentration of scores at 100, indicating good health and joint function. The mean value for total score from the WOMAC surveys of patients was 79.08 (19.9). These plots help illustrate the unequal distribution of values for the EQ-5D-5L and for the WOMAC, with our data set.

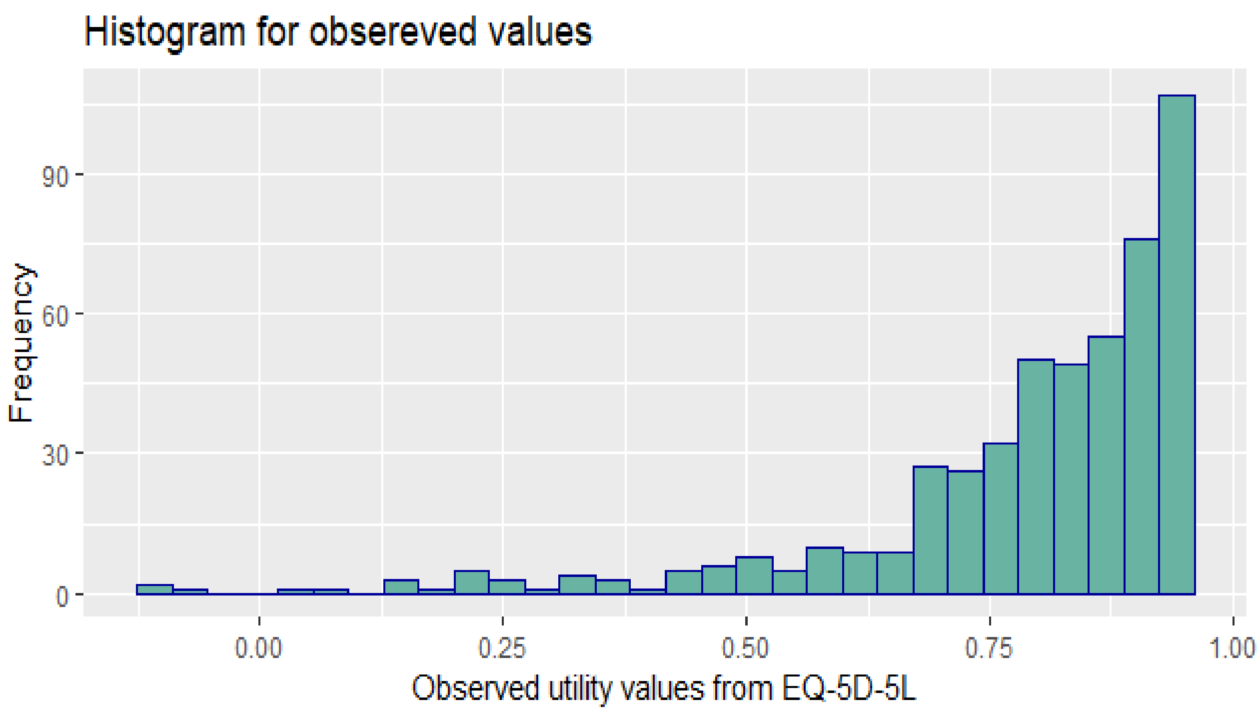

Figure 3.2. Represents the distribution of EQ-5D-5L utility index values within the sample population and its frequency.

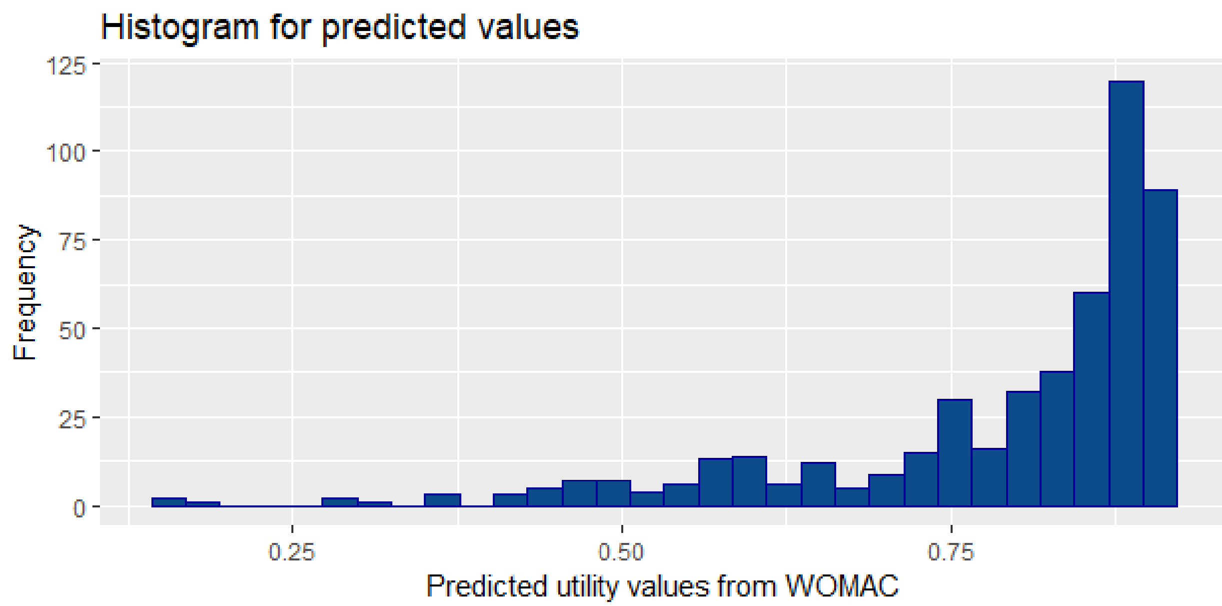

Figure 3.3. Distribution of total WOMAC scores across the sample population.

3.3.2 Linear Regression analysis

Coefficients	Estimate	5%	95%	Standard Error of measurement
Intercept	-0.0000632	-0.05007984	0.05013337	0.0305
Predicted EQ-5D-5L utility scores from WOMAC`	1.00068	0.93756961	1.06232743	0.054
Residual	0.119			
Multiple R^2	0.5825			
Adjusted R^2	0.5816			

Table 3.4Linear regression. 95% CI for coefficients for the model. The associated p value for the model was < 0.001.

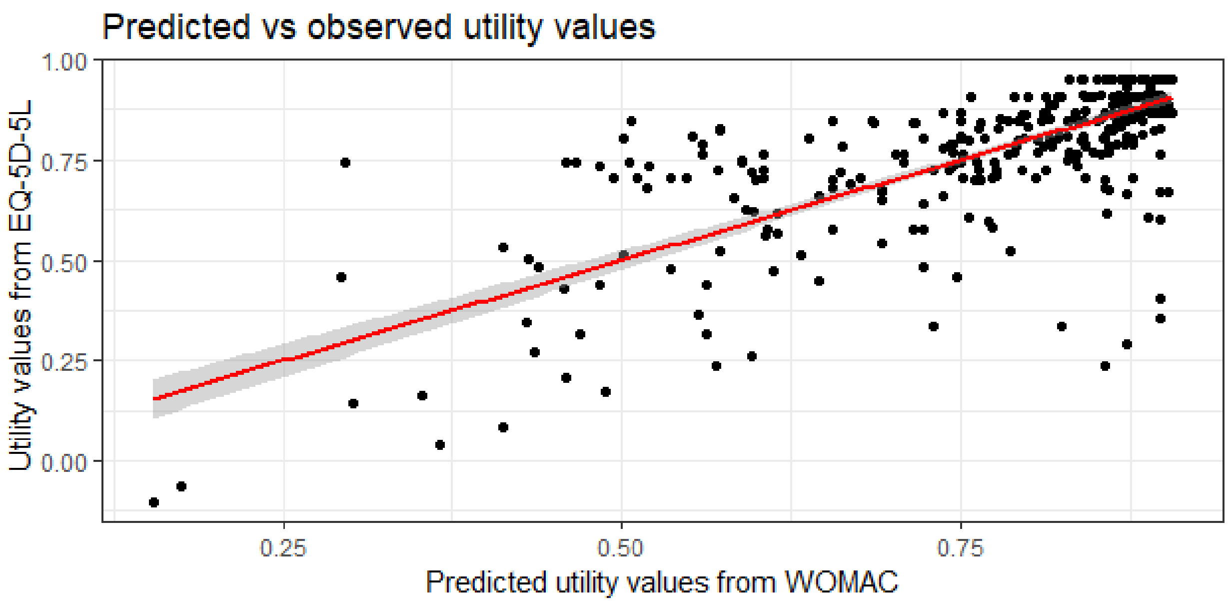

Figure.3.4 Linear regression between predicted and observed utility scores, with a 95% confidence band around the regression line, for the overall sample.

The converted utility value from the WOMAC questionnaires were a significant predictor of utility values provided by the EQ-5D-5L ($p<0.001$). The standard error of measurement was 0.054 and a residual error of 0.119 for the model. This means that the predicted values had an error of 0.054 from the true score and that the model predicts utility values with an error of 0.119 on average.

3.4 Mean Difference in utility scores

The mean difference between the two utility scores on the paired t-test was 0.000087 with a 95% CI (-0.0103 – 0.0105) p= 0.987. Since the p-value is higher than 0.05, we cannot reject the null hypothesis that the true difference between the two variables is equal to 0. This is also evident when looking at the mean difference value of 0.000087, which means that there is little or no difference in the means. With similar being presented across the subgroups as seen in table 3.5.

	Mean difference	Lower CI	Upper CI	P values
Total Sample	0.000087	-0.010	0.010	0.987
Subgroup 1	0.028	-0.013	0.069	0.182
Subgroup 2	-0.0054	-0.018	0.0073	0.403
Subgroup 3	0.016	-0.0058	0.039	0.1438
Subgroup 4	-0.030	-0.067	0.0060	0.1003

Table 3.5 Mean difference and 95% CI around mean difference for the total sample and for subgroups.

3.5 Bland Altman Plots

Figure 3.5 shows a Bland Altman plot which compares the means of the two scores and the differences between the two scores.

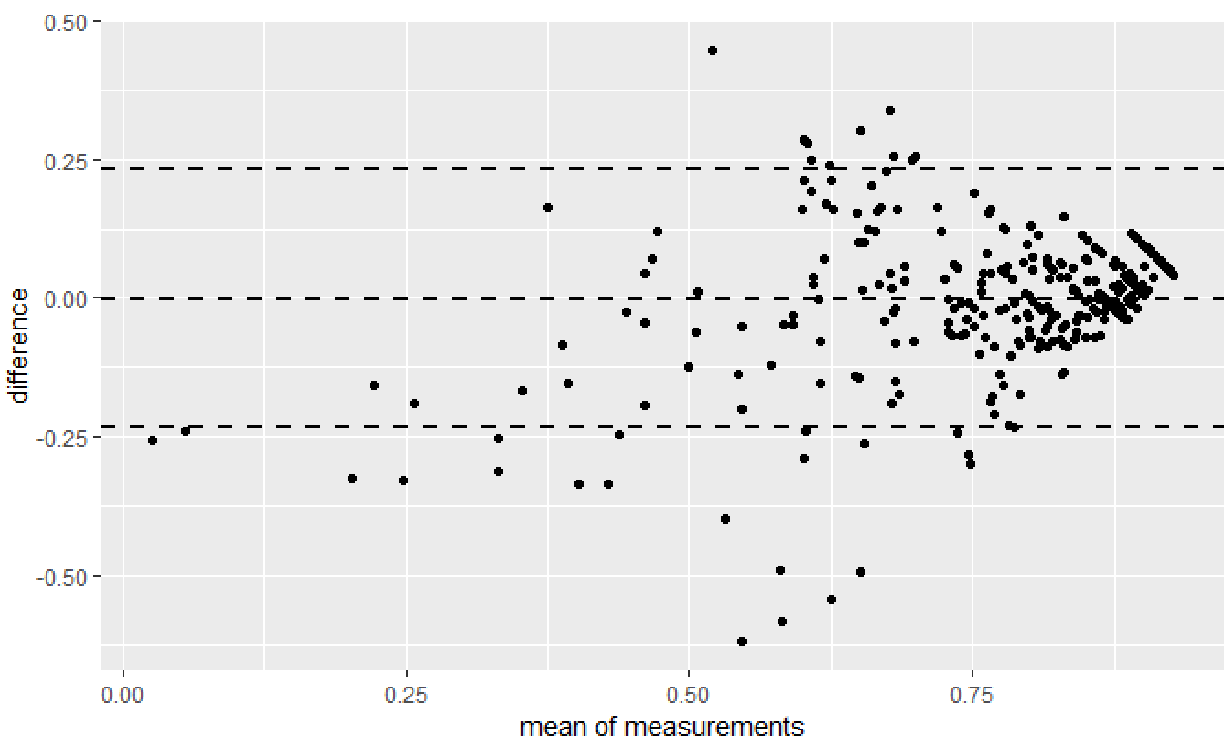

Figure. 3.5 Bland Altman plot comparing the two methods for the entire sample. Since there is a large cluster of points close to the zero line, we can interpret it as both methods having good agreement in producing utility values.

As shown in Figure 3.5, the largest density of points exists between the bias lines and the two limits, and a few points that exist outside the limits indicating that the two utility scores have good agreement.

Additionally, when looking at the associated Bland Altman plots for each of the subgroups, the subgroups with higher ICC values such as subgroups 2 and 3, have more data points that fall within the 95% confidence limits than subgroups 1 and 4. Their points are closer to the middle line of no difference, indicating that the amount of residual error present within those points is lower than those outside the limits.

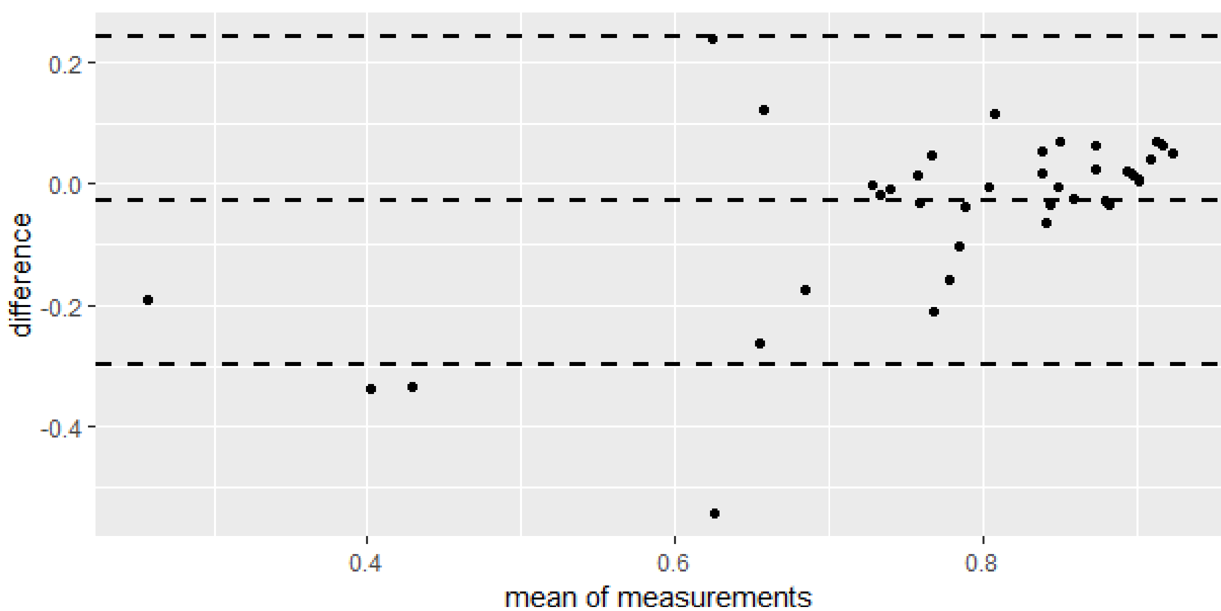

Figure 3.6. Bland Altman plot for subgroup 1.

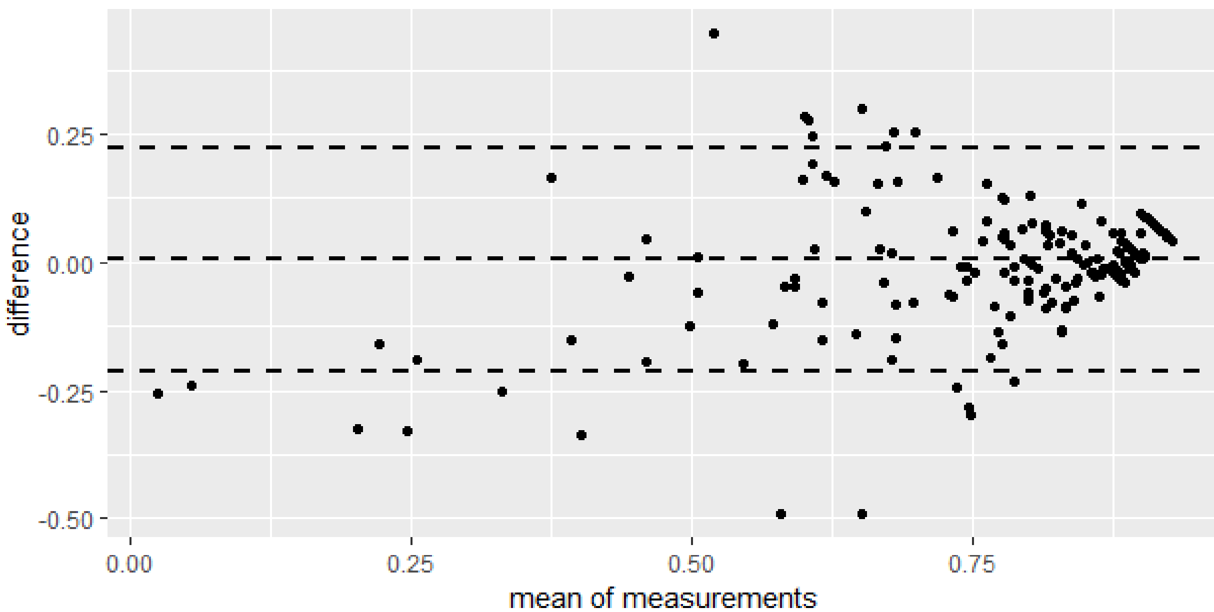

Figure3.7. Bland Altman plot for subgroup 2.

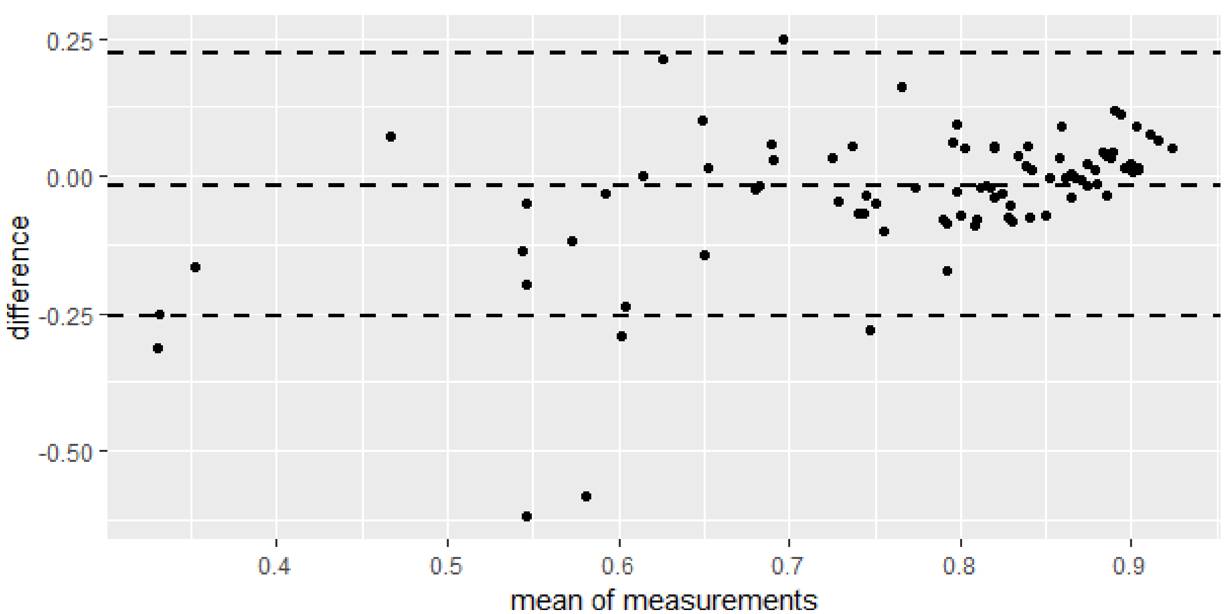

Figure 3.8. Bland Altman plot for subgroup 3.

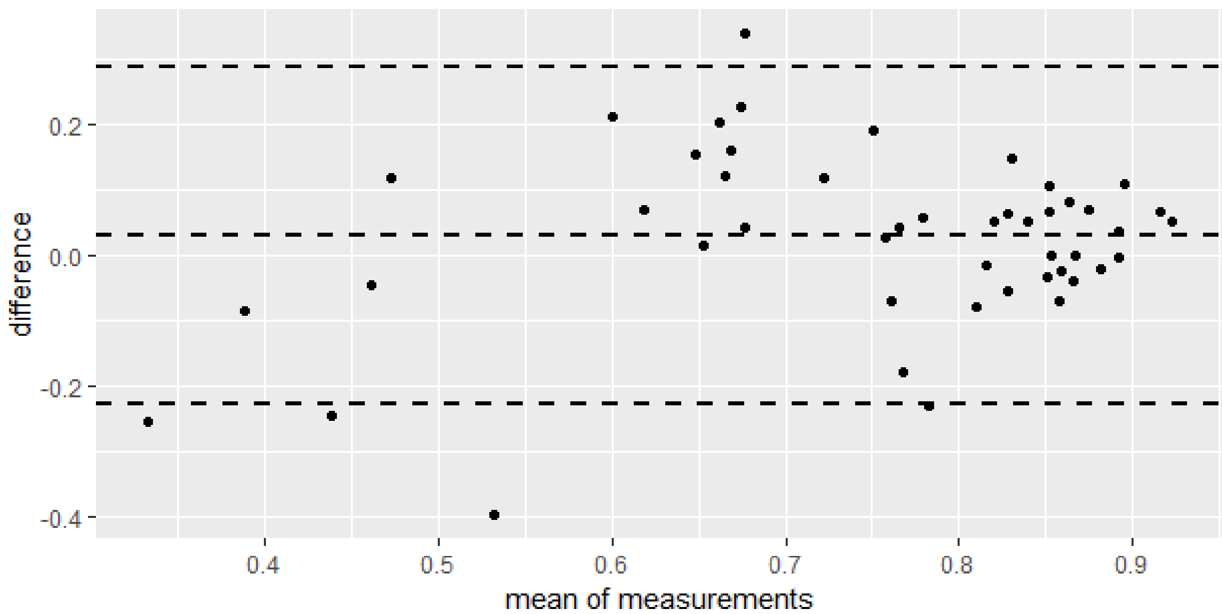

Figure 3.9. Bland Altman plot for subgroup 4.

3.6 Outliers

A total of 33 outliers were identified within the dataset with 10 outliers being commonly reported for both the predicted utility values and the EQ-5D-5L utility scores. These outliers were identified using the quantile function in R. The lower boundary was defined as 2.5%, and the upper boundary was defined as 97.5%. Therefore, any points that fell outside the boundary for both the surveys were identified as outliers. A table has been provided in Appendix D which summarizes the findings for the outliers for both surveys, while the boxplot provided in Figure 3. 10 provides a visual display of the number of outliers present within each of the surveys.

Individuals identified as outliers were mostly from subgroup 2 with 26 patients. All of the outliers either had low health profiles or high health profiles. Similarly, these individuals also had low predicted utility values based on the GLM model. This could be due to the model either underpredicting or overpredicting scores for individuals with low or high health states.

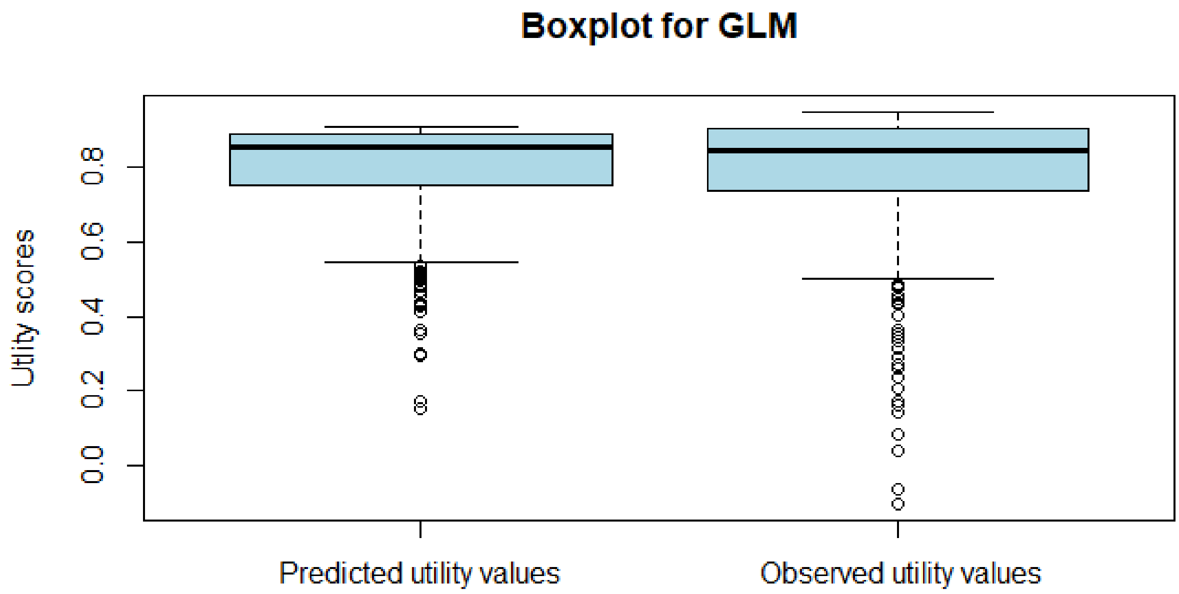

Figure 3.10. Boxplots for the predicted utility values from the WOMAC survey and the observed utility values for the EQ-5D-5L for our data.

When the outliers were removed, the ICC value for the total sample fell to 0.78 and 95% CI of 0.74 to 0.82. The decrease in agreement is due to the lower number of participants present in the data, which now stands at 464 participants. The Bland Altman plot for the new dataset, displayed below in figure 3.11 shows fewer points present below the lower agreement band. Similarly, the data points that were common outliers for both the EQ-5D-5L and the predicted utility scores from the WOMAC had an ICC value of 0.415 and 95% CI between -1.591 to 0.868. These results indicate that there is poor agreement between the data points that were commonly identified as outliers for both surveys. Additionally, the Bland Altman plot for the common outliers has been provided in figure. 3.12.

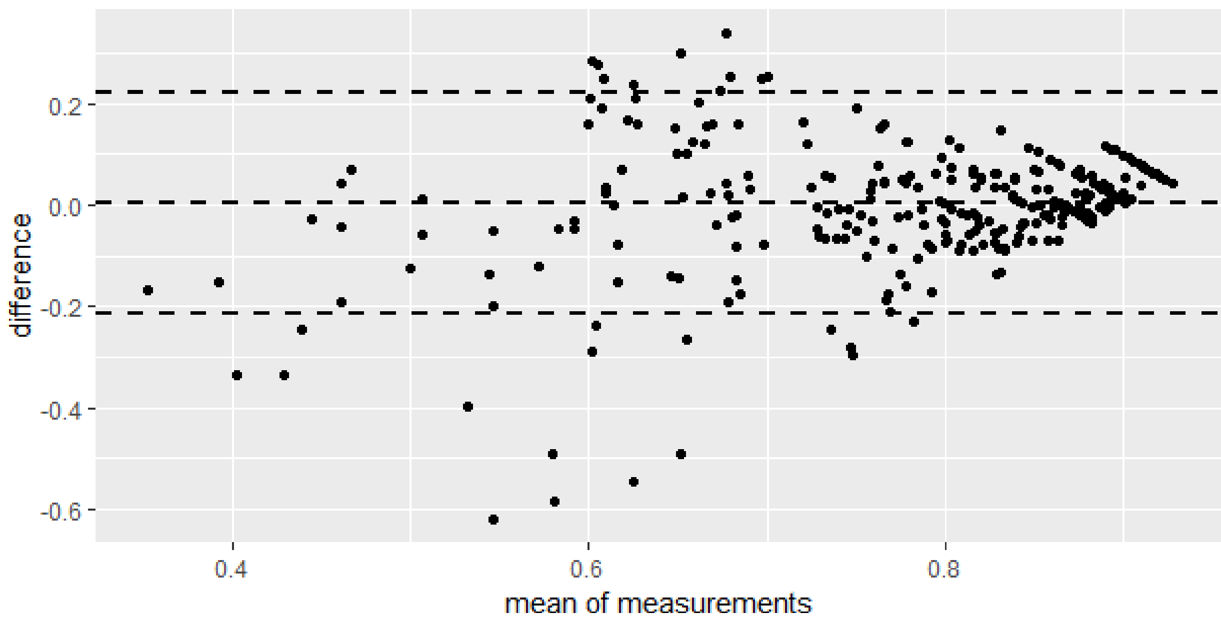

Figure 3.11. Bland Altman plot with outliers removed.

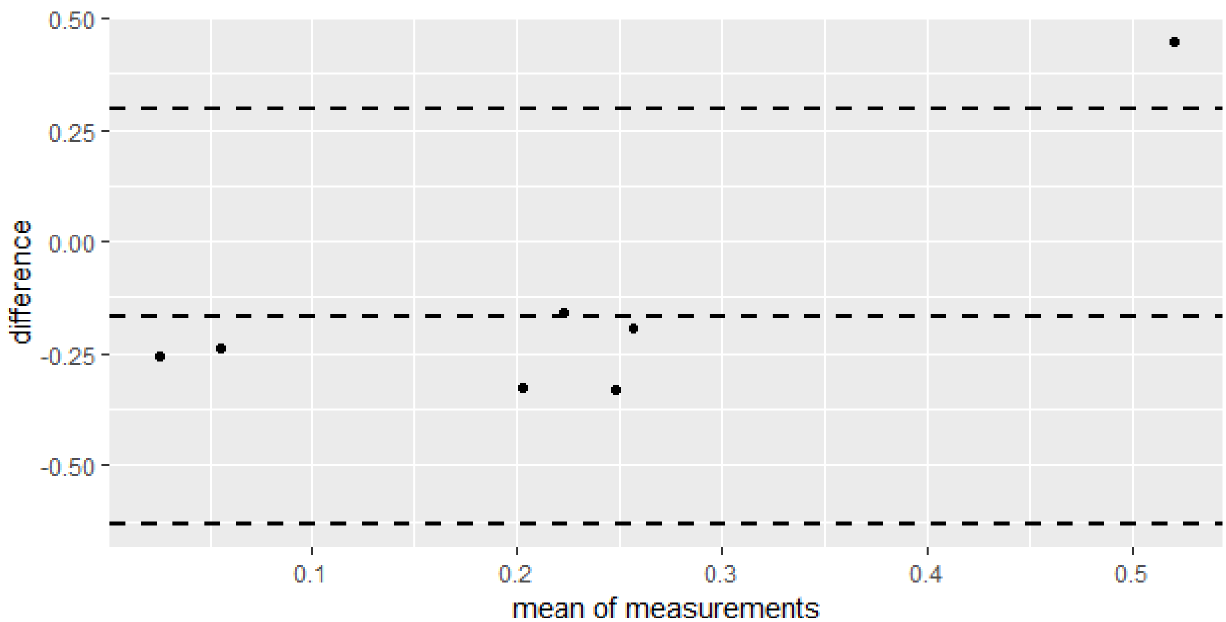

Figure 3.12. Bland Altman plot for outliers that were common for both surveys.

Chapter 4.0

4.0 Discussion

The purpose of this study was to measure the agreement between converted utility scores and EQ-5D-5L utility scores. We found good agreement between the two measurement instruments for the sample and found moderate to good agreement for subgroups. The paired t-test indicates that there is an extremely small difference between the two mean scores that was not statistically significant. When combined with the ICC value for the sample population, of 0.85, with a 95% CI (0.82 – 0.87), it is evident that there is good agreement between the WOMAC predicted utility values and the observed utility scores from the EQ-5D-5L. However, the ICC values for the subgroups were 1 year post-op and those over 10 years post-op had wider confidence intervals compared to the other groups, indicating greater uncertainty. These groups had fewer patients, which may have contributed to the greater uncertainty, despite the small mean difference values reported in Table 3.5.

We found that individuals who had lower utility index values and low total scores on the WOMAC had lower agreement. The mean ICC value for outliers was 0.41 with a 95% CI between -1.6 to 0.87, which indicates poor agreement between the WOMAC predicted utility scores and the EQ-5D-5L observed scores. The utility scores generated by the EQ-5D-5L represent the HRQOL weight multiplied by time to derive QALYs for each course of action within health economic evaluations, CUA in particular. While our results indicate that there is good agreement between the predicted and observed utility scores, the actual QALYs being generated will differ, as demonstrated by Barton (2008). Our results indicate that the HRQOL values being generated from the WOMAC will be much closer to the actual values than those previously generated from older mapping models, however, they will not be an exact match to the actual utility values. Therefore, in clinical practice, mapping algorithms can be an effective tool in identifying an approximation of the utility of TJR in patients who have undergone the procedure. But as Barton, Xie, and Bilbao have noted, mapping algorithms, including the algorithm used in this study can overpredict certain health states, as seen in the outliers for low health states (Barton, et al., 2008 ; Xie et al., 2010 ; Bilbao et al. 2020). Therefore, when interpreting the results from mapped surveys, caution must be applied for using this method to identify utility.

Our results are consistent with the ICC values provided by Bilbao and colleagues (2020) (ICC= 0.826). Their study consisted of patients diagnosed with hip and knee OA and compared agreement between the observed utility values and predicted utility values from the EQ-5D-5L and WOMAC scores from their baseline to the values generated at a second timepoint, 6 months later. The baseline results were used to identify mapping algorithms, while their 6-month time point results were used to internally validate their model, along with the relevant statistics for prediction and agreement. Their study produced 3 types of mapping algorithms, of which their preferred models were the GLM and beta regression models. The GLM and Beta regression models had good ICC values with 0.826 (0.800 – 0.851) and 0.830 (0.804 – 0.855), indicating good agreement. Their adjusted R^2 value for the GLM was found to be 0.61. Our study added to their results, by using their mapping algorithm on a Canadian population and thereby externally validating their GLM. Additionally, we used a Canadian EQ-5D-5L TTO value set (Xie et al., 2016).

Similarly, Xie and colleagues (2010) mapped the WOMAC onto the EQ-5D-3L within a Singaporean population diagnosed with knee OA , using a Japanese EQ-5D-3L scoring algorithm. Their study developed mapping algorithms based on the Ordinary Least Squares(OLS) and Censored Least Absolute Deviation (CLAD)s methods. Their preferred model was the OLS with WOMAC domain scores being used as variables since it had the best adjusted R^2 value of 0.499. The primary focus of their study was the predicted accuracy between the predicted values and the observed EQ-5D-3L utility values. Their study also provided information on individual and group differences of the predicted and observed scores. While knowing the absolute differences in values can be helpful, it does not provide an exact scale or magnitude of the agreement, like the ICC used in our study.

Barton and colleagues (2008), also developed mapping algorithms for converting the WOMAC to the EQ-5D-3L utility scores, within the UK population using the UK value set. Their study was comprised of individuals diagnosed with knee pain who were a part of the Lifestyle Interventions for Knee Pain (LIKP) study that compared four sets of interventions aimed at reducing knee pain. The primary outcome measures were the WOMAC and the EQ-5D-3L surveys, with participants being asked to complete the surveys at multiple time points such as baseline, 6 months, 12 months, and 24 months. Their adjusted R^2 value was 0.313 for their preferred model. Their study used the utility scores within a CUA to identify the QALY gains and the observed and predicted incremental cost per QALY. They identified that there were the differences in the incremental cost per QALY estimates based on the utility value generated from the WOMAC to the actual utility values from the EQ-5D-3L.

4.1 Strengths and Limitations

Our study is the first to incorporate a mapping algorithm for the WOMAC to obtain utility scores for the EQ-5D-5L using the Canadian EQ-5D-5L TTO value set (Xie et.al., 2016). Additionally, our study was large enough to run a GLM and thus provides further evidence for Bilbao's GLM by contributing to the external validity of the mapping algorithm.

In terms of limitations for the study, our study did not implement all 3 mapping algorithms constructed by Bilbao (2020), therefore, we could not compare the agreement statistic across all models based on a Canadian population. Implementing more complex models such as the Beta regression model may account for more errors or better prediction of utility values.

Further, as our sample consisted entirely of post-operative total joint replacement surgery patients, the majority had high utility values, and our sample did not have a representative distribution of possible utility scores. Therefore, future studies are needed to evaluate agreement among other orthopaedic patient populations, such as those preoperative, who may have lower health states.

Chapter 5

5.0 Conclusion

Our findings indicate that the practice of using mapping algorithms to ascertain utility scores from disease-specific surveys such as the WOMAC can produce accurate results with a SEM for predicted values of 0.054. While the utility index values generated from mapping algorithms do not provide perfect agreement with utility scores from preference-based surveys, they provide values that are closer to the observed utility scores than previous mapping algorithms published by Xie (2010) and Barton (2008), as seen in the predictive ability of the Bilbao model in the adjusted R^2 values.

5.1 Future Directions

Future studies can identify the impact of different utility values generated from mapping algorithms on the incremental cost utility ratios to understand the underlying implications of using mapping algorithms to determine the value of health interventions.

REFERENCE

Allen, K. D., Thoma, L. M., & Golightly, Y. M. (2022). Epidemiology of osteoarthritis. *Osteoarthritis and cartilage*, *30*(2), 184–195. https://doi.org/10.1016/j.joca.2021.04.020

American Academy of Orthopaedic Surgeons Management of Osteoarthritis of the Knee (NonArthroplasty) Evidence-Based Clinical Practice Guideline. https://www.aaos.org/oak3cpg Published 08/31/2021)

Barton, G. R., Sach, T. H., Jenkinson, C., Avery, A. J., Doherty, M., & Muir, K. R. (2008). Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores?. *Health and quality of life outcomes*, *6*, 51. https://doi.org/10.1186/1477-7525-6-51

Bilbao A, García-Pérez L, Arenaza JC, García I, Ariza-Cardiel G, Trujillo-Martín E, Forjaz MJ, Martín- Fernández J. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. Qual Life Res. 2018 Nov;27(11):2897-2908. doi: 10.1007/s11136-018-1929-x. Epub 2018 Jul 5. PMID: 29978346.

Bilbao, A., García-Pérez, L., Arenaza, J. C., García, I., Ariza-Cardiel, G., Trujillo-Martín, E., Forjaz, M. J., & Martín-Fernández, J. (2018). Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Quality of life research : an international journal of quality of life* *aspects of treatment, care and rehabilitation*, *27*(11), 2897–2908. https://doi.org/10.1007/s11136-018-1929-x

Birtwhistle, R., Morkem, R., Peat, G., Williamson, T., Green, M. E., Khan, S., & Jordan, K. P. (2015). Prevalence and management of osteoarthritis in primary care: an epidemiologic cohort study from the Canadian Primary Care Sentinel Surveillance Network. *CMAJ open*, *3*(3), E270–E275.

Canadian Institute for Health Information. Hip and Knee Replacements in Canada: CJRR Annual Statistics Summary, 2018–2019. Ottawa, ON: CIHI; 2020.

Dakin, H., Gray, A., & Murray, D. (2013). Mapping analyses to estimate EQ-5D utilities and responses based on Oxford Knee Score. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*, *22*(3), 683–694. https://doi.org/10.1007/s11136-012-0189-4

Dakin H, Abel L, Burns R, Yang Y. 2018. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. Health and Quality of Life Outcomes. 16:31.

Dakin H, Abel, L, Burns R, Yang Y, 2020. HERC database of mapping studies, Version 8.0 (Last updated: 14th October 2020). Available at: http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies.

Devlin, N., Parkin, D., & Janssen, B. (2020). *Methods for Analysing and Reporting EQ-5D Data*. Springer.

Drummond, M.E., Schulpher, M.J.J, Torrance, G.W. et.al. Methods for the economic evaluation of health care programmes, 4th edition. Oxford University Press, 2015.

Giavarina D. (2015). Understanding Bland Altman analysis. *Biochemia medica*, *25*(2), 141–151. https://doi.org/10.11613/BM.2015.015

Hawker GA. Osteoarthritis is a serious disease. Clin Exp Rheumatol. 2019 Sep-Oct;37 Suppl 120(5):3-6.

Janssen, M. F., Bonsel, G. J., & Luo, N. (2018). Is EQ-5D-5L Better Than EQ-5D-3L? A Head-to-Head Comparison of Descriptive Systems and Value Sets from Seven Countries. *PharmacoEconomics*, *36*(6), 675–697. https://doi.org/10.1007/s40273-018-0623-8

Jin, Al Sayah F, Ohinmaa A, Marshall DA, Johnson JA. Responsiveness of the EQ-5D-3L and EQ-5D-5L in patients following total hip or knee replacement. Qual Life Res. 2019 Sep;28(9):2409-2417. doi: 10.1007/s11136-019-02200-1. Epub 2019 May 14. PMID: 31089988.

KELLGREN, J. H., & LAWRENCE, J. S. (1957). Radiological assessment of osteoarthritis. *Annals of the rheumatic diseases*, *16*(4), 494–502. https://doi.org/10.1136/ard.16.4.494).

Kiadaliri, A. A., & Englund, M. (2016). Assessing the external validity of algorithms to estimate EQ- 5D-3L from the WOMAC. *Health and quality of life outcomes*, *14*(1), 141. https://doi.org/10.1186/s12955-016-0547-y

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kolasinski, S. L., Neogi, T., Hochberg, M. C., Oatis, C., Guyatt, G., Block, J., Callahan, L., Copenhaver, C., Dodge, C., Felson, D., Gellar, K., Harvey, W. F., Hawker, G., Herzig, E., Kwoh, C. K., Nelson, A. E., Samuels, J., Scanzello, C., White, D., Wise, B., … Reston, J. (2020). 2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee. *Arthritis care & research*, *72*(2), 149–162. https://doi-org.proxy1.lib.uwo.ca/10.1002/acr.24131)

Kraus, V. B., Blanco, F. J., Englund, M., Karsdal, M. A., & Lohmander, L. S. (2015). Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. Osteoarthritis and cartilage, 23(8), 1233–1241. https://doi.org/10.1016/j.joca.2015.03.036

Lundgren-Nilsson, Å., Dencker, A., Palstam, A., Person, G., Horton, M. C., Escorpizo, R., Küçükdeveci, A. A., Kutlay, S., Elhan, A. H., Stucki, G., Tennant, A., & Conaghan, P. G. (2018). Patient-reported outcome measures in osteoarthritis: a systematic search and review of their use and psychometric properties. *RMD open*, *4*(2), e000715. https://doi.org/10.1136/rmdopen- 2018-000715

McConnell, S., Kolopack, P. and Davis, A.M. (2001), The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. Arthritis & Rheumatism, 45: 453-461.

McConnell, S., Kolopack, P., & Davis, A. M. (2001). The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. *Arthritis and rheumatism*, *45*(5),453–461.https://doi.org/10.1002/1529-0131(200110)45:5<453::aid-art365>3.0.co;2-w

Nwachukwu, B. U., Bozic, K. J., Schairer, W. W., Bernstein, J. L., Jevsevar, D. S., Marx, R. G., & Padgett, D. E. (2015). Current status of cost utility analyses in total joint arthroplasty: a systematic review. *Clinical orthopaedics and related research*, *473*(5), 1815–1827. https://doi.org/10.1007/s11999-014-3964-4

Osteoarthritis Research Society International (OARSI) (2016). Standardization of Osteoarthritis definitions. Retrieved from: https://oarsi.org/research/standardization-osteoarthritis-definitions

Primeau, C. A., Zomar, B. O., Somerville, L. E., Joshi, I., Giffin, J. R., & Marsh, J. D. (2021). Health Economic Evaluations of Hip and Knee Interventions in Orthopaedic Sports Medicine: A Systematic Review and Quality Assessment. *Orthopaedic Journal of Sports Medicine*. https://doi.org/10.1177/2325967120987241

Sharif, B., Garner, R., Hennessy, D., Sanmartin, C., Flanagan, W. M., & Marshall, D. A. (2017). Productivity costs of work loss associated with Osteoarthritis in Canada from 2010 to 2031. *Osteoarthritis and cartilage*, *25*(2), 249–258. https://doi.org/10.1016/j.joca.2016.09.011).

Sharif, B., Kopec, J., Bansback, N., Rahman, M. M., Flanagan, W. M., Wong, H., Fines, P., & Anis, A. (2015). Projecting the direct cost burden of osteoarthritis in Canada using a microsimulation model. *Osteoarthritis and cartilage*, *23*(10), 1654–1663. https://doi.org/10.1016/j.joca.2015.05.029

Tarride, J.-E., Haq, M., O'Reilly, D.J., Bowen, J.M., Xie, F., Dolovich, L. and Goeree, R. (2012), The excess burden of Osteoarthritis in the province of Ontario, Canada. Arthritis & Rheumatism, 64: 1153-1161. https://doi.org/10.1002/art.33467

Wailoo, A., Hernandez Alava, M., & Escobar Martinez, A. (2014). Modelling the relationship between the WOMAC Osteoarthritis Index and EQ-5D. *Health and quality of life outcomes*, *12*, 37.

Wisløff, T., Hagen, G., Hamidi, V., Movik, E., Klemp, M., & Olsen, J. A. (2014). Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. *PharmacoEconomics*, *32*(4), 367–375. https://doi.org/10.1007/s40273-014-0136-z

White, D. K., & Master, H. (2016). Patient-Reported Measures of Physical Function in Knee Osteoarthritis. *Rheumatic diseases clinics of North America*, *42*(2), 239–252. https://doi.org/10.1016/j.rdc.2016.01.005

Whitehead, S. J. & Ali, S. (2010). Health Outcomes in economic evaluations: the QALY and utilities, British Medical Bulletin, 96(1), 5-21. https://doi.org/10.1093/bmb/ldq033

Xie, F., Pullenayegum, E. M., Li, S. C., Hopkins, R., Thumboo, J., & Lo, N. N. (2010). Use of a disease- specific instrument in economic evaluations: mapping WOMAC onto the EQ-5D utility index. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *13*(8), 873–878. https://doi.org/10.1111/j.1524-4733.2010.00770.x

Xie, F., Pullenayegum, E., Gaebel, K., Bansback, N., Bryan, S., Ohinmaa, A., Poissant, L., Johnson, J. A., & Canadian EQ-5D-5L Valuation Study Group (2016). A Time Trade-off-derived Value Set of the EQ-5D-5L for Canada. *Medical care*, *54*(1), 98–105. https://doi.org/10.1097/MLR.0000000000000447

Yu, S. P., & Hunter, D. J. (2015). Managing osteoarthritis. *Australian prescriber*, *38*(4), 115–119. https://doi.org/10.18773/austprescr.2015.039

Appendices

Appendix A Ethics approval

Western Research

Date: 13 September 2021

To: Dr. Jacquelyn Marsh

Project ID: 118680

Study Title: Validity of Quality of Life Measurement in Economic Evaluations in Total Joint Replacement: Agreement between Western Ontario and McMasters' Osteoarthritis Index (WOMAC) and the EuroQOL 5D (EQ-5D) Utility scores

Application Type: HSREB Initial Application

Review Type: Delegated

Meeting Date / Full Board Reporting Date: 21/Sept/2021

Date Approval Issued: 13/Sep/2021

REB Approval Expiry Date: 13/Sep/2022

Dear Dr. Jacquelyn Marsh

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study as described in the WREM application form, as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. **All other required institutional approvals and mandated training must also be obtained prior to the conduct of the study.**

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Demographics Survey	Online Survey	14/Aug/2021	1
Online surveys of Womac and EQ-5D surveys	Online Survey	15/Aug/2021	3
Demographics Survey	Paper Survey	24/Aug/2021	1
Telephone script V2 September 7 2021	Telephone Script		
Paper copy WOMAC and EQ-5D surveys August 15 2021	Paper Survey	15/Aug/2021	3
LOI and consent form Sept 8 2021	Written Consent/Assent	08/Sep/2021	3
Protocol September 10 2021	Protocol	10/Sep/2021	3

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
Reference document for rationale	References	26/May/2021	1

No deviations from, or changes to, the protocol or WREM application should be initiated without prior written approval of an appropriate amendment from Western HSREB , except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University HSREB operates in compliance with, and is constituted in accordance with, the requirements of the TriCouncil Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions. Sincerely,

LAWSON APPROVAL

LAWSON APPROVAL NUMBER: R-21-565

PROJECT TITLE: Validity of Quality of Life Measurement in Economic Evaluations in Total Joint Replacement: Agreement between Western Ontario and McMasters' Osteoarthritis Index (WOMAC) and the EuroQOL 5D (EQ-5D) Utility scores

PRINCIPAL INVESTIGATOR: Dr. Jacquelyn Marsh

LAWSON APPROVAL DATE: Friday, 19 November 2021

ReDA ID: 11053

Overall Study Status: Active

Please be advised the above project was reviewed by Lawson Administration and the project was approved. Your official approval document can be found in the documents section of your study in ReDA.

Appendix B

EQ-5D-5L

Under each heading please tick the **ONE** box that best describes your health TODAY.

Mobility

I have no problems in walking about ☐

I have slight problems in walking about ☐

I have moderate problems in walking about ☐

I have severe problems in walking about ☐

I am unable to walk about ☐

Self care

I have no problems washing or dressing myself ☐

I have slight problems washing or dressing myself ☐

I have moderate problems washing or dressing myself ☐

I have severe problems washing or dressing myself ☐

I am unable to wash or dress myself ☐

Usual activities (eg. work, study, housework, family or leisure activities)

I have no problems doing my usual activities ☐

I have slight problems doing my usual activities ☐

I have moderate problems doing my usual activates ☐

I have severe problems doing my usual activities ☐

I am unable to do my usual activities ☐

Pain/ discomfort

I have no pain or discomfort ☐

I have slight pain or discomfort ☐

I have moderate pain or discomfort ☐

I have severe pain or discomfort ☐

I have extreme pain or discomfort ☐

Anxiety/Depression

I am not anxious or depressed ☐

I am slightly anxious or depressed ☐

I am moderately anxious or depressed ☐

I am severely anxious or depressed ☐

I am extremely anxious or depressed ☐

Appendix C

WOMAC

A. Think about the ***pain*** you felt in the hip/knee during the last 4 months

Question: How much pain do you have?	None	Mild	Moderate	Severe	Extreme
1. Walking on a flat surface	0	1	2	3	4
2. Going up or down stairs	0	1	2	3	4
3. At night while in bed, pain disturbs your sleep	0	1	2	3	4
4. Sitting or lying	0	1	2	3	4
5. Standing upright	0	1	2	3	4

B. Think about the ***stiffness (not pain)*** you felt in your hip/knee during the last 4 weeks. Stiffness is a sensation of ***decreased*** ease in moving your joint.

	None	Mild	Moderate	Severe	Extreme
6. How ***severe*** is your ***stiffness after first awakening*** in the morning?	0	1	2	3	4
7. How ***severe*** is your stiffness after sitting, lying or resting l***ater in the day?***	0	1	2	3	4

C. Think about the ***difficulty*** you had in doing the following physical activities due to your hip/knee during the last 4 weeks. By this we mean your ***ability to move around and look for yourself.***

Question: what degree of difficulty	None	Mild	Moderate	Severe	Extreme

do you have?					
8. Descending stairs	0	1	2	3	4
9. Ascending stairs	0	1	2	3	4
10 Rising from sitting	0	1	2	3	4
11. Standing	0	1	2	3	4
12. Bending to the floor	0	1	2	3	4
13. Walking on a flat surface	0	1	2	3	4
14. Getting in and out of a car, or on or off a bus	0	1	2	3	4
15. Going shopping	0	1	2	3	4
16. Putting on your socks or stockings	0	1	2	3	4
17. Rising from bed	0	1	2	3	4
18. Taking off your socks or stockings	0	1	2	3	4
19. Lying in bed	0	1	2	3	4
20. Getting in or out of a bath	0	1	2	3	4
21. Sitting	0	1	2	3	4
22. Getting on or off the toilet	0	1	2	3	4
23. Performing heavy domestic duties	0	1	2	3	4

24. Performing light domestic duties	0	1	2	3	4

Appendix D

Outliers tables

Table 6 Outliers identified based on the utility from the EQ-5D-5L utility index values

Study ID	Age	Demographic	EQ-5D-5L health profile	Eq-5d-5l utility index	WOMAC total scores	Predicted utility score from Womac
Hip007	59.5589	Male	43443	0.205684816	36.04	0.458320919
Hip084	58.05479	Female	41444	0.174430248	42.04850006	0.487615785
Hip131	63.62466	Male	54554	-0.064662124	4.801469803	0.174630389
Knee040	61.7863	Male	43443	0.205684816	36.04	0.458320919
Knee041	65.81096	Male	43443	0.205684816	36.04	0.458320919
Knee047	44.81096	Female	54555	-0.102300456	3.188240051	0.153875545
Knee048	45.54247	Female	54555	-0.102300456	3.188240051	0.153875545
Knee070	69.14795	Female	44443	0.143931465	24.90439987	0.300941073
Knee179	64.2274	Male	54541	0.16088293	22.7852993	0.352262059
Knee180	65.11233	Male	54541	0.16088293	22.7852993	0.352262059
Knee198	68.51233	Male	55543	0.039785301	28.11179924	0.366017335
Knee250	59.73699	Female	43444	0.082788317	35.0705986	0.412517011

Table 7 . Outlier identified based on the predicted utility values derived from the WOMAC survey.

Study ID	Age	Demographic	EQ-5D-5L health profile	Eq-5d-5l utility index	WOMAC total scores	Predicted utility score from Womac
Hip018	59.65479	Male	11122	0.866965901	90.55000305	0.9065585
Hip059	63.90959	Male	11122	0.866965901	95.27500153	0.903598438
Hip060	68.22466	Male	11122	0.866965901	92.65000153	0.903598438
Hip089	63.64932	Female	11221	0.885146969	97.90000153	0.903598438
Hip131	63.62466	Male	54554	-0.064662124	4.801469803	0.174630389
Knee022	70.4411	Female	21121	0.865736296	81.625	0.899880688
Knee047	44.81096	Female	54555	-0.102300456	3.188240051	0.153875545
Knee048	45.54247	Female	54555	-0.102300456	3.188240051	0.153875545
Knee070	69.14795	Female	44443	0.143931465	24.90439987	0.300941073
Knee073	67.38356	Female	33441	0.345388032	28.86619949	0.429938079
Knee081	78.99452	Female	42221	0.671676063	97.90000153	0.903598438
Knee166	67.41644	Male	41433	0.457124353	21.13380051	0.293478548
Knee179	64.2274	Male	54541	0.16088293	22.7852993	0.352262059
Knee180	65.11233	Male	54541	0.16088293	22.7852993	0.352262059
Knee198	68.51233	Male	55543	0.039785301	28.11179924	0.366017335
Knee220	72.69863	Male	11111	0.948967801	97.90000153	0.903598438
Knee227	69.19726	Male	31331	0.743590266	23.3484993	0.296180264

Knee239	61.51507	Male	11111	0.948967801	93.69999695	0.906627813
Knee250	59.73699	Female	43444	0.082788317	35.0705986	0.412517011
Knee288	59.58082	Female	32241	0.531941441	32.4455986	0.412517011
Knee289	60.11781	Female	32241	0.531941441	32.4455986	0.412517011

Curriculum Vitae

Sanju Valamparampil, BHSc

Education

Master of Science (MSc) September 2020 - Present
Faculty of Health Sciences, Health and Rehabilitation Sciences program, Western University, London, ON

Regulatory affairs (Graduate certificate) September 2019- April 2020
Algonquin college, Perth, ON

Clinical Trials Management (Post degree diploma) September 2018- 2019
Western University, London, ON

Bachelor's in Health Sciences
September 2014 – 2018
Faculty of Health Sciences, Western University, London, ON

Research Experience

Thesis Sept 2020 – April 2022

Western University, London, Ontario

Supervisor : Dr. Jacquelyn Marsh ; Advisory Committee : Dr. Lyndsay Somerville and Dr. Edward Vasarhelyi.

Conducted a study comparing between the agreement of the EQ-5D-5L utility values to those utility values generated from the WOMAC.

Responsible for ethics submissions, and data collection.

Data management Assistant August 2019 - February 2020
Michael G. DeGroote Pain Clinic, Hamilton Health Sciences, Hamilton, ON

- Helped create databases and data quality checks on Redcap
- Created data management logs on excel for tracking programs
- Experienced in using Meditech and Sovera
- Helped create manuals for data entry and data quality process

Clinical Research assistant May 2019 – August 2019
Michael G. DeGroote Pain Clinic, Hamilton Health Sciences, Hamilton, ON

- Helped develop databases, data quality and data entry

- Collected consent from participants of various programs and helped to create data entry manuals and coordinating manuals.
- Ensured the timely entry of data and helped prepare study and information packages

PRESENTATIONS

Poster Presentation

Valamparampil, S. & Marsh, J. Validity of Quality- of-Life Measurement in Economic Evaluations: Agreement between Western Ontario and McMasters Osteoarthritis Index (WOMAC) and the EUROQOL 5D (EQ5D) utility scores. London, ON, 2021.

Teaching Assistant roles

Western University, London, ON

HS 4400B – Advanced Health policy

Professor : Maxwell Smith

HS3400 - Health policy

Professor : Denise Grafton