Investigating the Evolution of Model-Based Health Economic Evaluations and Changes in Health-Related Quality of Life in Osteoarthritis

by

Ting Zhao

Menzies institute for Medical Research

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy (Medical Studies)

University of Tasmania

December 2021

PhD Research Supervisors

Primary Supervisor

Prof Andrew J. Palmer
Professor of Medical Research
Founding Chair, Health Economics Research Unit Menzies Institute for Medical Research, University of Tasmania
Phone: +61 3 6226 7729
Email: Andrew.Palmer@utas.edu.au

Co-Supervisors

Tania Winzenberg
Professor in Chronic Disease Management
Menzies Institute for Medical Research, University of Tasmania
Phone: +61 3 6226 7770
Email: Tania.Winzenberg@utas.edu.au

Dawn Aitken
Associate Professor, Musculoskeletal Health and Diseases Theme
Menzies Institute for Medical Research
University of Tasmania
Phone: +61 3 6226 7769
Email: Dawn.Aitken@utas.edu.au

and

Dr. Barbara de Graaff
Menzies Institute for Medical Research, University of Tasmania
Phone: +61 3 6226 4262
Email: barbara.degraaff@utas.edu.au

Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by any other person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signature: Date: March 26th, 2021

Statement of authority to access

This thesis may be available for loan. Copying of any part of this thesis is prohibited for two years from the date this statement was signed; after that time, limited copying is permitted in accordance with the Australian Copyright Act of 1968.

Signature: Date: March 26th, 2021

Statement regarding published work contained within this thesis

The publishers of the papers comprising Chapters 3, 4 and 5 hold the copyright for that content, and access to the material should be sought from the respective journals. The remaining non-published content of the thesis may be made available for loan and limited copying in accordance with the Copyright Act 1968.

Signature: Date: March 26th, 2021

Statement of ethical conduct

The research associated with this thesis abides by the international and Australian codes on human experimentation, and the rulings of the Safety, Ethics, and Institutional Biosafety Committees of the University.

Ethical approval for Tasmanian Older Adult Cohort (TASOAC) study was obtained from the Southern Tasmanian Health and Medical Human Research Ethics Committee, and all study participants provided informed consent prior to participation.

Signature: Date: March 26th, 2021

Statement of co-authorship

This thesis includes papers for which Ting Zhao (TZ) was not the sole author. TZ was the first author in the research of each manuscript; however, she was assisted by the co-authors. The names, institutional affiliations and contributions of each co-author are detailed below.

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Candidate: Ting Zhao, Menzies Institute for Medical Research, University of Tasmania.

Author 1: Andrew J. Palmer, Menzies Institute for Medical Research, University of Tasmania.

Author 2: Tania Winzenberg, Menzies Institute for Medical Research, University of Tasmania.

Author 3: Dawn Aitken, Menzies Institute for Medical Research, University of Tasmania.

Author 4: Barbara de Graaff, Menzies Institute for Medical Research, University of Tasmania

Author 5: Graeme Jones, Menzies Institute for Medical Research, University of Tasmania

Author 6: Hasnat Ahmad, Menzies Institute for Medical Research, University of Tasmania

Author 7: Qing Xia, Menzies Institute for Medical Research, University of Tasmania

Contribution of work by co-authors for each paper:

Study 1 reported in Chapter 3:

Zhao, Ting, Hasnat Ahmad, Barbara de Graaff, Qing Xia, Tania Winzenberg, Dawn Aitken, and Andrew J. Palmer. "A Systematic Review of the Evolution of Health-Economic Evaluation Models of Osteoarthritis." Arthritis Care & Research (2020).

Credit author statement

TZ, BdeG, DA, TW and AP conceived and designed the methods. TZ and QX conducted literature search and screening, TZ and HA conducted the data extraction and reporting quality assessment. All authors were involved in the manuscript preparation. All authors read and approved the final manuscript.

Study 2 reported in Chapter 4:

Zhao, Ting, Tania Winzenberg, Barbara de Graaff, Dawn Aitken, Hasnat Ahmad, and Andrew J. Palmer. "A systematic review and meta-analysis of health state utility values for osteoarthritis-related conditions." Arthritis Care & Research (2020).

Credit author statement

TZ, BdeG, HA, DA, TW and AP conceived and designed the methods, and critically reviewed/edited the manuscript. TZ and HA conducted literature search, screening and data extraction. TZ conducted the meta-analyses and wrote the manuscript. All authors read and approved the final manuscript.

Study 3 reported in Chapter 5:

Zhao, Ting, Hasnat Ahmad, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Graeme Jones, and Andrew J. Palmer. "Cross-sectional and temporal differences in health-related quality of life of people with and without osteoarthritis: a 10-year prospective study." Rheumatology (2021).

Credit author statement

G.J. designed the TASOAC study that provided the data for our analyses. TZ, AP and HA conceived and designed the initial concept and method for the present study. DA, TW, A.P and BdeG modified and improved the original concept and methods. TZ accomplished the

setting up of the database, performed necessary data analyses and drafted the manuscript. All authors were involved in review of the manuscript and its final approval.

Study 4 reported in Chapter 6:

Zhao, Ting, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Hasnat Ahmad, Graeme Jones, and Andrew J. Palmer. "The impact of comorbidities on health-related quality of life of people with osteoarthritis over ten years" Rheumatology (2021).

Credit author statement

GJ designed the TASOAC study that provided the data for our analyses. TZ, AP and HA conceived and designed the initial concept and method of this manuscript, DA, TW, and BdeG helped in modifying and improving the original concept and methods. TZ accomplished the setting up of database and performed necessary data analyses and drafted the manuscript. All authors were involved in the manuscript review and final approval.

Signature: Date: March 26th, 2021

Candidate, Ting Zhao
Menzies Institute for Medical Research

Signature: Date: March 31st, 2021

Primary supervisor, Professor Andrew J Palmer
Menzies Institute for Medical Research
University of Tasmania

Signature: Date: April 1st, 2021

Director, Distinguished Professor Alison Venn
Menzies Institute for Medical Research

Abstract

Osteoarthritis (OA) is one of the most common chronic joint diseases, and is characterized by joint pain, stiffness, swelling, loss of function and disability, which in turn, negatively impact individuals' health-related quality of life (HRQoL) and incur a substantial socio-economic burden. Currently, there is no cure for OA, but many treatments and approaches are available to help relieve disease symptoms and improve patients' HRQoL. Health economic evaluation models incorporate clinical, health economic and epidemiological data to compare alternative options in terms of both economic costs and clinical effectiveness to identify the interventions that are best value for money. Despite significant progress made in health economic modelling of OA since 1994 (when the first model-based health economic evaluation was performed), existing models are likely to suffer from limitations related to the choice of model input parameters, discount rates, and model health states/events. There is a need to perform a comprehensive review of health-economic evaluation models of OA to identify their key strengths and weaknesses to provide directions for improvement in the current modelling practice.

Health state utility values (HSUVs) measure the strength of a preference for a particular health state, represented as a number between 0 (death) and 1 (optimal health). They are an important input parameter for health economic evaluation models. However, no comprehensive database of OA-related HSUVs for patients with different affected joint sites undergoing different treatments is available to date, which makes the selection of HSUVs to be used in the modelling practice a challenging task. Therefore, it is imperative to generate a HSUVs database for various sub-groups of OA patients to guide HSUVs choices in future health economic modelling of OA interventions. Furthermore, OA and the associated comorbidities (numbers as well as patterns) are linked to worse health outcomes over time. However, the

quantification of the long-term impacts of OA and associated comorbidities on HRQoL has not been achieved to date and is an important research area that needs attention.

This PhD research thesis (comprising of 5 key chapters) aims to fill these evidence gaps by 1) synthesizing the strengths and weaknesses of existing OA health economic evaluation models; 2) generating a HSUVs database for OA-related conditions; 3) investigating the long-term changes in OA people's HRQoL; and 4) evaluating the impact of numbers and patterns of comorbidities on HRQoL and identifying the most prevalent and influential comorbidity patterns that impact HRQoL in people with OA over a ten year period. The following is a brief overview of each of the five chapters included in this research thesis.

Chapter 1 presents a general introduction to OA and health economics.

Chapter 2 presents a systematic review of all OA health economic evaluation models and evolution of modelling in the field of OA. This is the first study comprehensively reviewing the evolution of health-economic evaluation models of all OA interventions including prevention, core treatments, adjunct non-pharmacological interventions, pharmacological and surgical treatments to identify the key strengths and limitations facing existing OA health-economic evaluation models and provide directions for improvement for current modelling practice. OA health economic evaluation models have evolved and improved substantially over time, with the focus shifting from short-to-medium-term pharmacological decision-tree models to surgical-focused lifetime Markov models. Indirect costs of OA are frequently not considered, despite using a societal perspective. There was a lack of reporting the sensitivity of model outcomes to input parameters including discount rates, OA definition, and population parameters. Whilst the coverage of OA-related adverse events has improved over time, they are still not comprehensively captured in most health economics models of OA.

Chapter 3 presents a systematic review and meta-analysis of HSUVs of people with OA-related conditions. This is the first study comprehensively reviewing OA-related HSUVs and statistically meta-analyzing the HSUVs for different affected joint sites before and after various treatments. The systematic review identified important areas where the current evidence is lacking, namely under-represented multi-attribute utility instruments, geographical locations/ethnicities, affected OA joint sites and treatment options. The meta-analyses generated a HSUVs database for OA patients with different affected joint sites undergoing different treatments that may be applied in future health economic modelling of OA interventions.

Chapter 4 describes the impact of OA on HRQoL in the forms of HSUVs and health-dimension scores and investigates the longitudinal changes in HRQoL of people with OA compared to those without OA using an Australian population-based longitudinal cohort. Compared to participants without OA, HSUs for those with OA were 0.07 (95% confidence interval: -0.09, -0.05) units lower on average over ten years. HSUs for participants with knee and/or hip OA were similar to those with other types of OA at 2.5 years follow-up and then diverged, with HSUs of the former being as much as 0.09 units lower than the latter. Those with OA had lower scores for psychological wellness, independent living and social relationships compared to those without OA. Independent living and social relationships were mainly impacted by knee and/or hip OA with the effect on the former increasing over time. In summary, OA negatively impacts multiple facets of HRQoL, but with different intensity and timing. Interventions to improve HRQoL should be tailored to specific OA types, health dimensions, and times. Support to maintain psychological wellness should be provided irrespective of OA type and duration. However, support to maintain independent living could be more relevant to knee and/or hip OA patients living with the disease for longer.

People with OA are more likely to have comorbidities than people without OA. Comorbidities are associated with worse health outcomes and increased economic burden. Chapter 5 presents the results of investigating the impact of numbers and patterns of comorbidities on HRQoL and identifying the most prevalent and influential comorbidity patterns that impact HRQoL in people with OA over ten years. Having more comorbidities negatively impacted OA patients' long-term HRQoL. Compared with comorbidity-free OA participants, the HSUV of those with 2 or ≥3 comorbidities were -0.07 and -0.13 units lower respectively over ten years, largely driven by reduced scores for independent living, social relationships, and psychological wellness. The types and combinations of comorbidities vary in effect sizes and health dimensions influenced. Comorbidity patterns including 'cardiovascular and non-OA musculoskeletal' were most influential and were associated with up to 0.13 units lower HSUV, mostly through negative impacts on independent living (up to -0.12), psychological wellness (up to -0.08) and social relationship (up to -0.06). The optimal management and prevention of cardiovascular and non-OA musculoskeletal conditions may yield improvements in OA patients' HRQoL. The findings are also helpful to guide the adjustment of HSUVs input for comorbidity numbers and patterns in the future OA health economic models of Australians and similar populations with alternative comorbidity profiles.

In summary, this thesis synthesizes the strengths and weakness of existing OA health economic evaluation models and provides a HSUV database for OA-related conditions, which will be helpful in the development of an improved model and guiding the choice of HSUVs in the future. This thesis also bridged important research gaps by investigating the long-term changes in OA people's HRQoL and evaluating the impact of numbers and patterns of comorbidities on HRQoL and identifying the most prevalent and influential comorbidity patterns on HRQoL in people with OA over ten years, which will help to improve the future management of OA and generate the HSUVs inputs for modelling practice.

Acknowledgements

I have been immensely fortunate to find the opportunity to undertake my PhD in Australia at Menzies Institute for Medical Research, which is a leading medical research institute recognized worldwide for its research excellence. I would like to thank Menzies, and the University of Tasmania for this opportunity and financial support.

I would like to express my deepest thanks to my primary research supervisor, Professor Andrew J Palmer, for his consistent guidance, patience, support and motivation. He always motivated me and helped me whenever I was stuck or struggling with my work. He always provided insightful answers to my questions and offered invaluable advice. It was a great fortune to be able to work with a very encouraging and expert academic like Professor Palmer.

I am also grateful to my co-supervisors, Professor Tania Winzenberg, A/Professor Dawn Aitken, and Dr. Barbara de Graaff, for their guidance and support. Professor Tania Winzenberg is expert in conducting systematic reviews and data synthesis and with her great academic writing skills, her feedback on my work always provided me with new perspectives of thinking and interpretation. I am very thankful to A/Professor Dawn Aitken for her enthusiasm to facilitate and help students. She introduced me to the Musculoskeletal Group and provided me opportunities to communicate with and learn from team members. She was always efficient and responsive in providing her advice and support. I would like to thank Dr. Barbara de Graaff for her help involving me in the Health Economics Group and her support throughout my PhD journey. She was always approachable and ready to help. I wish to be able to continue researching with Professor Palmer, Professor Winzenberg, A/Professor Aitken, Dr. de Graaff and their research teams in the future.

I would also like to thank all the members of Health Economics Group for their support and help including Dr. Hasnat Ahmad, Dr. Stella Lartey, Dr. Julie Campbell, Miss Ingrid Cox, Miss Qing Xia and many others. Dr. Hasnat Ahmad's consistent feedback in an informal setting, complementing the advice and guidance provided by my research supervisors, has been instrumental in helping me achieve my research goals in a timely and efficient manner. I am also thankful to the Musculoskeletal Group members for the opportunities to communicate and learn from them. A special thanks to Mr. Petr Otahal for his expert advice in statistics. I would like to thank Dr Lei Si for providing me with the initial suggestion to do my PhD studies in Australia.

I would like to thank my dear parents, Mr. Chuansheng Zhao and Mrs. Xiuying Zhao, my brother, Mr. Long Zhao and his family, for their love, affection and support, without which, it would be impossible for me to complete my PhD in a foreign country. Also, thanks to all my friends for their help and company throughout my PhD journey. Finally, I would like to say a special thanks to myself for working hard and consistently throughout.

List of publications

Publications arising from work contained in this thesis:

1. Zhao, Ting, Hasnat Ahmad, Barbara de Graaff, Qing Xia, Tania Winzenberg, Dawn Aitken, and Andrew J. Palmer. "A Systematic Review of the Evolution of Health-Economic Evaluation Models of Osteoarthritis." Arthritis Care & Research (2020).

2. Zhao, Ting, Tania Winzenberg, Barbara de Graaff, Dawn Aitken, Hasnat Ahmad, and Andrew J. Palmer. "A systematic review and meta-analysis of health state utility values for osteoarthritis-related conditions." Arthritis Care & Research (2020).

3. Zhao, Ting, Hasnat Ahmad, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Graeme Jones, and Andrew J. Palmer. "Cross-sectional and temporal differences in health-related quality of life of people with and without osteoarthritis: a 10-year prospective study." Rheumatology (2021).

4. Zhao, Ting, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Hasnat Ahmad, Graeme Jones, and Andrew J. Palmer. “The impact of comorbidities on health-related quality of life of people with osteoarthritis over ten years” Rheumatology (2021).

Conference presentations arising from work contained in this thesis:

Accepted Conference Abstracts

1. Ting Zhao, Hasnat Ahmad, Qing Xia, Barbara de Graaff, Dawn Aitken, Tania Winzenberg, Andrew J. Palmer. A Systematic Review of the Evolution of Health Economic Models Developed for Osteoarthritis. Australia Health Economics Society, 2019.

2. Ting Zhao, Hasnat Ahmad, Dawn Aitken, Tania Winzenberg, Barbara de Graaff, Andrew J. Palmer. Health-Related Quality of Life Among People with Osteoarthritis: A Prospective Longitudinal Study. Health Technology Assessment international (HTAi), 2020.

3. Ting Zhao, Hasnat Ahmad, Dawn Aitken, Tania Winzenberg, Barbara de Graaff, Andrew J. Palmer. Health-related Quality of Life Among Older People with Osteoarthritis: A Prospective Longitudinal Study. Australian Rheumatology Association 2020 Annual Scientific Meeting.

4. Ting Zhao, Hasnat Ahmad, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Graeme Jones, Andrew J. Palmer. Estimating the impact of Comorbidities on Health-Related Quality of Life of People with Osteoarthritis. ISPOR Asia Pacific 2020 Conference.

5. Ting Zhao, Hasnat Ahmad, Barbara de Graaff, Dawn Aitken, Tania Winzenberg, Andrew J. Palmer. Assessing the Reporting Quality of Studies of Health Economic Evaluation Models in Osteoarthritis, ISPOR 2021 VIRTUAL

Manuscripts authored/co-authored during candidature, but external to this thesis material:

1. Ahmad H, Campbell JA, van der Mei I, Taylor BV, Zhao T, Palmer AJ. The increasing economic burden of multiple sclerosis by disability severity in Australia in 2017: results from updated and detailed data on types of costs. Multiple Sclerosis and Related Disorders. 2020 Jun 3:102247.

2. Ahmad, Hasnat, Ingrid van der Mei, Bruce V. Taylor, Ting Zhao, Qing Xia and Andrew J. Palmer. Does health-related quality of life differ between people with relapse onset versus progressive onset multiple sclerosis, *Multiple Sclerosis and Related Disorders* 54 (2021): 103138.

List of abbreviations

List of abbreviations	Full term
ABS	Australian Bureau of Statistics
ACR	American College of Rheumatology
AEs	Adverse events
AQoL	Assessment of Quality of Life
BMI	Body mass index
CEA	Cost-effectiveness analysis
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence interval
COX-2	Cyclooxygenase-2
CUA	Cost utility analysis
CV	Cardiovascular
DALYs	Disability-adjusted life years
DES	Discrete event simulation
DMTs	Disease modifying therapies
DSA	Deterministic sensitivity analysis
EQ-5D	EuroQoL-5-Dimension
GI	Gastrointestinal
GPs	General practitioners
HAQ-DI	the Health Assessment Questionnaire Disability Index
HEEs	Health-economic evaluations
HEEMs	Health-economic evaluation models
HRQoL	Health-related quality of life
HSUVs	Health state utility values
HTA	Health Technology Assessment
HUI	Health Utility Index
ICER	Incremental cost-effectiveness ratio
ISPOR	International Society for Pharmacoeconomics and Outcomes research
KL scale	Kellgren–Lawrence scale
LMM	Linear mixed model
MAUIs	Multi-attribute utility instruments
Me	Metabolic disease
MID	Minimum clinically important difference
MRI	Magnetic resonance imaging
Ms	Musculoskeletal
NHS	National Health Survey
NICE	National Institute for Health and Care Excellence
NSAIDs	Nonsteroidal anti-inflammatory drugs
OA	Osteoarthritis
OR	Odd ratio
PBAC	Pharmaceutical Benefits Advisory Committee
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PROSPERO	International Prospective Register of Systematic Reviews
PSA	Probabilistic sensitivity analysis

QALYs	Quality-adjusted life years
QWB	Quality of Well-Being
RACGP	Royal Australian College of General Practitioners
Re	Respiratory disease
RR	Relative risk
SD	Standard deviation
SE	Standard error
SF-6D	Short-Form-6-Dimension
SG	Standard Gamble
TASOAC	Tasmanian Older Adult Cohort
TTO	Time Trade-off
WTP	Willingness-to-pay
WOMAC	Western Ontario and McMaster Osteoarthritis Index

TABLE OF CONTENTS

Table of contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Preface

Chapter 1 presents an overview of this research thesis. The chapter briefly describes the background and rationale for the four studies presented in this thesis, the data platform used, the key aims and objectives of the included studies and the thesis outline.

1.2 Background and rational for the studies

Osteoarthritis (OA) is one of the most prevalent joint diseases worldwide which is characterized by joint pain, stiffness, loss of function and with disease progression leading to disability. Due to the increasing risk factors, for example the ageing population and increasing obesity rates, the prevalence of OA is predicted to rise further (1). Due to its chronic nature and high prevalence, OA poses a significant economic and humanistic burden in terms of direct and indirect costs and negative impacts on individuals' health-related quality of life (HRQoL) (2).

Interventions to ease OA symptoms, postpone or halt disease progression include lifestyle, non-medical, medical, and surgical interventions. In chronic diseases such as OA, healthcare policy decisions should be informed by long-term health and economic evidence. In the field of health economics, health economic evaluation models are commonly used to compare alternative therapeutic options in terms of long-term economic costs and effectiveness to identify the interventions that are best value for money. The development of robust health economic evaluation models requires careful consideration of several important elements, including the nature and extent of the available disease-related evidence and clinical complexities; aims, scope, methodological framework, perspective, and time horizon of the model; and data availability and other factors such as within-model synthesis of evidence. Due to differences in those elements, existing health economic evaluation models have varying scopes, methodological frameworks, structures, divisions of model events, and input data sources.

Health economic evaluation models have been used in the field of OA for a long time and have evolved substantially over time, both because of methodological developments and increasing data availability. These models use clinical and economic data to identify current and upcoming health interventions that are effective and safe, as well as being good value for

money. To date, no study has synthesised the evolution of health economic evaluation models of all OA interventions, with an emphasis on their strengths and weaknesses and study gaps to improve the health economic modelling of OA.

Health state utility values (HSUVs) are a key input parameter to many health economic evaluation models. However, no systematic review and meta-analysis has yet been conducted to summarize estimates of OA-related HSUVs. The generation of a HSUVs database for OA patients with different affected joint sites undergoing different treatments is imperative to guide HSUVs choices in future OA health economic models. Finally, HRQoL impacts of OA have not been well researched in Australia, so no locally derived reliable estimates of HSUVs are available to be used as inputs to a health economic model for Australians with OA.

The correct implementation of multi-state health economic evaluation models not only relies on an appropriate and comprehensive model design but also on the availability of reliable and robust estimates of HSUVs and other key input parameters (e.g., disease management costs, and transition probabilities of disease progression). Therefore, the first key aim of this research thesis is to investigate the evolution, strengths and weaknesses of health economic evaluation models of OA to provide directions for improvement in OA health economic modelling practices internationally and in Australia. The second key aim is to ensure the availability of a comprehensive database of reliable and robust HRQoL inputs to be used in future health economic models of OA, internationally and in Australia. A final aim of this thesis is to address the paucity of data on HSUVs impacts of OA in Australia. The following sections describe the rational of each of the four studies included in this thesis.

1.2.1 Rational for study one (reported in chapter 3)

Health economic evaluation models have been widely used in the field of OA to facilitate decision makers in identifying the interventions that are best value for money. Since the publication of first health economic evaluation model of OA in 1994 (3), numerous modelling studies have been conducted, particularly in western developed nations (4). OA models differ from each other in multiple ways and have developed over time, both because of methodological advances and the availability of better data. Understanding the general characteristics of OA models, such as the target population, model structure, and type of questions answered, can provide future investigators with a systematic and broad view of the modelling landscape that has been adopted in the field and lead to significant improvements in OA health economic modelling practice. Previous systematic reviews have investigated the

OA health economic evidence (4, 5); however, the scope of those systematic reviews was often limited to a specific treatment type and none of the previously conducted systematic reviews have synthesised the evolution of health economic evaluation models of all OA interventions.

Thus, there is a need to comprehensively investigate the evolution of health economic evaluation models used for all forms of OA interventions, with an emphasis on their strengths and weaknesses and study gaps to inform the future development of robust OA health economic models of various forms and treatments of OA, internationally and in Australia.

1.2.2 Rational for study two (reported in chapter 4)

Health-state utility values (HSUVs) are one of the key input parameters for health economic evaluation models to calculate quality-adjusted life years (QALYs). HSUVs measure an individual or society's preference for a particular health state and are represented as a number between 0 (death) and 1 (perfect health). The HSUV estimates for the same health states may differ between studies due to several factors, for example, the difference in utility measurements, the choice of respondents, sample size and quality of studies. There is a growing interest in meta-analytic methods that pool HSUVs collected across a number of studies (6). These methods are capable of generating precise estimates of the measure of interest and estimates of uncertainty surrounding the estimated values. To date, no systematic review and meta-analysis has summarized the estimates of OA-related HSUVs.

Thus, a systematic review and meta-analysis is needed to 1) comprehensively review OA-related HSUVs to identify the areas where the current evidence is lacking; and 2) to meta-analyse OA-related HSUVs based on different affected joint sites, treatments, and utility measures to generate a HSUVs database to guide input choices in future health economic modelling practice.

1.2.3 Rational for study three (reported in chapter 5)

As has been noted in Section 1.1.2, HSUVs measure the individual's or society's preference for a particular health state and are likely to differ from one population to another. Studies assessing the impact of OA on HSUVs have mostly been undertaken in Europe and the United States, and there is a paucity of data on HSUVs impacts of OA in Australia. For example, the latest study estimating the impact of OA on HSUVs based on Australian data was published in 2015, however, this study only focused on knee and/or hip OA and included only a small number (n=21) of community participants (7). Moreover, no study has investigated the

longitudinal changes in the HSUVs of Australians with OA compared with those without OA, which is important given the chronic nature of this health condition. Finally, the investigation of the physical and psychosocial health drivers of longitudinal changes in HSUV scores of Australians with OA compared with those without have not been performed.

To bridge these important evidence gaps, a study is needed to investigate the cross-sectional and longitudinal difference in HSUVs and health-dimension scores of Australians with OA compared with those without. The HSUVs estimates are helpful to guide input choices in the future health economic evaluation models of OA for Australian and similar populations.

1.2.4 Rational for study four (reported in chapter 6)

People with OA are more likely to have comorbidities than people without OA. Comorbidities are one of the important factors that lead to the lower HRQoL of OA people (8). Some studies have investigated the impact of comorbidities on HSUVs of OA people, with most using the total number of comorbidities (comorbidity count) as the comorbidity measure. Whilst comorbidity count provides an understanding of the potential additive impact of multiple comorbid disease on HSUVs, its reliance on an implicit and rather unrealistic assumption of a uniform impact of all comorbidities makes this comorbidity classification scheme less useful. Additionally, no previous study has investigated the long-term effects of comorbidities on HSUVs and individual health dimensions, which is important to be investigated given the chronic nature of this health condition.

Thus, a research study is needed to identify the most prevalent OA-related comorbidities and to investigate the impact of numbers and patterns of comorbidities on long-term HSUVs and individual health-dimension scores of people with OA. Findings from this research can also guide the adjustment of HSUVs inputs for comorbidity patterns in the future OA health economic models of Australians and similar populations.

1.3 Introduction of the data platforms used for the studies in this thesis

Study three (reported in chapter 5) and study four (reported in chapter 6) of this thesis were based on the Tasmanian Older Adult Cohort (TASOAC) study. TASOAC is a prospective, population-based study aimed at identifying the environmental, genetic, and biochemical factors associated with the development and progression of osteoarthritis at multiple joint sites

(e.g., hand, knee, hip, and spine). Participants between the ages of 50 and 80 years were randomly selected from the electoral roll in Southern Tasmania with an equal number of men and women. Eleven-hundred participants were initially enrolled in the study, and 1,099 attended a baseline clinic between March 2002 and September 2004. At the phase 2 follow-up (approximately 2.7 years from the baseline), 875 participants provided the data. Phase 3 follow-up data was collected from 769 participants approximately 5 years from the baseline. Phase 4 follow-up data was collected from 568 participants approximately 10 years from the baseline. Participants provided information about their sociodemographic characteristics, OA diagnosis, affected OA joint sites, HRQoL, comorbidity profiles and others. Full details of the TASOAC data used in studies 3 and 4 is provided in the methods sections of Chapters 5 and 6.

1.4 The main objectives of the studies included in this thesis

With the rationale of Section 1.1 in mind, the key objectives of each of the four studies included in thesis are listed below.

Study one (reported in Chapter 3)

To comprehensively synthesise the evolution of health-economic evaluation model for all OA interventions (i.e.: preventions, core treatments, adjunct non-pharmacological interventions, pharmacological and surgical treatments), with an emphasis on their strengths and weaknesses and study gaps to inform the future development of robust OA health economic models of various forms and treatments of OA, internationally and in Australia.

Study two (reported in Chapter 4)

1) To systematically review OA-related HSUVs studies to identify the areas where the current evidence is lacking.

2) To meta-analyse the OA-related HSUVs to generate a HSUVs databases for OA patients with different affected joint sites undergoing different treatments to guide HSUVs choices in future health economic models of OA in a global context.

Study three (reported in Chapter 5)

To address the paucity of Australian data on cross-sectional and temporal HRQoL impacts of OA and to ensure the availability of locally driven HSUVs inputs to be used in future health economic models of OA for Australian and similar populations.

Study four (reported in Chapter 6)

To examine the long-term contribution of numbers and patterns of comorbidities on HRQoL (in terms of HSUVs and dimensional scores) of Australians with OA.

1.5 Outline of the thesis

The overall aim of this research thesis is to improve the current OA health economic modelling practices in terms of the model design and quality/variety of HSUV model inputs, internationally and in Australia. Chapter 1 presented here provides a brief overview of the thesis (including the background and rational for the four included studies, data platform used, and key objectives of each included studies). Chapter 2 provides a detailed introduction of OA, health economics, and the application of health economics in the field of OA. It starts by introducing the OA definitions used in the literature, OA risk factors, comorbidities of OA, OA prevalence, and OA management, followed by the introduction of health economics, health economic evaluation models and associated key concepts including the health-related quality of life (HRQoL). The last section of chapter 2 provides an overview of health economics of OA including the health economic evaluation models used for OA, and the studies of OA impacts on HSUVs. Chapter 3 documents study one that provides the systematic review of OA health economic evaluation models. Chapter 4 documents study two that provides the systematic review and meta-analysis of OA-related HSUVs. Chapter 5 documents study three that assesses the cross-sectional and longitudinal changes in HSUVs and health dimension scores of OA compared with those without using Australian data. Chapter 6 documents study four that investigate the long-term impacts of comorbidity count and patterns on HSUVs and health dimension scores of people with OA using Australia data. Chapter 7 then summarises the key findings of the included studies and discusses their key strengths and weaknesses. A final section provides direction for future research, based on the findings from this thesis as well as those from the existing literature in the field.

1.6 Reference

1. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. The Lancet. 2019;393(10182):1745-59.

2. Xie F, Kovic B, Jin X, He X, Wang M, Silvestre C. Economic and humanistic burden of osteoarthritis: a systematic review of large sample studies. Pharmacoeconomics. 2016;34(11):1087-100.

3. McKell C, Stewart A. Cost-minimisation analysis comparing topical versus systematic NSAIDS in the treatment of mild osteoarthritis of the superficial joints. British Journal of Medical Economics. 1994;7:137-46.

4. Kamaruzaman H, Kinghorn P, Oppong R. Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature. BMC musculoskeletal disorders. 2017;18(1):1-17.

5. Wielage RC, Myers JA, Klein RW, Happich M. Cost-effectiveness analyses of osteoarthritis oral therapies: a systematic review. Applied health economics and health policy. 2013;11(6):593-618.

6. Petrou S, Kwon J, Madan J. A practical guide to conducting a systematic review and meta-analysis of health state utility values. Pharmacoeconomics. 2018;36(9):1043-61.

7. Ackerman IN, Bucknill A, Page RS, Broughton NS, Roberts C, Cavka B, et al. The substantial personal burden experienced by younger people with hip or knee osteoarthritis. Osteoarthritis and cartilage. 2015;23(8):1276-84.

8. Wesseling J, Welsing PM, Bierma-Zeinstra SM, Dekker J, Gorter KJ, Kloppenburg M, et al. Impact of self-reported comorbidity on physical and mental health status in early symptomatic osteoarthritis: the CHECK (Cohort Hip and Cohort Knee) study. Rheumatology. 2013;52(1):180-8.

Chapter 2 Introduction of osteoarthritis and health economics

2.1 Preface

Chapter 2 provides a background on the topics that are important to this thesis, including the definition of osteoarthritis (OA), its types, diagnosis, management, and prevalence. The chapter also provides a comprehensive overview of the key topics related to health economic evaluation models of OA, including the types and components of popular health economic evaluation models, their uses, and a few key inputs to these models including HRQoL, HSUVs, health dimension scores and others. A summary of the previous literature surrounding health economic evaluation models of OA and the key model input parameter of HSUVs is also presented at the end of Chapter 2 to highlight the key evidence gaps that this research thesis aims to bridge.

2.2 Introduction to osteoarthritis

Osteoarthritis (OA) is the most common joint disease, characterized biologically by degeneration of articular cartilage, limited intraarticular inflammation with synovitis, and changes in peri-articular and subchondral bone; and symptomatically by joint pain, stiffness, swelling, loss of function, and disability (1, 2). OA impacts physical as well as psychosocial health aspects of OA patients, with pain being the key problem for most people with OA (2).

2.2.1 Diagnosis of OA in research

OA case is commonly defined as radiographic OA, clinical (radiographic + symptomatic) OA or self-reported OA in the literature (3). Radiographic OA considers radiographic images-based pathophysiological joint signs (4). It can be determined using alternative scoring systems (e.g., Kellgren–Lawrence (KL) scale, Joint space width method, Croft index, American college of rheumatology criteria) (5-8), and the assessment of individual radiographic features. The radiographic definition is commonly used for knee and hip OA and less reliable for other joint sites (9).

Clinical OA considers both radiographic and joint symptoms related to the pathology (i.e., pain, stiffness and loss of function) (3). The concept of clinical OA is attractive to clinicians as it combines information on structural damage with the patient's symptoms. The most

commonly used clinical criteria for defining OA is developed by the American College of Rheumatology (ACR) in early 1990's (10). The ACR's clinical approach to defining OA considers medical history, laboratory test results and physical examination.

Previous studies have commonly used patients' self-reports of OA diagnosis to define OA cases (11). Self-reported OA does not require a clinical examination or radiographic imaging, making this a good option for large population studies and those conducted in community settings (12). However, individuals may not able to distinguish OA from other forms of arthritis (e.g., rheumatoid arthritis), or from other non-arthritic diseases (e.g., osteoporosis) (9). Notably, the agreement between radiographic, clinical and self-reported diagnoses of OA has been investigated previously and found modest (13).

2.2.2 Risk factors for OA

The risk factors for OA have been well researched and established, with age being the most evident risk factor (14). The incidence of OA increases with age due to the cumulative exposure to various risk factors and biological age-related changes in the joint structures (15, 16). Other common OA risk factors include female sex, higher body mass index (BMI), joint injury, joint deformity, occupational factors, sports participation, and genetics. Female sex and obesity are shown to be moderate to strong risk factors for knee OA, whereas their effects for hip OA are found to be less pronounced (17, 18). Joint deformity is a well-established risk factor for OA, for example, knee malalignment has been shown as a moderate to strong risk factor for knee OA (19), and hip deformities (e.g., cam deformity, acetabular dysplasia) are considered moderate to strong risk factors for hip OA (18). Positive association has been found between physical workload/occupational factors and OA. Whilst heavy lifting and frequent kneeling are shown as the risk factors for knee OA (20-23), workers in farming or the construction industry are shown to increase the risk of hip OA (20, 24). A positive association between the intensity of physical sporting activities and hip OA has also been reported (25). Finally, the contribution of genetics in OA is estimated to be between 40% and 80%, with a stronger contribution in hand and hip OA than knee OA (26).

The joint-specific effects of risk factors for development of OA have been summarized in the literature (27) and are shown in Table 2.1.

Table 2.1 The risk factors for development of osteoarthritis

Risk Factor	Hip OA	Knee OA	Hand OA
Obesity	(+)	+	(+)
Age	+	+	+
Female sex	(+)	+	+
Ethnicity (vs Caucasian)			
Chinese	−	+	−
Genotype	+	+	+
Bone mineral density	+	+	+
Grip Strength			+
Quadriceps Strength		(−)	

Note: +, good evidence of increased risk; (+), weak evidence of increased risk; blank, inconsistent or no evidence of increased risk; (−), weak evidence of protective effects; −, good evidence of protective effects. Source: Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. British medical bulletin. 2013;105(1):185-99

2.2.3 Comorbidities of OA

Comorbidity in a persons with OA is defined as any additional disease or disorder occurring concomitantly with OA (28). Comorbidities are common, especially in older and obese people, and are associated with clinical complexity, worse health outcomes, and increased economic burden (29). Due to the overlapping risk factors (e.g.: old age, and obesity) and other factors including the adverse effects of OA interventions (30), people with OA are more likely to have comorbidities than people without OA. A recently published study has systematically reviewed and meta-analysed the comorbidities in individuals with OA compared to those without (31). Forty-two studies were included in that review and the pooled prevalence of any chronic condition in OA patients was 67% (95% confidence interval [CI] 58–74), the pooled prevalence ratio for any comorbidity was 1.21 compared with those without OA, with leading systems that comorbidities occur were cardiovascular (35%), musculoskeletal (34%), neurologic (30%), and upper gastrointestinal (19%) and the leading comorbidities were hypertension (50%), dyslipidemia (48%), and back pain (33%), followed by thyroid disorder (26%) and depression (17%). (Figure 2.1)

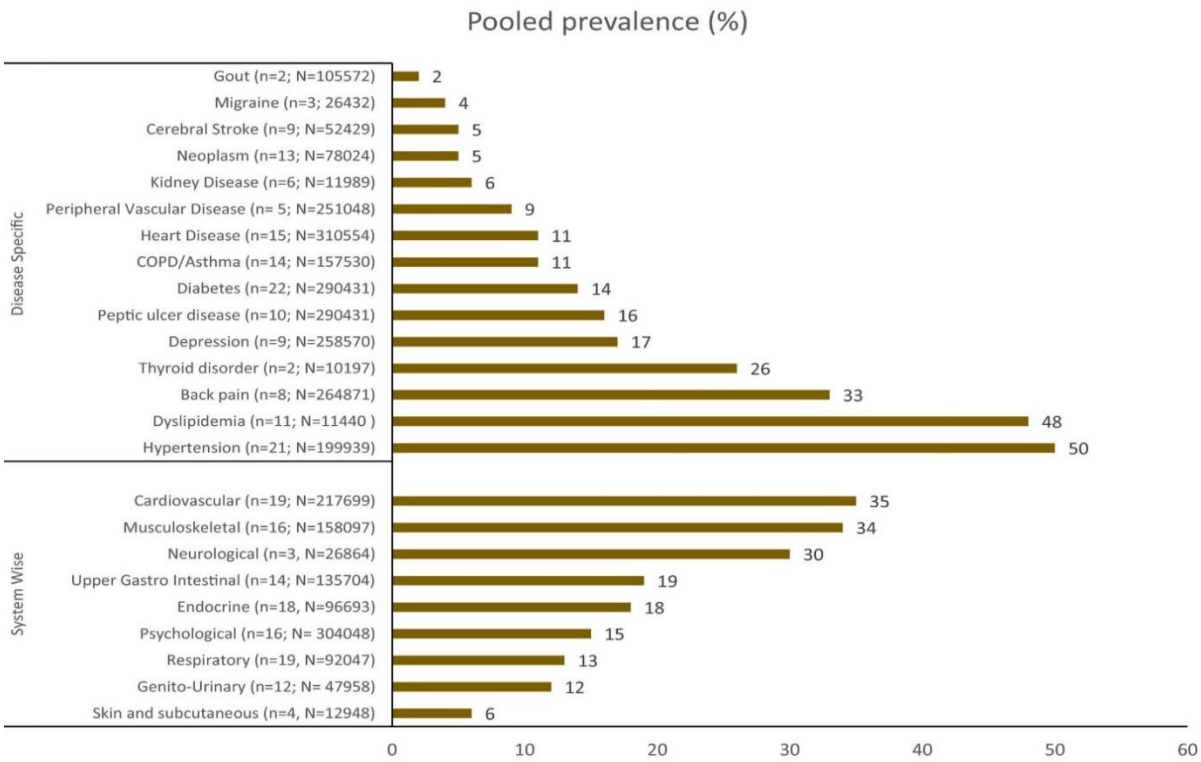


Figure 2.1. Prevalence (%) of comorbidities in individuals with osteoarthritis (disease and system specific). n = number of studies; N = number of participants; COPD = chronic obstructive pulmonary disease. Source: Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in Osteoarthritis: A systematic review and meta-analysis of observational studies. Arthritis care & research. 2020;72(7):991-1000

2.2.4 The prevalence of OA

The prevalence of OA may vary between studies due to differences in the OA definition used, as well as in the age, sex and geographic distributions of the study populations (3). A systematic review has been conducted for the prevalence of knee, hip and hand OA and provided the prevalence in terms of joint sites, age, sex, and countries (3, 32). It was indicated that the prevalence of radiographic OA is higher than symptomatic OA, the prevalence of knee and hand OA were higher than hip OA. Within each joint site, the prevalence of self-reported OA was similar with symptomatic OA. In terms of sex, knee or hand OA was more prevalent in women than in men, especially for symptomatic OA.

The overall prevalence of OA by joint sites has been meta-analysed, with the results showing the highest prevalence for hand OA (43.3%, 95% CI 42.7-42.9), followed by knee OA (23.9%, 95% CI 23.6-24.2) and hip OA (10.9%, 95% CI 10.6-11.2). In terms of sex distribution, hand and hip OA have similar prevalence for males and females, but knee OA was more prevalent in females than males (3).

OA is the most common form of arthritis in Australia. According to the Australian Bureau of Statistics (ABS) 2017–18 National Health Survey (NHS), 2.2 million (9.3%) Australians have OA (33) which represented 62% of all arthritic conditions in 2017–18. Although, OA affects people of all ages, the prevalence increases sharply from the age of 45 years. For instance, 22% and 36% of Australians over the ages of 45 years and 75 years suffer from OA, respectively. The number of Australians with OA is predicted to increase to 3.1 million by 2030 both because of population ageing and increasing rates of obesity in Australia, which will only exacerbate the health economic burden of this health condition (34).

2.2.5 Management of OA

Numerous guidelines for OA management have been developed (35-37). These guidelines have been systematically reviewed and critically evaluated (38, 39) to compare the degree of agreement between alternative sets of guidelines. The widely recommended non-pharmacologic management options may include education/self-management, exercise, weight loss, walking aids, and thermal modalities. The most recommended pharmacologic treatments include the first line acetaminophen/paracetamol and the second line nonsteroidal anti-inflammatory drugs (NSAIDs). Intra-articular corticosteroids were generally recommended for hip and knee OA. Joint replacement was recommended for late-stage OA; arthroscopy with debridement was not recommended for symptomatic knee OA. Controversy remains about the use of acupuncture, knee braces, heel wedges, intra-articular hyaluronans, and glucosamine/chondroitin (38).

Following the clinical practice guidelines, in general, a stepped-care approach is recommended for OA management, which means treatments for OA are arranged in order which begin with the safest and least invasive therapies before proceeding to more invasive, expensive therapies (Figure 2.2). Individual needs, risk factors and preferences modulate this approach. According to the recent RACGP guideline, education, exercises, and weight loss (to those overweight or obese) are recommended to be offered throughout the treatment of OA. Whilst all OA patients are recommended to receive at least some treatment from non-pharmacological and pharmacological treatments, surgical management is suggested to be reserved for those who do not improve with behavioural and pharmacologic therapy, and those with interactable pain and loss of function.

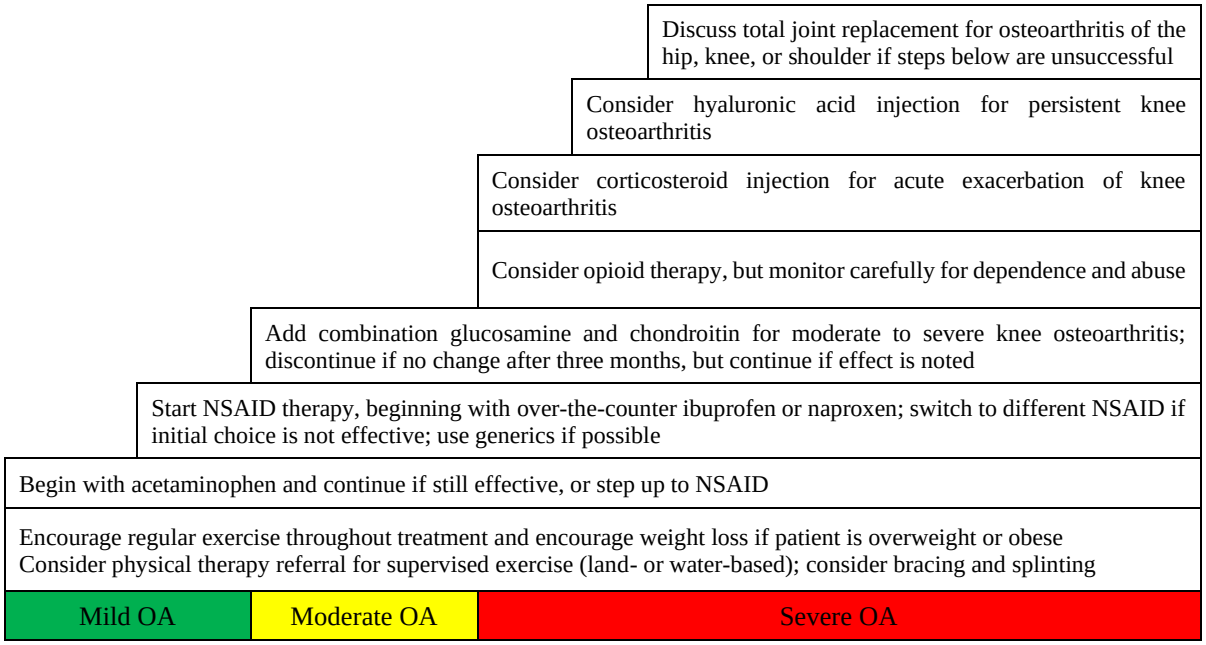


Figure 2.2. Recommended stepped-care approach for the treatment of OA. Source: Sinusas K. Osteoarthritis: diagnosis and treatment. American family physician. 2012 Jan 1;85(1):49-56.

In Australia, the Royal Australian College of General Practitioners (RACGP) has developed the guideline for the non-surgical management of knee and hip OA for use in the primary care setting by general practitioners (GPs) (40). The guideline presents the best available, current scientific evidence for OA interventions, covering all interventions other than joint replacement for the hip and knee. This guideline provides advice and recommendations for the management of people with knee and/or hip OA. The RACGP OA management guideline has a strong focus on self-management and non-surgical treatments to improve the health of people with knee and/or hip OA. However, despite these clear instructions, the non-pharmacological interventions including exercise, weight loss, physiotherapy remain underutilised/under-recommended by Australian GPs (41).

One of the important concerns related to the treatment of OA is the treatment side effects. Evidence for the side effects of OA treatments has been mainly investigated in pharmacological therapies (39). Oral NSAIDs have been shown to associate with 3-5 times higher risk of gastrointestinal (GI) side effects when compared with placebo or non-exposure (42), whereas topical NSAIDs resulted in no more GI adverse events than placebo [relative risk (RR)= 0.81, 95% CI 0.43, 1.56] (43) or non-exposure [odd ratio (OR)= 1.45, 95% CI 0.84, 2.50] (44). Cyclooxygenase-2 (COX-2) selective NSAIDs targeted to manage OA pain were associated with increased risks for upper GI as well as cardiovascular (CV) adverse events. (45). The use of misoprostol was associated with an increased risk of diarrhoea (RR=1.81, 95% CI 1.52, 2.61) (46).

2.3 Introduction to health economics

2.3.1 Health economics

Health economics is defined as the application of economic theory, models and empirical techniques to the analysis of decision making by individuals, health care providers and governments with respect to health and health care (Morris, Devlin & Parkin, 2007) (47). Health economics underpins decision-making in health policy and practice and is essential to generate evidence to ensure rational use of scarce healthcare resources (49). It was recognised as a discipline in 1963 with the publication of research article, entitled: Uncertainty and the Welfare Economics of Medical Care (48). Since then, it has become increasingly important due to the growing healthcare sector size and budget, increasing health consumer concerns and the substantial healthcare and healthcare industry linked economic elements (Figure 2.3) (49). In today's world, health economic evaluations serve as the foundation to generate evidence to facilitate healthcare decision making in many countries (50). For example, in Australia, the health economic evaluation has been used by the Pharmaceutical Benefits Advisory Committee (PBCA) to meet the legislative requirements in making funding recommendations for drugs to government.

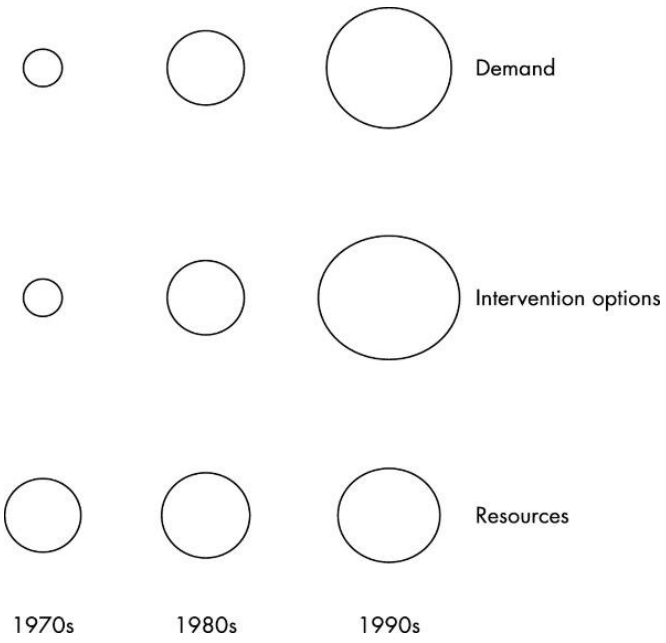


Figure 2.3. Increasing demands and intervention options on limited resources (area of each circle reflects size of each variable). Source: Kernick DP. Introduction to health economics for the medical practitioner. Postgraduate medical journal. 2003;79(929):147-50

2.3.2 Health economic evaluations

Economic evaluation is defined as "the systematic appraisal of costs and benefits of projects, normally undertaken to determine the relative economic efficiency of programs." In simple words, economic evaluation is the understanding and use of economic evidence in decision making. It is a comparative analysis of alternative options in terms of both their costs and outcomes incorporating clinical, epidemiological, and economic data. Health economic evaluations have been widely used to help decision makers (including clinicians, governments, payers, patients and other stake holders) to identify the health interventions that are effective as well as being good value for money. In some countries, the use of cost-effectiveness analysis has been institutionalised for decision making for the public subsidies for medicine purchase (51, 52). There are two levels of economic evaluations: 1) partial and 2) full. Partial economic evaluation measures program or disease costs/outcomes but does not involve a comparison with alternative options and does not relate costs to outcomes (e.g.: cost-of-illness analysis and program cost analysis). On the other hand, full economic evaluations compare two or more public health interventions through the examination of costs of inputs and outcomes. Full economic evaluations include cost-benefit, cost-effectiveness, cost-utility, cost-consequences, and cost-minimization analyses (Table 2.2).

Table 2.2. Types of health economic evaluations

Type	Description	Cost measurement	Outcome measurement
Partial economic evaluations			
Cost of illness	disease economic burden	$	–
Program cost analysis	Net program cost	$	–
Full economic evaluations			
Cost-benefit analysis	Compares different programs with different outcomes (e.g., health vs. other area)	$	$
Cost-effectiveness analysis	Compares interventions with the same outcomes	$	Single "natural" unit outcome measure
Cost-utility analysis	Compares interventions with different health outcomes	$	Multiple outcomes—life-years adjusted for quality-of-life
Cost-minimization analysis	Compares the costs of alternative interventions that have equal effects	$	Equivalence demonstrated or assumed in comparative groups

Cost-consequences analysis	Lists separately all the direct and indirect costs and catalogues different outcomes of all alternatives, with no specific preference for one costing approach/outcome measure (as is the case for cost-effectiveness analysis or cost-utility analysis)	$	Multi-dimensional listing of outcomes

The most commonly used form of economic evaluation in health sector is the cost-effectiveness analysis (CEA) (53). CEA compares the relative costs and outcomes of different interventions by calculating the incremental cost-effectiveness ratio (ICER), which is calculated by dividing the difference un costs between two interventions by the difference in their effectiveness and represents the incremental costs associated with one additional units of effectiveness gained. The calculated value of ICER is then compared with a ceiling ratio (more commonly known as willingness-to-pay [WTP] threshold) to draw conclusions about the cost-effectiveness of an intervention of interest. The CEA is the same as cost utility analysis (CUA) when the unit of effectiveness measure is quality adjusted life years (QALYs).

2.3.3 Health economic evaluation models

The costs of an intervention and its associated outcomes can be recorded (or estimated) by a clinical trial-based evaluation. However, it has been well recognised that economic evaluations based on single clinical trials suffer from several challenges as they may fail to consider all important comparators and relevant inputs or be conducted over a short time which may not be enough to predict long-term disease outcomes (54). Additionally, by relying on a single trial, we are likely to ignore important evidence from meta-analyses, other trials, and observational studies. As a result of these limitations, there has been extensive consideration of the appropriate design and analysis of trials for economic evaluation models (55).

The modelling studies use initial values obtained from clinical trials, and then mathematical models to synthesize all downstream information regarding treatment process, costs, and outcomes. The roles and applications of modelling are to (56):

1. extend single trial results,

2. answer policy questions by considering multiple evidence,

3. generalise results from one context to others,

4. inform research design; and

5. model analyses related uncertainties.

The most used health economic evaluation modelling methods include decision trees, multi-state Markov models, and discrete-event simulation models.

2.3.3.1 Decision tree model

Decision tree models use distinct branches to represent all possible pathways for patient(s). It consists of a series of 'nodes', with each node taking the form of a 'choice' between alternative interventions or a 'probability' of occurrence or non-occurrence of an event. All branch segments and ends are assigned with relevant outcomes and costs that are combined using branch possibilities. The tree is 'rolled back' to a decision node, where costs and outcomes associated with alternative treatment options are compared. Decision trees are generally appropriate if the time horizon of the model is short, events are non-recurrent and the mortality of patients is similar across comparators (57). Figure 2.4 shows an example of decision tree model for OA.

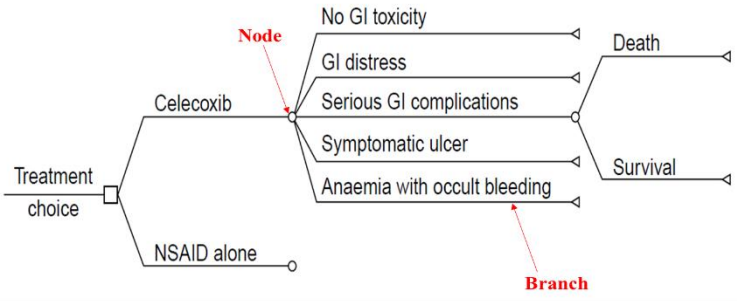


Figure 2.4 The decision tree model example. Source: Zabinski RA, et al. An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada. Pharmacoeconomics. 2001 Nov;19(1):49-58.

2.3.3.2 Markov model

The Markov model is the most common form of health economic evaluation models of healthcare interventions at present. They are suited to decisions where the timing of events is longer, and events are recurrent. Therefore, these models are appropriate where the strategies being evaluated are of a sequential or repetitive nature. A Markov model comprises a finite set of health states. The states are such that at any given time interval, the individual will be in only one health state. The number and nature of the states are governed by the decision problem. Individuals move ('transition') between disease states as their condition changes over time or remain in the same state. Time itself is considered as discrete time periods called 'cycles' (typically one year), and the likelihood of patients moving from one health state to another over one cycle is represented as 'transition probabilities'. Rewards, such as costs, life-

years, or QALYs, are assigned to each health state and earned at the end of each cycle and can be compared between cohorts receiving alternative treatments.

There are two types of commonly used Markov model: Markov cohort models and individual-level Markov microsimulation models. A Markov cohort model runs an infinitely large group of individuals, assumed to have identical characteristics, all at once through the model and provides an exact deterministic calculation (i.e.: a precisely determined outcome given a set of initial conditions and parameters). The transition probabilities of disease progression in Markov cohort models are assumed to only depend on the current health state at any given cycle and not on the history of previous events. This is known as Markovian (memoryless-ness) assumption, an inherent limitation facing such models. Individual level Markov model randomly walks one individual at a time through the model, calculating that patient's outcome values. Outcomes are generated for each individual and are used to estimate the distribution of an outcome for a sample of potentially heterogeneous individuals. As everyone is different, the uncertainty is at individual level in Markov microsimulation models. Also, because these models are based on Monte Carlo simulation, any two trials using the same inputs can return very different results due to randomization at the chance nodes in the model. An advantage of individual level Markov model is its ability to model individual characteristics by more accurately capturing individual clinical pathways and to evaluate dynamic intervention strategies (where future decisions depend on current and past patient characteristics). However, compared with cohort Markov models, individual level Markov models are computationally intensive, highly data-required and more difficult to debug. Figure 2.5 shows an example of Markov model.

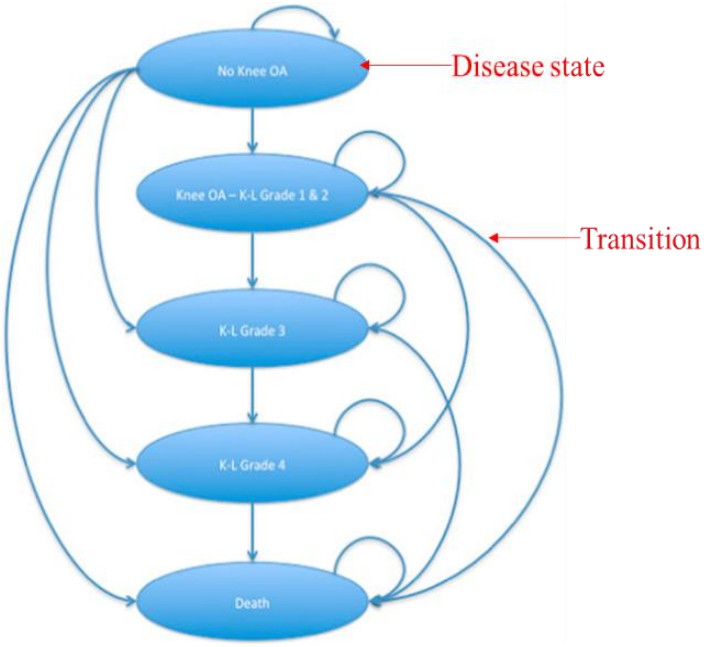


Figure 2.5 The Markov model example. Source: Karmarkar TD, et.al. A fresh perspective on a familiar problem: examining disparities in knee osteoarthritis using a Markov model. Medical care. 2017 Dec;55(12):993.

2.3.3.3 Discrete event simulation

Discrete event simulation (DES) models are characterized by their ability to represent complex behaviour within, and interactions between individuals, populations, and their environments. A DES model simulates individual patient's experience over time, keeping a track of patient's life course of events and their impacts. In comparison with aggregate models without interaction (e.g., decision trees or Markov models), DES can be more advantageous to model complex systems at the individual level instead of cohort level. Unlike Markov models that are focused on health states, DES is conceptualised around events. In a DES, the model moves forward in time to the point at which the next event is experienced, which can provide more flexibility with respect to analysing input data. DES modelling approach also come with the ability to model decision making under circumstances of constrained or limited resources. In DES inter-health states movements may occur at varying times, so time-to-event distributions are important. DES models are constructed for using individual patient level data, and findings are then aggregated over time to produce estimates for the whole cohort of patients. Events and health states are associated with resource use and HSUVs, and event probabilities are based on individual patient characteristics. DES is suitable for modelling complex health conditions with multiple types of events/health states (e.g., diabetic complications) or situations where the patient's history is important (58). The DES has been recommended as a preferred technique for health economic evaluations (59) due to its advantages over Markov

models including the ability to model queuing for limited resources, capture individual patient histories, accommodate complexity and uncertainty, represent time flexibly, model competing risks, and accommodate multiple events simultaneously. However, it's implementation can be more complex and challenging due to the potential for model overspecification, increased data requirements, specialized expensive software, and increased model development, validation, and computational times, as well as decreased transparency.

2.3.4 Some concepts in health economic evaluations

2.3.4.1 Perspective

The perspective is the point of view adopted when deciding which types of costs and health benefits are to be included in an economic evaluation. The commonly used perspectives in the literature include societal perspective, healthcare system perspective, patient's perspective and hospital perspective. The broadest perspective is societal perspective which considers a full range of direct and indirect costs associated with different interventions, particularly the productivity losses arising from patients' inability to work or those associated with absenteeism/presenteeism. There is no certain answer to the questions which perspective is better or which perspective should be adopted, the choice of perspective should therefore depend on the purpose of economic evaluation, the availability of information, the local culture, the political landscape and other factors (60). For instance, Polimeni *et al*. (2013) recommends societal perspective for healthcare interventions for illnesses where morbidity and long-absences from work are probable (61), whereas the UK National Institute for Health and Care Excellence (NICE) recommends a perspective of 'National Health service and personal and social services', recognising that the societal perspective may bias against those not in work due to aging or poor health.

2.3.4.2 Time horizon

The time horizon used for an economic evaluation is the duration over which health outcomes and costs are calculated. The choice of time horizon depends on the nature of the disease, the intervention under consideration and the purpose of the analysis (62). For example, longer time horizons are more appropriate for chronic conditions associated with on-going medical management in which the costs and benefits occur over time. In comparison, a shorter time horizon is more suitable for acute conditions of short-term nature.

2.3.4.3 Discount rate

Costs and health outcomes considered in health economic evaluations often occur at different points in time. In the context of economic evaluation, the future costs and health outcomes are usually valued less than their current value. Therefore, economic evaluations need to adjust the future costs and outcomes for the values if they occur currently by discounting using a suitable discount rate. Most commonly used discount rate in health economic evaluations is 3%, which is accordance with guidelines developed for USA (63). A discount rate of 3.5% for costs and health outcomes is recommended by UK's NICE guidelines for economic evaluations (64).

To date, no consensus has been reached on the choice of discount rate, and if the same discount rate should be used for both costs and outcomes. Controversies also remain on the choice of discounting models (e.g., constant vs hyperbolic discounting). Therefore, economic guidelines explicitly recommend sensitivity analyses with alternative discount rates to examine the robustness of economic evaluation findings to the choice of discount rates (65).

2.3.4.4 Sensitivity analysis

The results obtained from health economic evaluation models are often sensitive to the choice of inputs and modelling assumptions. Transferring the model findings from one setting to another may result in additional uncertainty owing to differences in economic and health care contexts (66). Other common sources of uncertainty to health economic model estimates include: uncertainty in the sample data, uncertainty relating to extrapolation and the generalisability of the results, and uncertainty relating to analytical methods (67).

Sensitivity analysis is an important component of economic evaluations used to assess uncertainty facing health economic model results, and to provide information to decision-makes about the robustness of their model-based decision. Key model inputs as well as assumptions concerning the model (individually or in combination with each other) are varied and the impact on the results recorded.

There are three types of sensitivity analysis to handle uncertainty facing economic evaluation results (66).

1. **Deterministic sensitivity analysis (DSA)**, which is conducted by varying the value of each parameter. In case of one-way DSA, only one parameter is varied at one time.

However, multi-way DSA assesses the impact of varying multiple parameters at one time.

2. **Extreme scenario analysis**, in which several important parameters are set under two extreme (the best and the worst) scenarios.
3. **Probabilistic sensitivity analysis (PSA**), which generates a distribution of outputs based on input parameters' distributions).

2.3.5 Key inputs into health economic models

The key inputs into health economic models include the costs of disease states, events and treatment costs, transition probabilities and health-related quality of life (HRQoL).

Costs may include the direct (healthcare/non-healthcare) and indirect (lost productivity) costs of a specific disease or disease state or event expressed in monetary terms. The direct costs include medical costs such as the costs of diagnostic tests, hospitalisations, nursing costs, and GPs/specialists consultation costs; and non-medical costs including the healthcare-related travel costs and informal care costs. The most frequently considered indirect costs are those due to productivity losses due to the impact of illness on patients and caregivers related to on-the-job productivity decreases (presenteeism), absenteeism, wage and leisure time losses. The categories of costs included in an evaluation depend on the study purpose, adopted perspective, study population etc (68).

Transition probabilities of disease progression are another important input to health economic evaluation models. Transition probabilities measure the likelihood of patients moving from one health state to another over one model cycle and range between 0: impossible to 1: certain (see also Section 2.2.3.2).

Health-related quality of life (HRQoL) of patients is a further key driver of economic evaluation models, and has been the focus of an ample volume of research during past decades (69).The concept of HRQOL is subjective and multidimensional. It has evolved since the 1980s to encompass physical, psychosocial, and other (e.g.: occupational and somatic sensational) aspects of life that are shown to affect health. HRQoL can be measured in many ways, and is most commonly reflected as health state utility values (HSUVs) in health economic evaluations (70). HRQoL can also be reflected as health dimension scores assessing physical, psychological and other health aspects of lives of patients (71, 72).

As the economic costs and transition probabilities are not the key focus of this research thesis, the remainder of this section will be focused on the key model input parameter of HRQoL only.

2.3.5.1 Health State Utility values (HSUVs)

As has been noted above, HRQoL is mostly reflected as HSUVs in health economic evaluation models. HSUVs measure the strength of a preference for a particular health state, and typically range between 0: equivalent to dead and 1: perfect health. Negative HSUVs are possible, indicating a health state worse than death. HSUV(s) are preference based measures that are typically used to calculate quality-adjusted life years (QALYs) - a preferred measure of clinical effectiveness in health economic evaluations (73).

HSUVs can be obtained through several methods (73) including direct methods such as the Standard Gamble (SG), Time Trade-off (TTO) and the rating scales (RS) and indirect methods that involve the use of preference-based multi-attribute utility instruments (MAUIs). MAUIs are formed of a generic HRQoL questionnaire and an accompanying formula or set of weights (or "tariffs") elicited from a sample of the general population for converting responses into HSUVs. The most used MAUIs include the EuroQoL-5-Dimension (EQ-5D [3 level and 5 level) instruments, Short-Form-6-Dimension (SF-6D), Health Utility Index (HUI) and Assessment of Quality of Life (AQoL) suite of instruments (Table 2.3). HSUVs measured by different MAUIs may differ due to the difference in their descriptive systems, scale effects and utility formulas (74). Finally, mapping techniques are used to transform non-preference based HRQoL measures into HSUVs.

Table 2.3. Comparisons of the dimensions and content of multi-attribute utility instruments.

	QWB	15D	EQ-5D		SF-6D	HUI 3	AQoL-8D
			EQ-5D-3L	EQ-5D-5L			
Year (preference weights available)	1976	1989	1995	2012	2002	2002	2009
Countries of origin	USA	Finland	Europe/UK		UK/USA	Canada	Australia
Dimensions	3 items plus 27 symptoms	15	5		6	8	8
Items		15	5		6	8	35 (25 items capture psychosocial dimensions of health)
Response levels	2-3	4-5	3	5	4-6	5-6	4
States defined	945	3.1×10^{10}	243	3,125	18,000	972,000	2.37×10^{23}
Total time taken	/	4 minutes	1 minutes		2.5 minutes	3 minutes	5.5 minutes
Dimension type[a]							
- Physical							
Mobility/activity	**	**	**		**	*	*
Self-care			*				
Dexterity						*	

Energy		*			*		*
Cognition						*	
Pain/discomfort		*		*	*	*	*
Senses (vision, hearing, speech)		***				***	*
- Psychosocial							
Social function	*	*			*		*
Mental function						*	
Mental health		**		*	*		**
Satisfaction							*
% difference between utilities of two instruments attributable to the difference in their descriptive system[b]							
15D	NR	/	NR	64.3%	43.0%	71.4%	27.4%
EQ-5D-5L	NR	/	NR	/	76.7%	85.5%	80.8%
SF-6D	NR	/	NR	/	/	50.9	58.8
HUI 3	NR	/	NR	/	/	/	101.6%
AQoL-8D	NR	/	NR	/	/	/	/
Number of studies used in present review	0	2	21	5	2	0	3

The number of asterisk [*] in the table represents the number of covered health dimensions in an instrument.
a. referred from "*Richardson (2011). Review and Critique of Health Related Multi Attribute Utility Instruments. Centre for Health Economics. Research paper 2011(64).*"
b. referred from "*Richardson (2015). Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. Qual Life Res. 2015 Aug;24(8):2045-53*."
Abbreviation: 15D, 15-dimensional questionnaire; AQoL, Assessment of Quality of Life; EQ-5D, EuroQoL five-dimensions scale; HUI, Health Utilities Index; QWB, Quality of Well-Being Scale; SF-6D, Short Form 6-dimension.
Source: Xia Q, Campbell JA, Ahmad H, Si L, de Graaff B, Otahal P, Palmer AJ. Health state utilities for economic evaluation of bariatric surgery: A comprehensive systematic review and meta-analysis. Obesity Reviews. 2020 Aug;21(8):e13028.

Whilst the International Society for Pharmacoeconomics and Outcomes research (ISPOR) recommends the use of locally derived model input parameter, analysts usually lack time and resources to estimate primary HSUVs for all health states of interest. Therefore, they tend to rely on systematic reviews and meta-analytic methods to pool HSUVs collected across several studies. Meta-analytical estimates are considered useful as they provide common source of values for economic evaluations across a clinical area or targeting a specific population group (75).

2.3.5.2 The AQoL-4D multi-attribute utility instrument (MAUI)

The Assessment of Quality of Life-4D (AQoL-4D) is a well validated (76), commonly used MAUIs. A unique feature of the AQoL-4D is that the utility weights have been derived from an Australian population sample (77) and norms of the Australian population are available (78), which makes it predominantly adopted in Australia-based research. The AQoL-4D MAUI can provide dimension scores and an overall index of the health state utility.

AQoL-4D questionnaire consists of 12 items covering four dimensions. Each dimension has three items with four response levels. The four dimensions and their corresponding items are: 1) independent living (self-care, activities of daily living, and mobility), 2) social relationships (social isolation, relationship and family role), 3) physical senses (sight, hearing and communication), and 4) psychological wellbeing (sleep, anxiety and pain) (Supplement 2.2).

Items scores are combined according to the AQoL-4D algorithm to calculate dimension scores and HSUVs (79) (Table 2.4).

Table 2.4 Items dis-utility values

Dimension	Item	Response level			
		1	2	3	4
Independent Living	1	0.000	0.154	0.403	1.000
	2	0.000	0.244	0.343	1.000
	3	0.000	0.326	0.415	1.000
Social Relationships	4	0.000	0.169	0.396	1.000
	5	0.000	0.095	0.191	1.000
	6	0.000	0.147	0.297	1.000
Physical Senses	7	0.000	0.145	0.288	1.000
	8	0.000	0.253	0.478	1.000
	9	0.000	0.219	0.343	1.000
Psychological Wellness	10	0.000	0.107	0.109	1.000
	11	0.000	0.141	0.199	1.000
	12	0.000	0.104	0.312	1.000

Each dimension disutility values (dvD) are estimated from the following equations:

$dvD_{\text{independent living}}$ = (1.0989*(1-(1-0.6097*dvQ1) *(1-0.4641*dvQ2) *(1-0.5733*dvQ3)))

$dvD_{\text{social relationships}}$ = (1.0395*(1-(1-0.7023*dvQ4) *(1-0.6253*dvQ5) *(1-0.6638*dvQ6)))

$dvD_{\text{physical senses}}$ = (1.6556*(1-(1-0.2476*dvQ7) *(1-0.2054*dvQ8) *(1-0.3382*dvQ9)))

$dvD_{\text{psychological wellness}}$ = (1.2920*(1-(1-0.1703*dvQ10) *(1-0.2554*dvQ11) *(1-0.6347*dvQ12)))

Each dimension scores are derived from the dvD and calculated using the equation:

Dimension score=1- dvD

The overall index of HSU is calculated using the equation:

HSU=(1.04*(1-(0.841*dvD1)) *(1-(0.855*dvD2)) *(1-(0.931*dvD3)) *(1-(0.997*dvD4))) - 0.04

2.3.6 Health economic evaluation reporting guideline

The transparent and consistent reporting of health economic evaluations is crucial to allow scrutiny of study findings and ultimately, better health decisions. With the increasing number of publications available, and high opportunity costs from decisions based on misleading study findings, transparency and clarity in reporting become even more important. The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) is the most widely used

reporting guideline for health economic evaluations, which includes recommendations for six main categories with 24 items (Supplement 2.1).

2.4 The Health Economics of OA

OA poses a substantial health economic burden to patients, their caregivers and society. The increasing prevalence of OA over time only exacerbates the situation. Most of the available OA treatments are symptom-specific and are largely ineffective in limiting OA's progression. Some non-pharmaceutical therapies are introduced to improve/maintain functional performance through lifestyle changes. Disease modifying therapies (DMTs) may also be introduced in the future. However, DMTs are likely to be costly and may differ in terms of their impacts. It is therefore imperative to examine the cost-effectiveness of various existing and upcoming OA treatments to identify the health interventions that are the best value for money.

2.4.1 The application of health economic evaluation models in the field of OA

The evaluation of policies for the treatment of chronic illnesses such as OA pose significant challenges, as health and economic consequences accrue over long periods, while evidence on the effectiveness of interventions is usually obtained from clinical trials or epidemiological studies of limited duration. Thus, the health economic evaluation models have been widely used to evaluate the OA interventions to capture all the possible outcomes. Health economic evaluations models incorporate clinical, health economic and epidemiological data and sophisticated statistical methodologies. Since the publication of the first model-based health economic evaluation of OA in 1994 (80), numerous modelling studies have been conducted, particularly in Western developed nations. In a systematic review of cost-effectiveness of surgical interventions of OA, 16 model-based studies were included compared with 6 trial-based (81). Another systematic review of cost-effectiveness analyses of OA oral therapies included 28 OA health economic evaluation models (82). In both reviews, they found the development of models that decision-tree models were generally replaced by Markov models or occasionally discrete event simulation in response to changing treatment landscapes, data availability as well as methodological advances in the application of decision science to health care. However, the scope of these systematic reviews of OA health-economic evidence was

limited to a specific treatment type (e.g.: surgical, oral therapies), and none of these reviews systematically reviewed all features of health economic evaluation models.

Building a model requires consideration of important elements including the complexity of the clinical area, the available evidence related to the disease, as well as other issues such as the scope or boundaries of the model, the appropriate time horizon, the perspective of the analysis, the availability of data, and a formal synthesis of evidence within the model (83). OA models can vary in their methodological framework, model structures, division of model events and input data sources with the development of all those elements as well as the methodological advances. Understanding the general characteristics of OA models, such as the target population, model structure, and type of questions answered, can provide future investigators with a systematic and broad view of the modelling landscape that has been adopted in the field of OA and lead to significant improvements in modelling practice (84). However, no previous systematic review has synthesised the evolution of health-economic evaluation models of all OA interventions, with an emphasis on their strengths and weaknesses and study gaps to inform the future health economic modelling of OA, this has been addressed in Chapter 3 of this thesis.

2.4.2 The studies on OA-related HSUVs

Numerous studies have been conducted to investigate the OA-related HSUVs, (85, 86).Systematic/literature review on OA-related HSUVs has been conducted to provide an overview of OA-related HSUVs (85, 86). The EQ-5D was found to be the most used MAUI to derive HSUVs; OA-related HSUVs spanned widely across studies and differed by measures. Moreover, observational studies were more commonly used to collect HSUVs than clinical trials. However, there are limitations facing previous reviews, the literature review on OA health utilities conducted by Ruchlin *et al*. only included randomized controlled trials or observational studies that published before 2006 and focused on the utility data for pain-related outcomes, it also did not provide information about joint-specific utilities (85). The systematic review on OA humanistic burden conducted by Xie *et al*. provided a brief OA utility ranges measured by different instruments, however, it only included studies published between 2006-2016 with sample size ≥1000 and did not included studies associated with OA treatments resulting in only five related studies were included (86). Furthermore, no meta-analysis has been conducted to synthesise the HSUVs for OA-related conditions to generate a HSUVs database for OA patients with different affected joint sites undergoing different treatments to

guide HSUVs choices in future health economic modelling of OA interventions, this has been addressed in Chapter 4 of this thesis.

2.4.3 The studies on OA-related HSUVs based on Australian population

HSUVs measure the population preference which is likely to vary from population to population. There were limited studies investigating the HSUVs of OA in Australia. From the 2007 Australian National Survey of Mental Health and Wellbeing, the population norms for arthritis were 0.81 (0.22) based on AQoL-4D, however, no OA-specific norms been investigated. The latest study that estimated the HSUVs of Australians with OA was published in 2015 using data from a multicentre, cross-sectional survey, focused on younger hip or knee OA based on AQoL-4D (87). However, only small number (n=21) of community-based participants were included in that study. As a chronic condition, OA is expected to associate with long-term impairment on HRQoL, however, no study has investigated longitudinal changes in HSUVs of people with OA compared to those without OA and investigated the impacted health dimensions in OA population that derived the HSUVs impairment, this has been addressed in Chapter 5 of this thesis. People with OA are associated with higher risk of comorbidities which also contribute to the impacts on HSUVs. However, there is lack of studies investigating the comorbidity profile of OA population to identify the most prevalent comorbidity patterns and their associations with HSUV impairments, this has been addressed in Chapter 6 of this thesis.

2.5 Reference

1. Rousseau J-C, Delmas PD. Biological markers in osteoarthritis. Nature Clinical Practice Rheumatology. 2007;3(6):346-56.

2. Kean W, Kean R, Buchanan W. Osteoarthritis: symptoms, signs and source of pain. Inflammopharmacology. 2004;12(1):3-31.

3. Pereira D, Peleteiro B, Araujo J, Branco J, Santos R, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. Osteoarthritis and cartilage. 2011;19(11):1270-85.

4. Gupta KB, Duryea J, Weissman BN. Radiographic evaluation of osteoarthritis. Radiologic Clinics. 2004;42(1):11-41.

5. Culvenor AG, Engen CN, Øiestad BE, Engebretsen L, Risberg MA. Defining the presence of radiographic knee osteoarthritis: a comparison between the Kellgren and Lawrence system and OARSI atlas criteria. Knee Surgery, Sports Traumatology, Arthroscopy. 2015;23(12):3532-9.

6. Felson DT, Nevitt MC, Yang M, Clancy M, Niu J, Torner JC, et al. A new approach yields high rates of radiographic progression in knee osteoarthritis. The Journal of rheumatology. 2008;35(10):2047-54.

7. Salaffi F, Carotti M, Stancati A, Grassi W. Radiographic assessment of osteoarthritis: analysis of disease progression. Aging clinical and experimental research. 2003;15(5):391-404.

8. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology. 1991;34(5):505-14.

9. Nelson A, Jordan J. Defining osteoarthritis: a moving target. Osteoarthritis and cartilage. 2012;20(1):1-3.

10. Altman R. Criteria for classification of clinical osteoarthritis. The Journal of rheumatology Supplement. 1991;27:10-2.

11. Singh G, Miller JD, Lee FH, Pettitt D, Russell MW. Prevalence of cardiovascular disease risk factors among US adults with self-reported osteoarthritis: data from the Third National Health and Nutrition Examination Survey. population. 2002;7:17.

12. Peeters GG, Alshurafa M, Schaap L, de Vet HC. Diagnostic accuracy of self-reported arthritis in the general adult population is acceptable. Journal of clinical epidemiology. 2015;68(4):452-9.

13. Parsons C, Clynes M, Syddall H, Jagannath D, Litwic A, van der Pas S, et al. How well do radiographic, clinical and self-reported diagnoses of knee osteoarthritis agree? Findings from the Hertfordshire cohort study. SpringerPlus. 2015;4(1):1-5.

14. Blagojevic M, Jinks C, Jeffery A, Jordan K. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthritis and cartilage. 2010;18(1):24-33.

15. Aigner T, Rose J, Martin J, Buckwalter J. Aging theories of primary osteoarthritis: from epidemiology to molecular biology. Rejuvenation research. 2004;7(2):134-45.

16. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clinics in geriatric medicine. 2010;26(3):355-69.

17. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan J, Protheroe J, Jordan K. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. Osteoarthritis and cartilage. 2015;23(4):507-15.

18. Hosnijeh FS, Kavousi M, Boer CG, Uitterlinden AG, Hofman A, Reijman M, et al. Development of a prediction model for future risk of radiographic hip osteoarthritis. Osteoarthritis and cartilage. 2018;26(4):540-6.

19. Tanamas S, Hanna FS, Cicuttini FM, Wluka AE, Berry P, Urquhart DM. Does knee malalignment increase the risk of development and progression of knee osteoarthritis? A systematic review. Arthritis Care & Research: Official Journal of the American College of Rheumatology. 2009;61(4):459-67.

20. Bierma-Zeinstra SM, Koes BW. Risk factors and prognostic factors of hip and knee osteoarthritis. Nature clinical practice Rheumatology. 2007;3(2):78-85.

21. Schouten JS, de Bie RA, Swaen G. An update on the relationship between occupational factors and osteoarthritis of the hip and knee. Current opinion in rheumatology. 2002;14(2):89-92.

22. Maetzel A, Mäkelä M, Hawker G, Bombardier C. Osteoarthritis of the hip and knee and mechanical occupational exposure--a systematic overview of the evidence. The Journal of rheumatology. 1997;24(8):1599-607.

23. Ezzat AM, Li LC. Occupational physical loading tasks and knee osteoarthritis: a review of the evidence. Physiotherapy Canada. 2014;66(1):91-107.

24. Harris EC, Coggon D. HIP osteoarthritis and work. Best practice & research Clinical rheumatology. 2015;29(3):462-82.

25. Lievense A, Bierma-Zeinstra S, Verhagen A, Bernsen R, Verhaar J, Koes B. Influence of sporting activities on the development of osteoarthritis of the hip: a systematic review. Arthritis Care & Research: Official Journal of the American College of Rheumatology. 2003;49(2):228-36.

26. Van Meurs J. Osteoarthritis year in review 2016: genetics, genomics and epigenetics. Osteoarthritis and cartilage. 2017;25(2):181-9.

27. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. British medical bulletin. 2013;105(1):185-99.

28. Hosseini K, Gaujoux-Viala C, Coste J, Pouchot J, Fautrel B, Rat A-C, et al. Impact of co-morbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis. Best Practice & Research Clinical Rheumatology. 2012;26(5):627-35.

29. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. The Annals of Family Medicine. 2009;7(4):357-63.

30. Seed SM, Dunican KC, Lynch AM. Osteoarthritis: A review of treatment options. Geriatrics. 2009;64(10).

31. Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in Osteoarthritis: A systematic review and meta-analysis of observational studies. Arthritis care & research. 2020;72(7):991-1000.

32. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. The Lancet. 2019;393(10182):1745-59.

33. Health AIo, Welfare. Osteoarthritis. Canberra: AIHW; 2020.

34. Ackerman IN, Pratt C, Gorelik A, Liew D. Projected burden of osteoarthritis and rheumatoid arthritis in Australia: a population-level analysis. Arthritis care & research. 2018;70(6):877-83.

35. McAlindon TE, Bannuru RR, Sullivan M, Arden N, Berenbaum F, Bierma-Zeinstra S, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis and cartilage. 2014;22(3):363-88.

36. Brosseau L, Taki J, Desjardins B, Thevenot O, Fransen M, Wells GA, et al. The Ottawa panel clinical practice guidelines for the management of knee osteoarthritis. Part two: strengthening exercise programs. Clinical rehabilitation. 2017;31(5):596-611.

37. Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, et al. Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. Arthritis and rheumatism. 1995;38(11):1541-6.

38. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM, editors. A systematic review of recommendations and guidelines for the management of osteoarthritis: the chronic osteoarthritis management initiative of the US bone and joint initiative. Seminars in arthritis and rheumatism; 2014: Elsevier.

39. Zhang W, Moskowitz R, Nuki G, Abramson S, Altman R, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis and cartilage. 2007;15(9):981-1000.

40. Brand C, Buchbinder R, Wluka A, Ruth D, McKenzie S, Jones K, et al. Guideline for the non-surgical management of hip and knee osteoarthritis: The Royal Australian College of General Practitioners; 2009.

41. Basedow M, Williams H, Shanahan EM, Runciman WB, Esterman A. Australian GP management of osteoarthritis following the release of the RACGP guideline for the non-surgical management of hip and knee osteoarthritis. BMC research notes. 2015;8(1):1-8.

42. Ofman JJ, MacLean CH, Straus WL, Morton SC, Berger ML, Roth EA, et al. A metaanalysis of severe upper gastrointestinal complications of nonsteroidal antiinflammatory drugs. The Journal of rheumatology. 2002;29(4):804-12.

43. Lin J, Zhang W, Jones A, Doherty M. Efficacy of topical non-steroidal anti-inflammatory drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials. Bmj. 2004;329(7461):324.

44. Evans J, McMahon AD, McGilchrist MM, White G, Murray FE, McDevitt DG, et al. Topical non-steroidal anti-inflammatory drugs and admission to hospital for upper gastrointestinal bleeding and perforation: a record linkage case-control study. Bmj. 1995;311(6996):22-6.

45. Curtis E, Fuggle N, Shaw S, Spooner L, Ntani G, Parsons C, et al. Safety of cyclooxygenase-2 inhibitors in osteoarthritis: outcomes of a systematic review and meta-analysis. Drugs & aging. 2019;36(1):25-44.

46. Capurso L, Koch M. Prevention of NSAID-induced gastric lesions: H2 antagonists or misoprostol? A meta-analysis of controlled clinical studies. 1991.

47. Morris S, Devlin N, Parkin D. Economic analysis in health care: John Wiley & Sons; 2007.

48. Arrow KJ. Uncertainty and the welfare economics of medical care. Uncertainty in economics. 1978:345-75.

49. Kernick DP. Introduction to health economics for the medical practitioner. Postgraduate medical journal. 2003;79(929):147-50.

50. Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. Jama. 2009;302(13):1437-43.

51. Drummond M, Jönsson B, Rutten F. The role of economic evaluation in the pricing and reimbursement of medicines. Health policy. 1997;40(3):199-215.

52. Duthie T, Trueman P, Chancellor J, Diez L. Research into the use of health economics in decision making in the United Kingdom—Phase II: Is health economics 'for good or evil'? Health Policy. 1999;46(2):143-57.

53. Udeh BL. Economic Evaluation Studies. Chest. 2020;158(1):S88-S96.

54. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial‐based economic evaluation for health care decision making? Health economics. 2006;15(7):677-87.

55. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation: Oup Oxford; 2006.

56. Brennan A, Akehurst R. Modelling in health economic evaluation. Pharmacoeconomics. 2000;17(5):445-59.

57. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. Journal of health services research & policy. 2004;9(2):110-8.

58. Higashi H, Barendregt JJ. Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis: discrete-event simulation model. PloS one. 2011;6(9):e25403.

59. Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? Value in health. 2010;13(8):1056-60.

60. Chalkidou K, Culyer AJ, Nemzoff C. Perspective in economic evaluations of health care interventions in low-and middle-income countries—one size does not fit all. CGD Policy Paper. 2018.

61. Polimeni JM, Vichansavakul K, Iorgulescu RI, Chandrasekara R. Why perspective matters in health outcomes research analyses. International Business & Economics Research Journal (IBER). 2013;12(11):1503-12.

62. Basu A, Maciejewski ML. Choosing a time horizon in cost and cost-effectiveness analyses. Jama. 2019;321(11):1096-7.

63. Haacker M, Hallett TB, Atun R. On discount rates for economic evaluations in global health. Health Policy and Planning. 2019;35(1):107-14.

64. Massetti M, Aballéa S, Videau Y, Rémuzat C, Roïz J, Toumi M. A comparison of HAS & NICE guidelines for the economic evaluation of health technologies in the context of their respective national health care systems and cultural environments. Journal of market access & health policy. 2015;3(1):24966.

65. Attema AE, Brouwer WB, Claxton K. Discounting in economic evaluations. Pharmacoeconomics. 2018;36(7):745-58.

66. Limwattananon S. Handling uncertainty of the economic evaluation result: sensitivity analysis. Journal of the Medical Association of Thailand. 2011;91(6):59.

67. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. Health economics. 1994;3(2):95-104.

68. Garrison Jr LP, Pauly MV, Willke RJ, Neumann PJ. An overview of value, perspective, and decision context—a health economics approach: an ISPOR Special Task Force report [2]. Value in Health. 2018;21(2):124-30.

69. Haraldstad K, Wahl A, Andenæs R, Andersen JR, Andersen MH, Beisland E, et al. A systematic review of quality of life research in medicine and health sciences. Quality of Life Research. 2019;28(10):2641-50.

70. Torrance GW. Utility approach to measuring health-related quality of life. Journal of chronic diseases. 1987;40(6):593-600.

71. Ahmad H, van der Mei I, Taylor BV, Campbell JA, Palmer AJ. Measuring the health-related quality of life in Australians with multiple sclerosis using the assessment of quality of life-8-dimension (AQoL-8D) multi-attribute utility instrument. Multiple Sclerosis and Related Disorders. 2020;44:102358.

72. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clinical and Experimental Research. 2005;17(4):255-63.

73. Meregaglia M, Cairns J. A systematic literature review of health state utility values in head and neck cancer. Health and Quality of Life Outcomes. 2017;15(1):174.

74. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. Quality of Life Research. 2015;24(8):2045-53.

75. Petrou S, Kwon J, Madan J. A practical guide to conducting a systematic review and meta-analysis of health state utility values. Pharmacoeconomics. 2018;36(9):1043-61.

76. Hawthorne G, Richardson J, Osborne R, McNeil H. The Australian quality of life (AQoL) instrument: initial validation: Centre for Health Program Evaluation Melbourne, Vic.; 1997.

77. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Annals of medicine. 2001;33(5):358-70.

78. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Australian and New Zealand journal of public health. 2005;29(2):136-42.

79. Hawthorne G, Richardson J, Day N. Using the Assessment of Quality of life (AQoL) version 1. Melbourne: Centre for Health Program Evaluation. 2001.

80. McKell C, Stewart A. Cost-minimisation analysis comparing topical versus systematic NSAIDS in the treatment of mild osteoarthritis of the superficial joints. British Journal of Medical Economics. 1994;7:137-46.

81. Kamaruzaman H, Kinghorn P, Oppong R. Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature. BMC musculoskeletal disorders. 2017;18(1):1-17.

82. Wielage RC, Myers JA, Klein RW, Happich M. Cost-effectiveness analyses of osteoarthritis oral therapies: a systematic review. Applied health economics and health policy. 2013;11(6):593-618.

83. Ramos MCP, Barton P, Jowett S, Sutton AJ. A systematic review of research guidelines in decision-analytic modeling. Value in Health. 2015;18(4):512-29.

84. Zafari Z, Bryan S, Sin DD, Conte T, Khakban R, Sadatsafavi M. A systematic review of health economics simulation models of chronic obstructive pulmonary disease. Value in Health. 2017;20(1):152-62.

85. Ruchlin HS, Insinga RP. A Review of Health-Utility Data for Osteoarthritis. Pharmacoeconomics. 2008;26(11):925-35.

86. Xie F, Kovic B, Jin X, He X, Wang M, Silvestre C. Economic and humanistic burden of osteoarthritis: a systematic review of large sample studies. Pharmacoeconomics. 2016;34(11):1087-100.

87. Ackerman IN, Bucknill A, Page RS, Broughton NS, Roberts C, Cavka B, et al. The substantial personal burden experienced by younger people with hip or knee osteoarthritis. Osteoarthritis and cartilage. 2015;23(8):1276-84.

2.6 Supplements

Supplement 2.1. CHEERS checklist-Items to include when reporting economics evaluations of health interventions.

Section/item	Item No	Recommendation	Reported on page No/line No
Title and abstract			
Title	1	Identify the study as an economic evaluation or use more specific terms such as "cost-effectiveness analysis", and describe the interventions compared.	_____
Abstract	2	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base-case and uncertainty analyses), and conclusions.	_____

Introduction			
Background and objectives	3	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for health policy or practice decisions.	_____ _____
Methods			
Target population and subgroups	4	Describe characteristics of the base-case population and subgroups analyzed including why they were chosen.	_____
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	_____
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	_____
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	_____
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	_____
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	_____
Choice of health outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	_____
Measurement of effectiveness	11a	Single study–based estimates: Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.	_____
	11b	Synthesis-based estimates: Describe fully the methods used for the identification of included studies and synthesis of clinical effectiveness data.	_____
Measurement and valuation of preference-based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	_____
Estimating resources and costs	13a	Single study–based economic evaluation: Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	_____
	13b	Model-based economic evaluation: Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	_____
Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.	_____
Choice of model	15	Describe and give reasons for the specific type of decision-analytic model used. Providing a figure to show model structure is strongly recommended.	_____

Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytic model.	_____
Analytic methods	17	Describe all analytic methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (e.g., half-cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.	_____
Results			
Study parameters	18	Report the values, ranges, references, and if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.	_____
Incremental costs and outcomes	19	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.	_____
Characterizing uncertainty	20a	Single study–based economic evaluation: Describe the effects of sampling uncertainty for estimated incremental cost, incremental effectiveness, and incremental cost-effectiveness, together with the impact of methodological assumptions (such as discount rate, study perspective).	_____
	20b	Model-based economic evaluation: Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.	_____
Characterizing heterogeneity	21	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.	_____
Discussion			
Study findings, limitations, generalizability, and current knowledge	22	Summarize key study findings and describe how they support the conclusions reached. Discuss limitations and the generalizability of the findings and how the findings fit with current knowledge.	_____
Other			
Source of funding	23	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other nonmonetary sources of support.	_____
Conflicts of interest	24	Describe any potential for conflict of interest among study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors' recommendations.	_____

Supplement 2.2 The AQoL-4D questionnaire

INDEPENDENT LIVING

1. Do I need any help looking after myself?

I need no help at all	1
Occasionally I need some help with personal care tasks	2
I need help with the more difficult personal care tasks	3
I need daily help with most or all personal care tasks	4

2. When doing household tasks: (For example, preparing food, gardening, using the video recorder, radio, telephone or washing the car)

I need no help at all	1
Occasionally I need some help with household tasks	2
I need help with the more difficult household tasks	3
I need daily help with most or all household tasks	4

3. Thinking about how easily I can get around my home and community:

I get around my home and community by myself without any difficulty	1
I find it difficult to get around my home and community by myself	2
I cannot get around the community by myself, but I can get around my home with some difficulty	3
I cannot get around either the community or my home by myself	4

SOCIAL RELATIONSHIPS

4. Because of my health, my relationships (for example: with my friends, partner or parents) generally:

Are very close and warm	1
Are sometimes close and warm	2
Are seldom close and warm	3
I have no close and warm relationships	4

5. Thinking about my relationship with other people:

I have plenty of friends, and am never lonely	1
Although I have friends, I am occasionally lonely	2
I have some friends but am often lonely for company	3
I am socially isolated and feel lonely	4

6. Thinking about my health and my relationship with my family:

My role in the family is unaffected by my health	1
There are some parts of my family role I cannot carry out	2
There are many parts of my family role I cannot carry out	3
I cannot carry out any part of my family role	4

PHYSICAL SENSES

7. Thinking about my vision, including when using my glasses or contact lenses if needed:

	I see normally	1
	I have some difficulty focusing on things, or I do not see them sharply. For example: small print, a newspaper, or seeing objects in the distance	2
	I have a lot of difficulty seeing things. My vision is blurred. For example: I can see just enough to get by with	3
	I only see general shapes or am blind. For example: I need a guide to move around	4
8. Thinking about my hearing, including using my hearing aid if needed:		
	I hear normally	1
	I have some difficulty hearing or I do not hear clearly. For example: I ask people to speak up, or turn up the TV or radio volume	2
	I have difficulty hearing things clearly. For example: Often I do not understand what is said. I usually do not take part in conversations because I cannot hear what is said	3
	I hear very little indeed. For example: I cannot fully understand loud voices speaking directly to me	4
9. When I communicate with others: (For example: by talking, listening, writing or signing)		
	I have no trouble speaking to them or understanding what they are saying	1
	I have some difficulty being understood by people who do not know me. I have no trouble understanding what others are saying to me	2
	I am only understood by people who know me well. I have great trouble understanding what others are saying to me	3
	I cannot adequately communicate with others	4
PSYCHOLOGICALWELLBEING		
10. If I think about how I sleep:		
	I am able to sleep without difficulty most of the time	1
	My sleep is interrupted some of the time, but I am usually able to go back to sleep without difficulty	2
	My sleep is interrupted most nights, but I am usually able to go back to sleep without difficulty	3
	I sleep in short bursts only. I am awake most of the night	4
11. Thinking about how I generally feel:		
	I do not feel anxious, worried or depressed	1
	I am slightly anxious, worried or depressed	2
	I feel moderately anxious, worried or depressed	3
	I am extremely anxious, worried or depressed	4

12. How much pain or discomfort do I experience?

None at all	1
I have moderate pain	2
I suffer from severe pain	3
I suffer unbearable pain	4

Chapter 3: A Systematic Review of the Evolution of Health-Economic Evaluation Models of Osteoarthritis

3.1 Preface

Chapter 3 presents Study 1 of this thesis which comprehensively investigates the evolution of health-economic evaluation models used for all forms of OA interventions, including preventions, core treatments, adjunct non-pharmacological interventions, pharmacological and surgical treatments; with an emphasis on their strengths and weaknesses and study gaps to inform the future development of robust OA health economic models of various forms and treatments of OA, internationally and in Australia.

The text in this chapter has been published in *Arthritis Care & Research* (Zhao, Ting, Hasnat Ahmad, Barbara de Graaff, Qing Xia, Tania Winzenberg, Dawn Aitken, and Andrew J. Palmer "A Systematic Review of the Evolution of Health-Economic Evaluation Models of Osteoarthritis." Arthritis Care & Research (2020) (Supplement 3A).

3.2 Abstract

Objective

To comprehensively synthesise the evolution of health-economic evaluation models (HEEMs) of all OA interventions including preventions, core treatments, adjunct non-pharmacological interventions, pharmacological and surgical treatments.

Methods

The literature was searched within health-economic/biomedical databases. Data extracted included: OA type, population characteristics, model setting/type/events, study perspective, comparators; and the reporting quality of the studies was assessed. The review protocol was registered at the International Prospective Register of Systematic Reviews (registration: CRD42018092937).

Results

Eighty-eight studies were included. Pharmacological and surgical interventions were the focus in 51% and 44% studies, respectively. Twenty-four studies adopted a societal perspective

(with increasing popularity after 2013), however most (63%) did not include indirect costs. Quality-adjusted life years (QALYs) was the most popular outcome measure since 2008. Markov models were used by 62% of studies, with increasing popularity since 2008. Until 2010, most studies used short-to-medium time horizons; subsequently a lifetime horizon became popular. Eighty-six percent of studies reported discount rate(s) (predominantly between 3% and 5%). Studies published after 2002 had a better coverage of OA-related adverse events (AEs). Reporting quality significantly improved after 2001.

Conclusions

OA HEEMs have evolved and improved substantially over time, with focus shifting from short-to-medium-term pharmacological decision-tree models to surgical-focused lifetime Markov models. Indirect costs of OA are frequently not considered, despite using a societal perspective. There was lack of reporting sensitivity of model outcome to input parameters including discount rate, OA definition, and population parameters. Whilst the coverage of OA-related AEs has improved over time, it is still not comprehensive.

3.3 Introduction

Approximately 240 million people globally were affected by osteoarthritis (OA) in 2016 (1) and its prevalance is projected to rise steadily (1-3). OA is characterised by joint pain, stiffness, swelling, loss of function and disability; which in turn, negatively impacts individuals' health-related quality of life (HRQoL) (4) and poses a significant economic burden to patients and society in terms of both direct (healthcare) and indirect costs (from lost productivity [early retirement/absenteeism/presenteeism]) (5-10).

Whilst there is no cure for OA, there are treatments available to ease OA symptoms and postpone disability progression. According to recent OA management guidelines (11), treatments include lifestyle (e.g.: exercise, weight management), non-medical (e.g.: heat packs, manual therapy), medical, and surgical interventions. In chronic diseases like OA, healthcare policy decisions should be informed by evidence of the long-term health and economic impacts. Given the scarcity of healthcare resources, it is critically important that the most cost-effective interventions are chosen.

Health-economic evaluations (HEEs) compare alternative treatment options in terms of both economic costs and clinical effectiveness to identify the interventions that are best value for money. The models used for these incorporate clinical, health-economic and epidemiological

data. HEEs include full (e.g.: cost-benefit, cost-effectiveness and cost-utility analysis) and partial (e.g.: cost of illness analysis) evaluations (Supplement 3.1) (12). Since the first model-based HEE of OA treatments was performed in 1994, numerous modelling studies have been conducted, particularly in Western developed nations (13-15). OA models vary in their methodological framework, model structures, division of model events, and input data sources (16-19). The availability of better quality clinical and epidemiological data, and methodological advances have contributed to the evolution of modelled HEEs over time (16, 17, 20).

The scope of previous systematic reviews of OA health-economic evidence was often limited to a specific treatment type (e.g.: surgical, pharmacological, physical) (21-23). None have synthesised the evolution of health-economic evaluation models (HEEMs) of all OA interventions. As the comprehensive investigation of the diverse HEEMs could lead to significant improvements in modelling practice (24, 25), we aimed to explore the evolution of HEEMs used for all forms of OA interventions, with an emphasis on their strengths and weaknesses and study gaps to inform the future development of an improved and overarching HEEM of OA.

3.4 Methods

This study is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (26). The protocol was registered at the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42018092937).

3.4.1 Literature search

Three biomedical (Medline via OvidSP, Embase via Ovid, and China National Knowledge Infrastructure) and three health-economic/economic databases (American Economic Association, the Centre for Reviews and Dissemination and the Cost-effectiveness Analysis Registry) were searched according to a search strategy that was defined in consultation with co-authors and a research librarian (Supplement 3.2). The literature was searched from each database's inception to July 2018. Reference lists of included studies and relevant reviews were hand searched.

3.4.2 Screening criteria

Title/abstract screening and full-text screening were performed in Covidence (27) by two reviewers (TZ and QX) independently based on predefined inclusion and exclusion criteria. Screening conflicts between the two reviewers were resolved through discussion with senior researchers (AP, HA and BdeG).

The inclusion criteria were studies: 1) in humans; 2) that reported the construction/application/validation of partial or full OA HEEM; 3) that were available as full-text; and 4) that were published in English/Chinese/German. Studies were included if they focused on arthritis populations including OA if they did not report the proportion of OA participants or if they reported on a cohort with ≥90% OA. Review articles, conference abstracts, comments and books were excluded. (Supplement 3.3)

Data extraction

A Microsoft Excel spreadsheet designed by co-authors was used to extract data by two reviewers (TZ and HA) independently. Discrepancies were resolved by consensus, and an additional reviewer (AP) was consulted in cases of no consensus. Data extracted included authors, publication year, study setting, OA type, targeted interventions and comparators, and information related to HEEMs (simulated population characteristics, study perspective, time horizon, discount rate, clinical effectiveness measures, cost inputs, model type, modelling software, health states, health events, and sensitivity analysis type). (Supplement 3.4)

3.4.3 Assessment of reporting quality

We assessed each study's reporting quality using the 24-item Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist (28). We adapted the evaluation methods published by de Graaff et al. (29). Twenty-four items were equally weighted with '1' referring to the item being well performed, and '0' otherwise. As not all items were applicable to all studies (e.g.: for a cost of illness analysis, "effectiveness" was not applicable), the quality scores were converted to percentages, adjusting the denominator to reflect the different number of applicable items. Studies were categorized into low (≤50%), moderate (50%-75%) and high reporting quality (>75%) groups (29).

3.4.4 Strategy for data synthesis

We adopted a narrative, descriptive synthesis approach (30) to assess and outline the evolution of HEEMs of OA interventions over time.

3.5 Results

3.5.1 Screening results

As shown in Figure 1, our search identified 1,683 potential references (1,498 from biomedical and 185 from economic databases). After removal of duplicates (n=424), 1,259 were left for title and abstract screening, which excluded 1,080. Of the 179 left, 96 were excluded during full-text screening (Figure 3.1). We identified an additional 5 studies through hand-searching, resulting in a total of 88 included studies (Supplement 3.5).

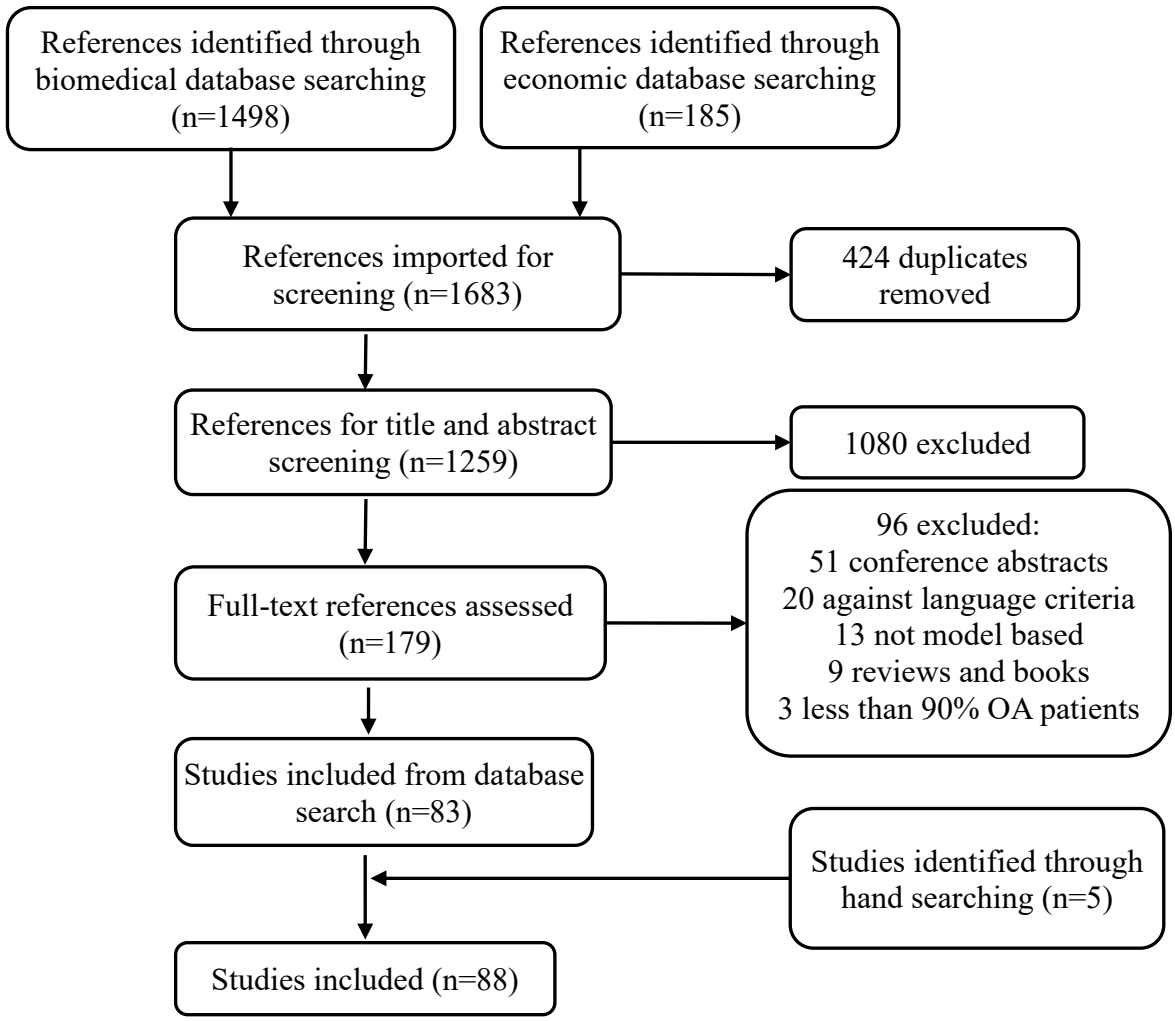


Figure 3.1. Flow chart results of study search based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses methodology

3.5.2 Year of publication

The first model-based HEE of OA treatment was published in 1994 (31) and more than half (51%, n=45) were published between 2013 and 2018 (Figure 3.2a).

3.5.3 Study settings

Almost half of the studies (49%, n=43) were conducted in the Americas, followed by Europe (36%, n=32) (Figure 3.2b). Most studies were performed in the United States (US) (40%, n=35), followed by the United Kingdom (UK) (18%, n=16) and Canada (7%, n=6). Only two studies were conducted in Australia and/or New Zealand. Seven did not report the study location. Studies from Australia, New Zealand, Saudi Arabia and China were published only after 2011.

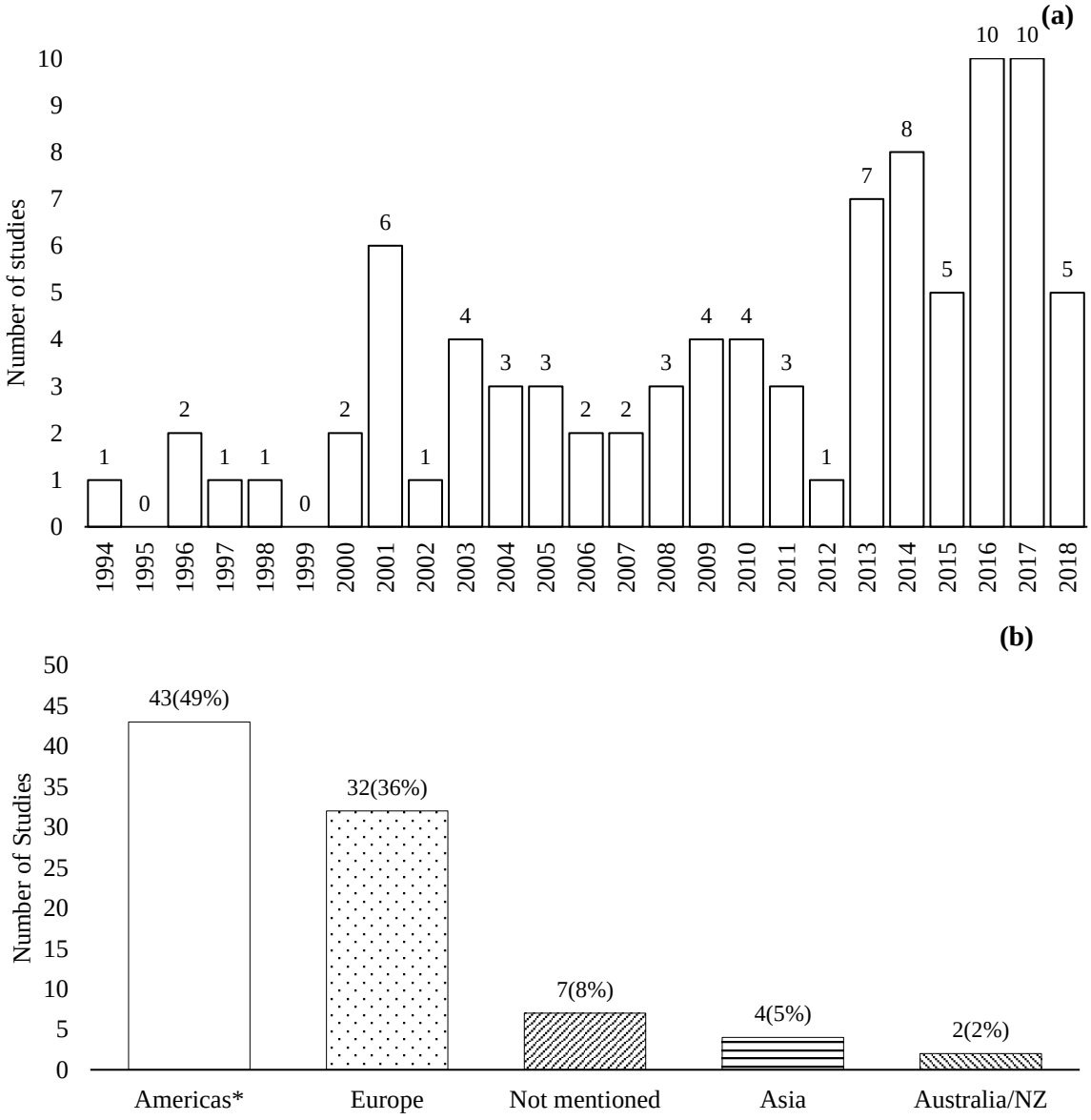


Figure 3.2. The distribution of published model-based osteoarthritis (OA) health economic evaluations by years of publication (a) and study settings (b). *Americas included both North and South America. NZ=New Zealand.

3.5.4 OA types and intervention options

Thirty-seven studies (42%) focused on knee OA and 13 (15%) on hip OA. One study each focused on ankle arthritis and glenohumeral OA. Thirty-six (41%) studies did not specify OA type, most of which targeted non-surgical treatments. Most studies focusing on a specific type of OA evaluated surgical treatments. Seventy-five percent (48/64) of post-2005 publications focused on a specific type of OA, compared to only 17% (4/24) of such studies pre-2005. (Figure 3.3a)

Of the 88 studies, 45 (51%) focused on pharmacological interventions and 39 (44%) focused on surgical interventions (Figure 3.3a, Supplement 3.6 [for the full list of interventions by OA type, intervention categorization, frequency, and references]). Four of 88 included studies assessed OA preventions (32-35), all of which were performed after 2014. Before 2005, all but two (16, 36) compared alternative pharmacological treatments, particularly cyclo-oxygenase 2 (COX-2) inhibitors selective nonsteroidal anti-inflammatory drugs (NSAIDs) with non-selective NSAIDs. Studies focusing on OA surgical treatments became popular thereafter.

3.5.5 Targeted populations

The target population was generally patients aged $\geq$40 years. Three studies specifically focused on older patients with mean ages of 74, 78 and $\geq$80 years (37-39). Of 45 pharmacological-focused studies, 17 analysed two patient subgroups separately: 1) low risk patients (younger and without prior history of upper gastrointestinal [GI]) events); and 2) high risk patients (older and with upper GI history).

3.5.6 Study perspectives and reporting of costs

Study perspectives were reported in 79 (90%) studies. Of these, 38 (48%) adopted a national healthcare perspective and 24 (30%) a societal perspective, however, 15 (63%) of these 24 studies did not consider indirect costs. Eleven studies adopted a third-party payer perspective, and single studies used patient and veteran health administration perspectives. In addition, four studies considered more than one perspective, all of which were published after 2013 (Figure 3.3b).

For pharmacological-focused studies, direct medical costs generally comprised of drug acquisition and adverse events (AEs) treatment costs, while for surgical-focused studies, direct costs were from primary/revision surgery. Indirect costs included productivity losses (n=6), time lost from work (n=2), lost wages (n=3), and caregivers' expenses (n=1).

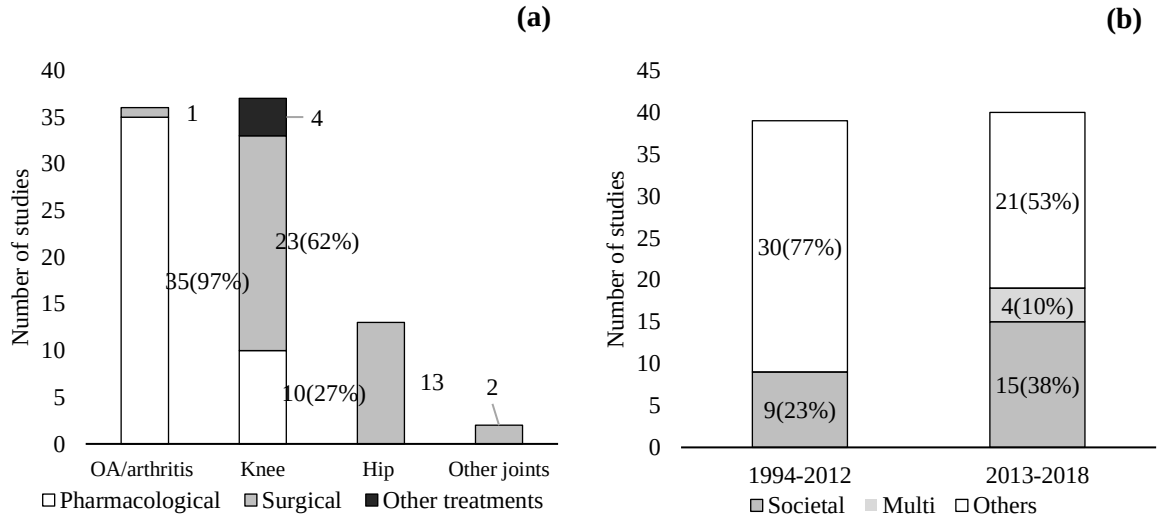


Figure 3.3. The distribution of included studies by focused OA types and intervention options (a) and the distribution of adopted perspectives by published time (n=79) (b). OA/arthritis includes OA without a specified site and arthritis comprising OA and rheumatoid arthritis. OA=osteoarthritis.

3.5.7 Effectiveness measurements

Eighty-nine percent (78/88) of studies reported the measurement of effectiveness, most commonly quality-adjusted life years (QALYs) (n=62), followed by multiple measures (n=7), and disease-specific effectiveness measures (n=6) (e.g.: AEs/complications averted, revision-free life years) (Figure 3.4a). The use of QALYs almost doubled, from 48% of studies published pre-2008 to 93% for those published after 2008 (Figure 3.4b).

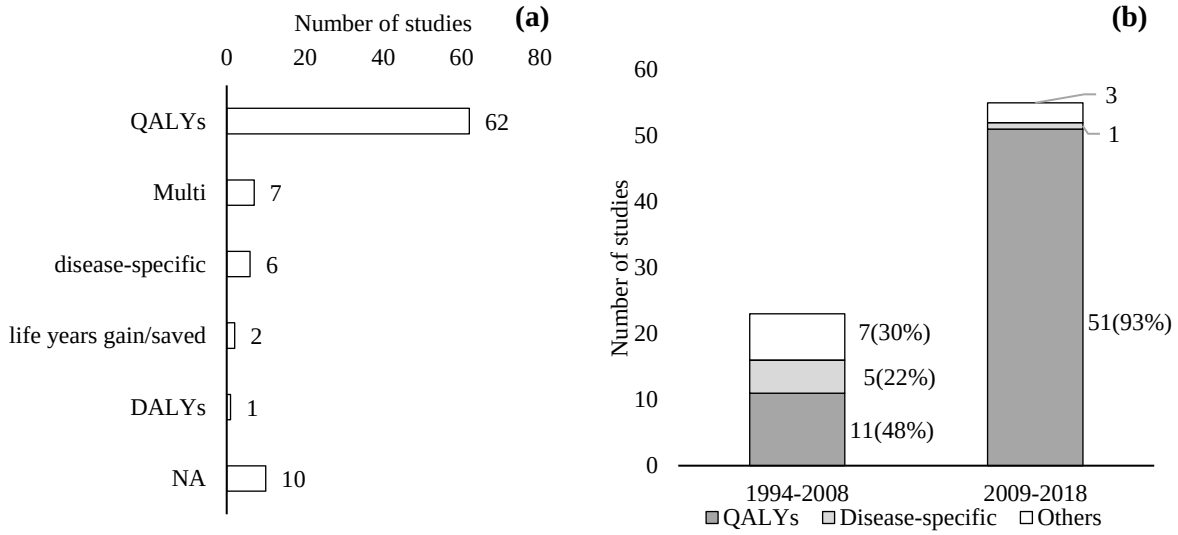


Figure 3.4. The effectiveness measures adopted in the included studies (a) and distribution by published time (n=78) (b). QALYs=quality-adjusted life years, DALYs=disability-adjusted life years, NA=not applicable.

3.5.8 Model types and computational software

Model type was reported in 87 (99%) studies. There were three key types: 1) Markov models (n=54, 62%), 2) decision-tree models (n=30, 34%), and 3) discrete-event simulation models (n=3, 3%) (Figure 3.5a). Of the 54 Markov model studies, most (n=44) were reported as Markov models, 8 reported the use of individua-level microsimulation models, and 2 studies reported as Markov cohort models. Markov models predominantly focused on surgical treatments, while pharmacological treatments were evaluated in 87% of decision-tree models (Figure 3.5a). The popularity of Markov models increased over time from 7% (pre-2008 period) to 88% (after 2008) (Figure 3.5b). There were four commonly used OA model structures : 1) OA policy (OAPoL) model (a Markov model to simulate the natural history of knee OA and predominantly used in the US) (19, 35, 38, 40-42); 2) the National Institute for Health and Care Excellence (NICE) model (a Markov model originally developed to compare NSAID/COX-2 inhibitor oral analgesics and subsequently extended to incorporate dose titration, discontinuation and AEs in addition to GI and cardiovascular (CV) AEs) (14, 18, 20, 43-46); 3) a model developed by Fitzpatrick (a Markov model aimed at evaluating hip OA surgical treatments) (47-51); and 4) a decision-tree model developed by Burke to compare NSAID/COX-2 inhibitor oral analgesics (52-55) (Supplement 3.7).

The adopted modelling software was mentioned in 57% (n=49) studies, with TreeAge being the most common (n=38), followed by Microsoft Excel (n=8), and Arena (n=1). Two studies

used more than one software, and no clear time trend was observed in the choice of software (Supplement 3.4).

3.5.9 Time horizon

Forty-eight (55%) studies ran the model over a lifetime horizon while 39 (45%) used short-to-medium-term horizons (from 2 weeks to 30 years). Markov model-based evaluations mostly adopted lifetime horizons (78%) while decision-tree evaluations largely used a pre-defined short-to-medium-term horizon (83%) (Figure 3.5c). The lifetime modelling horizon became more prevalent after 2010 (Figure 3.6a).

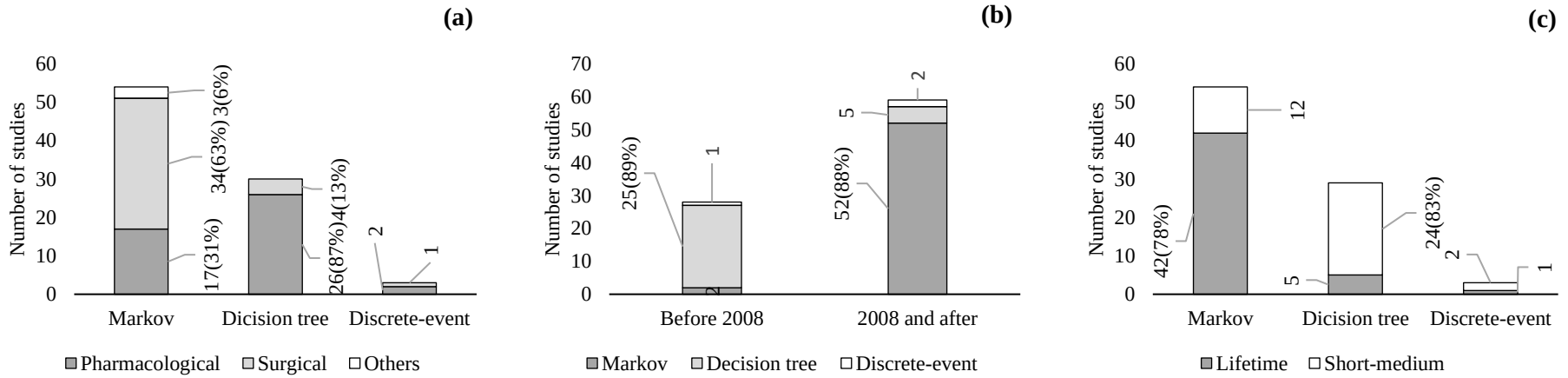


Figure 3.5. The number distribution of model types and treatments (a), by published time (b), and time horizon (c).

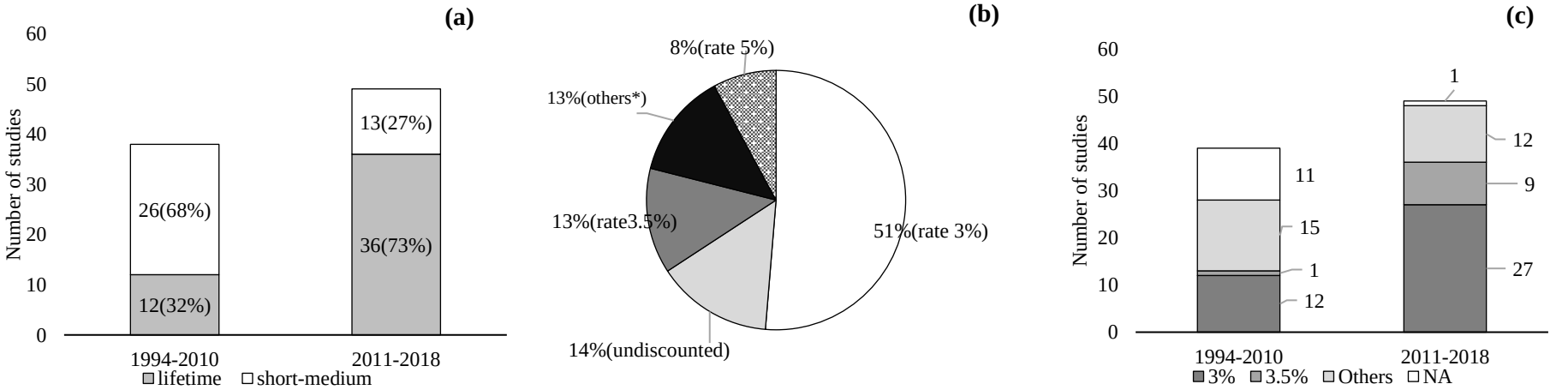


Figure 3.6. The distribution of adopted time horizon in included studies by published time (a), the percentage distribution of the adopted discount rates in the included modelling studies (n=76) (b), and the distribution of the adopted discount rates by published time (c). *Others included alternative discount rate for cost and outcomes, discount rate of 4%, 4.76%, and 6%.

3.5.10 Discount rates for costs and outcomes

Seventy-six (86%) studies reported the discount rate, which was most commonly 3% (n=39, 51%) for costs and outcomes (Figure 3.6b). Ten (13%) and six (8%) studies utilised 3.5% and 5% discount rates, respectively. Thirty (of 39) studies using a 3% discount rate were based in the US, and 9 (of 10) studies using a 3.5% discount rate were based in the UK. Seven used alternative discount rates for costs and outcomes. Twelve studies did not report use of a discount rate, with 11 of these published before 2010 (Figure 3.6c). The impact of varying discount rates on model outcomes was assessed by 15 of 88 (17%) included studies.

3.5.11 Model health states

Pharmacological-focused studies defined health states based on: 1) the occurrence of AEs/complications (43); 2) OA severity, the presence of joint pain, obesity and comorbidities (41); and 3) the Kellgren-Lawrence radiographic scale of OA (56). For surgical-focused studies, health states were defined based on: 1) the American College of Rheumatologists (ACR) functional status classification (47); 2) the event pathway following surgical treatment (57); 3) pain severity, postoperative complications, and subsequent surgical procedures (58); and 4) the Western Ontario and McMaster Osteoarthritis Index (WOMAC) scores (13).

All surgical-focused studies considered revision surgery, and 23 considered two or more surgical revisions. Nine studies also considered surgery-related complications (e.g.: infection, bleeding, dislocation).

For pharmacological-focused studies, the most important modelling event was medical AEs/regimen toxicity. Studies conducted before 2002 only included GI AEs, while post-2002 studies also considered CV AEs and renal toxicity. Twenty studies included CV AEs, and eleven considered discontinuation (mostly published after 2013). Only three studies considered treatment adherence (55, 59, 60).

3.5.12 Uncertainty analysis

Uncertainty analysis was reported in 86 (98%) studies, with most (n=47) conducting more than one type of sensitivity analysis. The adopted methods included: (a) deterministic (univariate and/or multivariate) sensitivity analysis (DSA) (n=68), (b) probability sensitivity analysis (PSA) (n=43), (c) threshold analysis (n=4), and (d) scenario analysis (n=2). Five studies did not report the type of sensitivity analysis. Whilst DSA was popular across the reporting periods,

PSA was first adopted in 2001 and became increasingly common thereafter. The most commonly evaluated parameters included costs, health state utilities (HSUs), probabilities of AEs, and treatment efficacy.

3.5.13 Reporting quality assessment

The reporting quality of all studies was assessed, with the exception of one study for which this was not applicable (61). The mean (standard deviation) score of all studies was 81% (9%). The reporting quality improved after 2001 and remained relatively high thereafter (Supplement 3.8a). Overall, 64 out of 87 studies exhibited a high reporting quality (>75%) (Supplement 3.8b). Whilst no study fully met CHEERS criteria, three achieved the highest reporting quality score of 96%. The CHEERS criteria were mostly met for the items of introduction, comparators, time horizon, resources and cost estimation, and uncertainty characterisation. However, the title (in terms of describing the interventions compared), abstract (in terms of the study perspective, setting, study inputs, and uncertainty analyses), and effectiveness measurement were commonly under-reported.

3.6 Discussion

This is the first study comprehensively reviewing the evolution of HEEMs for all forms of OA. Our review found that OA modelled evaluations are of a wide variety and have evolved substantially over time, with their emphasis and complexity shifting from pharmacological-focused short-to-medium term decision-tree models to surgical-focused lifetime Markov models. Existing HEEMs have limitations related to the choice of model input parameters, discount rates, and model health states/events. For instance, indirect costs related to OA were mostly not considered. Discount rates were mostly consistent with local guidelines, however, most studies failed to gauge the sensitivity of the model outcomes to discount rate changes. Most studies failed to consider important model events (e.g.: CV AEs), therapeutic adherence and treatment discontinuation. Despite clear guidelines (11), studies failed to pay adequate attention to lifestyle management, non-drug treatments and preventions. The reporting quality of included studies was reasonably satisfactory, however, the title, abstract, and effectiveness measures were mostly reported inadequately.

Cost categories considered in OA models should be consistent with the perspective (62). In this review, however, more than half (15 of 24) of the studies using a societal perspective failed to incorporate indirect costs, and of the 14 focused on patients of working age (< 65

years), only six considered productivity losses. Similarly, when modelling is for older populations (37-39), special consideration should be given to include the potentially large contribution of informal care costs (63), but only one study in this review included this. Given the increasing popularity of the societal perspective in recent years, we recommend future studies include all relevant costs, particularly the indirect costs from lost wages/productivity (64).

The focus of modelled evaluations of OA interventions changed over time. Prior to 2005, studies focused on pharmacological treatments; subsequently the focus shifted to surgical treatments. Our finding is consistent with the increasing popularity of surgical treatment for OA since 2000 (65) but the availability of relevant longitudinal data from national joint replacement registries may also have contributed to this rise (66). Most pharmacological studies focused on all types of OA combined and/or general arthritis patients, as there is no evidence of different treatment effects of drugs between OA types (43). In contrast, most studies of surgery focused on a specific type of OA (e.g.: knee and/or hip OA). Other (e.g.: ankle and glenohumeral) joints attracted limited attention, as might be expected due to the relatively low prevalence of OA at those joint sites (67, 68). As the studies focusing on ankle and glenohumeral joints were published before 2010, updated studies of these joint sites should be on the agenda for future research.

A limited number of HEEMs focused on lifestyle, non-drug treatments and preventions. This is despite guidelines (11, 69) recommending such interventions before pharmacological and surgical treatments. Because OA is a preventable, non-curable and progressive condition, it is critical that future studies investigate the impact of prevention and non-drug and non-surgical treatments on clinical and health-economic outcomes.

The most commonly used effectiveness measure was QALYs, with the proportion of studies reporting this increasing from 48% before 2008 to 93% thereafter. This is in accordance with national and international guidelines that recommend the use of QALYs (70). Importantly, the calculation of QALYs relies on HSUs that can be obtained from a variety of sources such as primary studies (37), systematic reviews (47) or randomized controlled trials (71). As different populations may value health states differently, caution is required when using non-locally derived HSUs (37, 72). Whilst disease-specific effectiveness measures (e.g.: revision free life years) provide information in more clinically relevant terms, these measures are not preference-based and also suffer from other limitations including the lack of comparability,

and difficulty in trading-off across different diseases (73). Thus, the transition to use of QALYs reflects best practice, and use of QALYs based on local HSUs is recommended in future studies.

More than half studies applied lifetime horizon; however, this only became popular after 2010. This change could be related to two factors: first, Markov models became more popular in recent years which tend to adopt lifetime horizons due to their ability to take into account the re-occurrence of model events (74); and second, CV AEs of pharmacological treatments were taken into account in recent years which, comparing with GI AEs, have a larger impact on mortality (60). As OA is a chronic condition with on-going medical management, we recommend use of lifetime horizons in future studies, especially when using a Markov model structure (75) and in situations where an intervention is expected to influence mortality rates (76).

Discount rates varied between study settings. The use of 3% and 3.5% discount rates in most US and UK based studies aligned well with the local guidelines (77, 78). However, 64% (9 out of 14) of studies conducted in other nations did not use discount rates from national guidelines (15, 47, 60, 79). Reporting of the discount rate has improved over time, which is as expected from the recent introduction of CHEERS statement for the development and appraisal of health economic evaluations (28). A significant majority (83%) of included studies did not conduct sensitivity analyses for discount rates, which contrasts recommendations (80). Future studies should choose a discount rate that is in-line with local guidelines and assess the sensitivity of the model outcomes to discount rate changes.

The choice of model events varied between studies and over time. Pharmacological-focused models evolved in their complexity by incorporating more AEs (such as CV AEs) and regimen discontinuation. However, the models of surgical treatments (with revision surgery being the most important model event) did not evolve to the same extent. Differences existed between studies in terms of the number of considered revisions. Considering the ten-yearly cumulative re-revision rates of primary total knee (22.8%) and hip replacement (21.5%) and the factors influencing these rates (81), the times of revision surgery should be decided based on the age of the target population and the surgery techniques of interest. The high rates of OA medical AEs and possibility of revision surgery impose significant additional costs (40), likely impacting the outcomes of cost-effectiveness analyses (22). Future HEEMs should therefore incorporate all relevant OA and treatment related events and complications/revisions.

The choice of modelling method varied between studies and over time, with Markov models becoming more popular after 2008. This is consistent with numerous decision analytic modelling guidelines (74, 76, 82), and suits the chronic nature of OA with the possibility of recurrent health events (75). Consistent with previous findings (83, 84), TreeAge was the single most popular modelling software followed by Microsoft Excel at all periods. Both software have been found equally reliable in conducting health-economic evaluations, and choice of software can depend on software availability, implementation skills, time constraints and end-user requirements (83, 84). We identified four popular OA HEEM structures, each with their own strengths and limitations. The original form of the NICE model included GI and CV AEs; however, treatment discontinuation and adherence were not included. The NICE model was extended in two studies by incorporating dose titration, discontinuation and additional AEs (44, 46). However, data on discontinuation is not easily available. The OAPoL model accounts for the inter-relationships among key variables such as the function of pain, obesity, and comorbidities. However, the inherent dependence of this model on scarce data for key variables limits its widespread use. The model developed by Fitzpatrick can be easily adapted to suit alternative settings and study objectives. However, it is considered too simple to fully evaluate the various outcomes between alternative surgeries (37, 85). Lastly, the model developed by Burke considered the severity of GI AEs and can easily be adapted to different settings (53, 54). However, it fails to consider other events such as CV AEs, discontinuation, and its time horizon of <1-year is not well suited to the chronic nature of OA.

PSA showed increasing popularity after 2001 which could be the result of an increasing awareness of PSA's importance in health economics over time (86) and advances in computational technologies. Limited studies evaluated the sensitivity of model outcomes to important input parameters including OA definitions and population parameters. We recommend considering these relatively neglected aspects in future modelling studies.

We recommend the use of Markov models due to their ability to incorporate repetitive (short and long-term) health events, including important medical AEs, therapeutic adherence, and discontinuation. Considering the short- or long-term nature of health events, the cycle length of Markov models should be adequate to represent the frequency of events and interventions (87). The choice of cycle length is determined by a number of factors, including the clinical problem, remaining life expectancy, availability of data, frequency of clinical follow-up, and computational efficiency (88). In cases where a relatively long (e.g., 1-year) cycle length is not adequate to capture the short-term impacts such as those of medical adverse events,

continuous-time multi-state models or discrete event simulation can be used. A previous study has shown that discrete-time multi-state models (MSMs) and continuous-time MSMs work equally well when the longitudinal observations are evenly distributed (89). However, when longitudinal observations are unevenly spaced, a discrete-time MSMs may be inaccurate. In this case, continuous-time MSMs may be better option. A continuous time MSM could be an alternative approach and should have a potential to being used much more by practitioners. However, the computational difficulties facing continuous-time models, limited availability of specialised software to build these models and poor model transparency has resulted in limiting their widespread use. Given that improper use of Markov models may result in biased estimations, perhaps some standardization in the reporting of MSM results and assumption verification is needed. Importantly, the probabilities of medical AEs, and time to events (e.g.: decision for joint replacement, revision surgery) depend on the history of previous states; however, no study in our review considered this important dimension. We therefore recommend future studies to integrate memory into their models to avoid problems associated with the Markov assumption of memoryless-ness, for example, the cohort-level Markov model can integrate memory feature by including tunnel states or tracker variables or building time dependency by incorporating an additional time dimension (90) or using the individual-level microsimulation models which incorporate individual clinical pathways and history of previous events by randomly walking one individual at a time through the model. As the choice of cohort-level Markov models or individual-based microsimulation models depends on data availabilities and other factors (e.g.: software availability), the use of individual level microsimulation may not always be possible. Future HEEMs should also benefit from the recent availability of MRI-based data on OA definitions, progression and MRI-based markers (19), and advances in new data science (that have enabled the use of machine learning-based patient-specific prediction models) (91).

This review found the reporting quality of studies has improved and has been reasonably satisfactory over time. Nonetheless, further improvements could be made, particularly, in relation to the quality of title, abstract, and effectiveness measures. As poor reporting may lead to costly decisions, future studies should ensure high transparency and reporting quality in all areas.

The strengths of our review include its comprehensive nature and the incorporation of assessment of reporting quality. The inclusion of all OA therapies builds on two existing systematic reviews of OA oral therapies and surgical interventions (21, 22). Furthermore, our

review summarised the development of OA models in terms of various model characteristics, which will help to evaluate the existing OA HEEMs and guide the development of a comprehensive gold standard HEEM of OA. Lastly, our assessment of reporting quality should be of interest to future researchers in improving the reporting quality of their health-economic modelling studies. A limitation is that our study did not cover studies published in languages other than English, Chinese and German. However, this may have a minimal impact on our key conclusions as only a small number of studies (n=20) were subject to language exclusions. A further limitation is that as we only had a small number of studies comparing OA preventative interventions, the review is predominantly focused on OA treatments. Future studies will need to consider the synthesis of OA prevention models as more evidence becomes available.

Conclusion

OA HEEMs are of a wide variety and evolved substantially over time. Furthermore, the number of modelled OA evaluations have rapidly increased in recent years. The focus of OA HEEs has shifted from short-to-medium-term pharmacological-focused decision-tree models to surgical-focused lifetime Markov models. We recommend future HEEMs use life-time Markovian model structures with memory integration and should also incorporate all relevant costs, model events, therapeutic adherence, discontinuation, appropriate discount rate and time horizon, and conduct sensitivity analyses for input parameters. Finally, we recommend improvements in relation to the reporting quality of (1) title, (2) abstract, and (3) effectiveness measurement.

Authors' contributions

TZ, BdeG, DA, TW and AP conceived and designed the methods. TZ and QX conducted literature search and screening, TZ and HA conducted the data extraction and reporting quality assessment. All authors were involved in the manuscript preparation. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the librarian researcher Michaela for her help on building the literature search strategy and Covidence Help Centre for facilitating the progression of literature screening.

3.7 References

1. Osteoarthritis Research Society International. Osteoarthritis: a serious disease. https://www.oarsi.org/sites/default/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf.
2. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity among adults and youth: United States, 2015-2016. NCHS Data Brief. 2017(288):1-8.
3. Sisson SB, Katzmarzyk PT. International prevalence of physical activity in youth and adults. Obes Rev. 2008;9(6):606-14.
4. Wilson R, Blakely T, Abbott JH. Radiographic knee osteoarthritis impacts multiple dimensions of health-related quality of life: data from the Osteoarthritis Initiative. Rheumatology (Oxford). 2018;57(5):891-9.
5. Huang Z, Ding C, Li T, Yu SP. Current status and future prospects for disease modification in osteoarthritis. Rheumatology (Oxford). 2018;57(suppl_4):iv108-iv23.
6. Abbott JH, Usiskin IM, Wilson R, Hansen P, Losina E. The quality-of-life burden of knee osteoarthritis in New Zealand adults: A model-based evaluation. PLoS One. 2017;12(10):e0185676.
7. Henderson JV, Harrison CM, Britt HC, Bayram CF, Miller GC. Prevalence, causes, severity, impact, and management of chronic pain in Australian general practice patients. Pain Med. 2013;14(9):1346-61.
8. March LM, Bachmeier CJ. Economics of osteoarthritis: a global perspective. Baillieres Clin Rheumatol. 1997;11(4):817-34.
9. Puig-Junoy J, Ruiz Zamora A. Socio-economic costs of osteoarthritis: A systematic review of cost-of-illness studies. Semin Arthritis Rheum. 2015;44(5):531-41.
10. Sharif B, Garner R, Hennessy D, Sanmartin C, Flanagan WM, Marshall DA. Productivity costs of work loss associated with osteoarthritis in Canada from 2010 to 2031. Osteoarthritis Cartilage. 2017;25(2):249-58.
11. The Royal Australian College of General Practitioners. Guideline for the management of knee and hip osteoarthritis. 2nd edn. East Melbourne, Vic: RACGP, 2018.

12. Rabarison KM, Bish CL, Massoudi MS, Giles WH. Economic evaluation enhances public health decision making. Frontiers in public health. 2015;3:164.

13. Losina E, Walensky RP, Kessler CL, Emrani PS, Reichmann WM, Wright EA, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. Arch Intern Med. 2009;169(12):1113-21; discussion 21-2.

14. Brereton N, Winn B, Akehurst R. The cost-effectiveness of celecoxib vs diclofenac in the treatment of osteoarthritis in the UK; an update to the NICE model using data from the CONDOR trial. J Med Econ. 2012;15(3):465-72.

15. Higashi H, Barendregt JJ. Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis: discrete-event simulation model. PLoS One. 2011;6(9):e25403.

16. Chang RW, Pellisier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. Jama. 1996;275(11):858-65.

17. Haglund U, Svarvar P. The Swedish ACCES model: predicting the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis. Rheumatology (Oxford). 2000;39 Suppl 2:51-6.

18. Latimer N, Lord J, Grant RL, O'Mahony R, Dickson J, Conaghan PG, et al. Cost effectiveness of COX 2 selective inhibitors and traditional NSAIDs alone or in combination with a proton pump inhibitor for people with osteoarthritis. BMJ. 2009;339:b2538.

19. Losina E, Daigle ME, Suter LG, Hunter DJ, Solomon DH, Walensky RP, et al. Disease-modifying drugs for knee osteoarthritis: can they be cost-effective? Osteoarthritis Cartilage. 2013;21(5):655-67.

20. Latimer N, Lord J, Grant RL, O'Mahony R, Dickson J, Conaghan PG. Value of information in the osteoarthritis setting: cost effectiveness of COX-2 selective inhibitors, traditional NSAIDs and proton pump inhibitors. Pharmacoeconomics. 2011;29(3):225-37.

21. Kamaruzaman H, Kinghorn P, Oppong R. Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature. BMC Musculoskelet Disord. 2017;18(1):183.

22. Wielage RC, Myers JA, Klein RW, Happich M. Cost-effectiveness analyses of osteoarthritis oral therapies: a systematic review. Appl Health Econ Health Policy. 2013;11(6):593-618.

23. Pinto D, Robertson MC, Hansen P, Abbott JH. Cost-effectiveness of nonpharmacologic, nonsurgical interventions for hip and/or knee osteoarthritis: systematic review. Value Health. 2012;15(1):1-12.

24. Si L, Winzenberg TM, Palmer AJ. A systematic review of models used in cost-effectiveness analyses of preventing osteoporotic fractures. Osteoporos Int. 2014;25(1):51-60.

25. Si L, Winzenberg TM, Jiang Q, Palmer AJ. Screening for and treatment of osteoporosis: construction and validation of a state-transition microsimulation cost-effectiveness model. Osteoporos Int. 2015;26(5):1477-89.

26. Moher D, Liberati A, Tetzlaff J, Altman DG, Group atP. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Annals of Internal Medicine. 2009;151(4):264-9.

27. Better systematic reivew management C, accessed: https://www.covidence.org/home, dated 29 January 2019.

28. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. Bmj. 2013;346:f1049.

29. de Graaff B, Neil A, Sanderson K, Si L, Yee KC, Palmer AJ. A systematic review and narrative synthesis of health economic studies conducted for hereditary haemochromatosis. Appl Health Econ Health Policy. 2015;13(5):469-83.

30. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version. 2006;1:b92.

31. McKell C, Stewart A. Cost-minimisation analysis comparing topical versus systematic NSAIDS in the treatment of mild osteoarthritis of the superficial joints. British Journal of Medical Economics. 1994;7:137-46.

32. Faucett SC, Geisler BP, Chahla J, Krych AJ, Kurzweil PR, Garner AM, et al. Meniscus root repair vs meniscectomy or nonoperative management to prevent knee osteoarthritis after medial meniscus root tears: clinical and economic effectiveness. Am J Sports Med. 2018:363546518755754.

33. Rongen JJ, Govers TM, Buma P, Grutters JP, Hannink G. Societal and economic effect of meniscus scaffold procedures for irreparable meniscus injuries. Am J Sports Med. 2016;44(7):1724-34.

34. Feeley BT, Liu S, Garner AM, Zhang AL, Pietzsch JB. The cost-effectiveness of meniscal repair versus partial meniscectomy: A model-based projection for the United States. The Knee. 2016;23(4):674-80.

35. Losina E, Burbine SA, Suter LG, Hunter DJ, Solomon DH, Daigle ME, et al. Pharmacologic regimens for knee osteoarthritis prevention: can they be cost-effective? Osteoarthritis Cartilage. 2014;22(3):415-30.

36. SooHoo NF, Kominski G. Cost-effectiveness analysis of total ankle arthroplasty. J Bone Joint Surg Am. 2004;86-a(11):2446-55.

37. Kunkel ST, Sabatino MJ, Kang R, Jevsevar DS, Moschetti WE. The cost-effectiveness of total hip arthroplasty in patients 80 years of age and older. J Arthroplasty. 2018;33(5):1359-67.

38. Katz JN, Smith SR, Collins JE, Solomon DH, Jordan JM, Hunter DJ, et al. Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities. Osteoarthritis Cartilage. 2016;24(3):409-18.

39. Slover J, Espehaug B, Havelin LI, Engesaeter LB, Furnes O, Tomek I, et al. Cost-effectiveness of unicompartmental and total knee arthroplasty in elderly low-demand patients. A Markov decision analysis. J Bone Joint Surg Am. 2006;88(11):2348-55.

40. Losina E, Usiskin IM, Smith SR, Sullivan JK, Smith KC, Hunter DJ, et al. Cost-effectiveness of generic celecoxib in knee osteoarthritis for average-risk patients: a model-based evaluation. Osteoarthritis Cartilage. 2018;26(5):641-50.

41. Smith SR, Katz JN, Collins JE, Solomon DH, Jordan JM, Suter LG, et al. Cost-effectiveness of tramadol and oxycodone in the treatment of knee osteoarthritis. Arthritis Care Res (Hoboken). 2017;69(2):234-42.

42. Losina E, Michl G, Collins JE, Hunter DJ, Jordan JM, Yelin E, et al. Model-based evaluation of cost-effectiveness of nerve growth factor inhibitors in knee osteoarthritis: impact of drug cost, toxicity, and means of administration. Osteoarthritis Cartilage. 2016;24(5):776-85.

43. Nasef SA, Shaaban AA, Mould-Quevedo J, Ismail TA. The cost-effectiveness of celecoxib versus non-steroidal anti-inflammatory drugs plus proton-pump inhibitors in the treatment of osteoarthritis in Saudi Arabia. Health Econ Rev. 2015;5(1):53.

44. Wielage RC, Patel AJ, Bansal M, Lee S, Klein RW, Happich M. Cost effectiveness of duloxetine for osteoarthritis: a Quebec societal perspective. Arthritis care & research. 2014;66(5):702-8.

45. Brereton N, Pennington B, Ekelund M, Akehurst R. A cost-effectiveness analysis of celecoxib compared with diclofenac in the treatment of pain in osteoarthritis (OA) within the Swedish health system using an adaptation of the NICE OA model. J Med Econ. 2014;17(9):677-84.

46. Wielage RC, Bansal M, Andrews JS, Klein RW, Happich M. Cost-utility analysis of duloxetine in osteoarthritis: a US private payer perspective. Appl Health Econ Health Policy. 2013;11(3):219-36.

47. Mujica-Mota RE, Watson LK, Tarricone R, Jager M. Cost-effectiveness of timely versus delayed primary total hip replacement in Germany: A social health insurance perspective. Orthop Rev (Pavia). 2017;9(3):7161.

48. Pulikottil-Jacob R, Connock M, Kandala NB, Mistry H, Grove A, Freeman K, et al. Has metal-on-metal resurfacing been a cost-effective intervention for health care providers?-A registry based study. PLoS One. 2016;11(11):e0165021.

49. Pulikottil-Jacob R, Connock M, Kandala NB, Mistry H, Grove A, Freeman K, et al. Cost effectiveness of total hip arthroplasty in osteoarthritis: comparison of devices with differing bearing surfaces and modes of fixation. Bone Joint J. 2015;97-b(4):449-57.

50. Mota RE. Cost-effectiveness analysis of early versus late total hip replacement in Italy. Value Health. 2013;16(2):267-79.

51. Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, et al. Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses. Health Technol Assess. 1998;2(20):1-64.

52. Chancellor JV, Hunsche E, de Cruz E, Sarasin FP. Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor, in Switzerland. Pharmacoeconomics. 2001;19(1):59-75.

53. Zabinski RA, Burke TA, Johnson J, Lavoie F, Fitzsimon C, Tretiak R, et al. An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada. Pharmacoeconomics. 2001;19 Suppl 1:49-58.

54. You JH, Lee KK, Chan TY, Lau WH, Chan FK. Arthritis treatment in Hong Kong--cost analysis of celecoxib versus conventional NSAIDS, with or without gastroprotective agents. Aliment Pharmacol Ther. 2002;16(12):2089-96.

55. Al MJ, Maniadakis N, Grijseels EW, Janssen M. Costs and effects of various analgesic treatments for patients with rheumatoid arthritis and osteoarthritis in the Netherlands. Value Health. 2008;11(4):589-99.

56. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494-502.

57. Chawla H, Nwachukwu B, Van Der List J, Eggman A, Pearle A, Ghomrawi H. Cost effectiveness of patellofemoral versus total knee arthroplasty in younger patients. The bone & joint journal. 2017;99(8):1028-36.

58. Konopka JF, Gomoll AH, Thornhill TS, Katz JN, Losina E. The cost-effectiveness of surgical treatment of medial unicompartmental knee osteoarthritis in younger patients: a computer model-based evaluation. The Journal of bone and joint surgery American volume. 2015;97(10):807.

59. Capel M, Tornero J, Zamorano JL, Oyaguez I, Casado MA, Sanchez-Covisa J, et al. Efficiency of naproxen/esomeprazole in association for osteoarthrosis treatment in Spain. Reumatol Clin. 2014;10(4):210-7.

60. Bessette L, Risebrough N, Mittmann N, Roussy J-P, Ho J, Zlateva G. Cost-utility of celecoxib use in different treatment strategies for osteoarthritis and rheumatoid arthritis from the Quebec healthcare system perspective. Journal of medical economics. 2009;12(3):246-58.

61. Wilson R, Abbott JH. Development and validation of a new population-based simulation model of osteoarthritis in New Zealand. Osteoarthritis Cartilage. 2018;26(4):531-9.

62. Garrison LP, Pauly MV, Willke RJ, Neumann PJ. An overview of value, perspective, and decision context—a health economics approach: an ISPOR Special Task Force report [2]. Value in Health. 2018;21(2):124-30.

63. Gupta S, Hawker GA, Laporte A, Croxford R, Coyte PC. The economic burden of disabling hip and knee osteoarthritis (OA) from the perspective of individuals living with this condition. Rheumatology (Oxford). 2005;44(12):1531-7.

64. Koenig L, Zhang Q, Austin MS, Demiralp B, Fehring TK, Feng C, et al. Estimating the societal benefits of THA after accounting for work status and productivity: a Markov model approach. Clinical Orthopaedics and Related Research®. 2016;474(12):2645-54.

65. Maradit Kremers H, Larson DR, Crowson CS, Kremers WK, Washington RE, Steiner CA, et al. Prevalence of total hip and knee replacement in the United States. J Bone Joint Surg Am. 2015;97(17):1386-97.

66. Ayers DC, Franklin PD. Joint replacement registries in the United States: a new paradigm. The Journal of bone and joint surgery American volume. 2014;96(18):1567.

67. Murray C, Marshall M, Rathod T, Bowen CJ, Menz HB, Roddy E. Population prevalence and distribution of ankle pain and symptomatic radiographic ankle osteoarthritis in community dwelling older adults: A systematic review and cross-sectional study. PLoS One. 2018;13(4):e0193662.

68. Cameron ML, Kocher MS, Briggs KK, Horan MP, Hawkins RJ. The prevalence of glenohumeral osteoarthrosis in unstable shoulders. The American journal of sports medicine. 2003;31(1):53-5.

69. Conditions NCCfC, Excellence NIfC, editors. Osteoarthritis: national clinical guidelines for care and management in adults2008: Royal College of Physicians.

70. Hiligsmann M, Cooper C, Arden N, Boers M, Branco JC, Brandi ML, et al., editors. Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Seminars in arthritis and rheumatism; 2013: Elsevier.

71. Bove AM, Smith KJ, Bise CG, Fritz JM, Childs J, Brennan GP, et al. Exercise, manual therapy, and booster sessions in knee osteoarthritis: cost-effectiveness analysis from a multicenter randomized controlled trial. Phys Ther. 2018;98(1):16-27.

72. Peersman G, Jak W, Vandenlangenbergh T, Jans C, Cartier P, Fennema P. Cost-effectiveness of unicondylar versus total knee arthroplasty: a Markov model analysis. Knee. 2014;21 Suppl 1:S37-42.

73. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied methods of cost-effectiveness analysis in healthcare: Oxford University Press; 2011.

74. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. Journal of health services research & policy. 2004;9(2):110-8.

75. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. NIHR Health Technology Assessment programme: Executive Summaries: NIHR Journals Library; 2004.

76. Drugs CAf, Health Ti. Guidelines for the economic evaluation of health technologies: Canada. Guidelines for the economic evaluation of health technologies: Canada: CADTH; 2006.

77. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. Jama. 1996;276(15):1253-8.

78. Smith WB, Steinberg J, Scholtes S, Mcnamara IR. Medial compartment knee osteoarthritis: age-stratified cost-effectiveness of total knee arthroplasty, unicompartmental knee arthroplasty, and high tibial osteotomy. Knee Surgery, Sports Traumatology, Arthroscopy. 2017;25(3):924-33.

79. Attema AE, Brouwer WBF, Claxton K. Discounting in economic evaluations. Pharmacoeconomics. 2018;36(7):745-58.

80. Adalsteinsson E, Toumi M. Benefits of probabilistic sensitivity analysis–a review of NICE decisions. Journal of market access & health policy. 2013;1(1):21240.

81. Australian Orthopaedic Association, National Joint Replacement Registry, Hip and Knee Arthroplasty, https://aoanjrr.sahmri.com/documents/10180/172286/Annual+Report+2014, Accessed: March 15 2019. .

82. Soto J. Health economic evaluations using decision analytic modeling: principles and practices—utilization of a checklist to their development and appraisal. International journal of technology assessment in health care. 2002;18(1):94-111.

83. Tsoi B, O'Reilly D, Jegathisawaran J, Tarride J, Blackhouse G, Goeree R. Systematic narrative review of decision frameworks to select the appropriate modelling approaches for health economic evaluations. BMC research notes. 2015;8(1):244.

84. Menn P, Holle R. Comparing three software tools for implementing markov models for health economic evaluations. Pharmacoeconomics. 2009;27(9):745-53.

85. Ghomrawi HM, Eggman AA, Pearle AD. Effect of age on cost-effectiveness of unicompartmental knee arthroplasty compared with total knee arthroplasty in the US. The Journal of bone and joint surgery American volume. 2015;97(5):396.

86. Baio G, Dawid AP. Probabilistic sensitivity analysis in health economics. Statistical Methods in Medical Research. 2015;24(6):615-34.

87. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force–3. Medical Decision Making. 2012;32(5):690-700.

88. Chhatwal J, Jayasuriya S, Elbasha EH. Changing cycle lengths in state-transition models: challenges and solutions. Medical Decision Making. 2016;36(8):952-64.

89. Wan L, Lou W, Abner E, Kryscio RJ. A comparison of time-homogeneous Markov chain and Markov process multi-state models. Communications in Statistics: Case Studies, Data Analysis and Applications. 2016;2(3-4):92-100.

90. Hawkins N, Sculpher M, Epstein D. Cost-effectiveness analysis of treatments for chronic disease: using R to incorporate time dependency of treatment response. Medical Decision Making. 2005;25(5):511-9.

91. Jamshidi A, Pelletier J-P, Martel-Pelletier J. Machine-learning-based patient-specific prediction models for knee osteoarthritis. Nature Reviews Rheumatology. 2019;15(1):49-60.

3.8 Supplements

Supplement 3A: The publication of "Chapter 3: A Systematic Review of the Evolution of Health-Economic Evaluation Models of Osteoarthritis"

Chapter 3A has been removed for copyright or proprietary reasons.

It is the following published article: Zhao, T., Ahmad, H., de Graaff, B., Xia, Q., Winzenberg, T., Aitken, D., Palmer, A. J., 2021., Systematic review of the evolution of health-economic evaluation models of osteoarthritis, Arthritis care & research, 73(11), 1617-1627

Supplement 3.1 Types of health economic evaluations

Economic evaluation is defined as "the systematic appraisal of costs and benefits of projects, normally undertaken to determine the relative economic efficiency of programs." In simple words, economic evaluation is the understanding and use of economic evidence in decision making. There are two levels of economic evaluations: 1) partial and 2) full. Partial economic evaluation measures program or disease costs/outcomes but does not involve a comparison with alternative options and does not relate costs to outcomes (e.g.: cost-of-illness analysis and program cost analysis). On the other hand, full economic evaluations compare two or more public health interventions through the examination of costs of inputs and outcomes. Full economic evaluations include cost-benefit, cost-effectiveness, cost-utility, cost-consequences, and cost-minimization analyses.

Type	Description	Cost measurement	Outcome measurement
Partial economic evaluations			
Cost of illness	disease economic burden	$	–
Program cost analysis	Net program cost	$	–
Full economic evaluations			
Cost-benefit analysis	Compares different programs with different outcomes (e.g., health vs. other area)	$	$
Cost-effectiveness analysis	Compares interventions with the same outcomes	$	Single "natural" unit outcome measure
Cost-utility analysis	Compares interventions with different health outcomes	$	Multiple outcomes—life-years adjusted for quality-of-life
Cost-minimization analysis	Compares the costs of alternative interventions that have equal effects	$	Equivalence demonstrated or assumed in comparative groups
Cost-consequences analysis	Lists separately all the direct and indirect costs and catalogues different outcomes of all alternatives, with no specific preference for one costing approach/outcome measure (as is the case for cost-effectiveness analysis or cost-utility analysis)	$	Multi-dimensional listing of outcomes

Supplement 3.2 The literature search strategy for Medline*

1. Cost effective*.mp.

2. Cost-effective*.mp.

3. Cost utility.mp.

4. cost-utility.mp.

5. Cost benefit.mp.

6. cost-benefit.mp.

7. Cost*1.mp.

8. quality-adjusted life years.mp.

9. Health-economic*1.mp.

10. Economic evaluation.mp.

11. Cost-Benefit Analysis/

12. Model*1.mp.

13. exp models, economic/

14. Osteoarthritis.mp.

15. OA.mp.

16. exp Osteoarthritis/

17. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

18. 12 or 13

19. 14 or 15 or 16

20. 17 and 18 and 19

Note: * The search strategy used for Medline was adapted to suite other databases.

Supplement 3.3

Inclusion criteria

1. Studies in humans;
2. Studies reporting the construction, validation and application of health economic evaluation models for OA interventions. Both the partial (cost of illness, program costs) and full (cost-effectiveness, cost-benefit, cost-utility, cost-minimization, and cost-consequences analysis) economic evaluation studies were included;
3. Studies that were available as full text;
4. Studies published in English, Chinese, or German languages;
5. Studies focusing on arthritis population including OA when the proportion of OA participants was not available or when the proportion of OA participants was available, and ≥90% participants in the sample had OA.

Exclusion criteria

1. Studies in animals;
2. Studies that did not report OA-related health-economic evaluation models (e.g.: non-model-based cost of illness or cost-effectiveness studies);
3. Conference abstracts;
4. When studies focused on arthritis population including OA and reported the proportion of OA participants being <90%;
5. Studies published in languages other than English, Chinese and German (e.g.: Spanish and Italian languages);
6. Review articles;
7. Comments; and
8. Books/book chapters.

Supplement 3.4 The extracted information from included studies

Title of the studies	Authors	Setting	OA type	Comparators	Population	Perspective	Time horizon	Discount rate	Effectiveness measures	Direct cost
Nabumetone compared with Ibuprofen and a weighted NSAID combination an economic evaluation	Akehurst (1998)	UK	arthritis	nabumetone with NSAIDs	not mentioned	NHS perspective	3 months	6	life year gained	the resource use and cost of NSAID and treating side effects
Costs and effects of various analgesic treatments for patients with rheumatoid arthritis and osteoarthritis in the Netherlands	A MJ (2008)	Netherlands	arthritis	various analgesic treatments (celecoxib, NSAIDs alone, NSAID +misoprostol, NSAID +H2RA, NSAID +PPI, Arthrotec)	RA and OA patients	societal perspective	6 months	not mentioned	"GI events averted" and "life-years saved"	only considers direct medical costs included drug costs and resource use associated with GI side effects
Economic benefit to society at large of total knee arthroplasty in younger patients a Markov analysis	Bedair (2014)	US	knee	TKA vs nonoperative treatment	fifty-year-old patients with end-stage OA	societal perspective	30 years (lifetime), one year cycle	3	NA	The direct costs associated with nonoperative management and those associated with TKA were considered

Cost-utility of celecoxib use in different treatment strategies for osteoarthritis and rheumatoid arthritis from the Quebec healthcare system perspective	Bessette (2009)	Canada	arthritis	celecoxib in three treatment strategies which were defined based on the occurrence of GI or CV events	two patients subpopulations: arthritis patients aged <65 years (at low/average risk of GI and CV events), those aged ≥65 years (at high risk of GI and CV events)	third-party payer	5 years, monthly cycle	3	QALY	The direct costs included medication, hospitalisations and Aes.
Exercise, Manual Therapy, and Booster Sessions in Knee Osteoarthritis Cost-Effectiveness Analysis from a Multicenter Randomized Controlled Trial	Bove (2017)	US	knee	exercise only (EX), EX plus booster sessions(B), EX plus manual therapy (MT), EX+MT+B	40 years old or older and met the American College of Rheumatology criteria for knee OA	societal perspective	2 years and 5 years, monthly cycle	3	QALYs	direct medical costs and nonmedical costs
The cost effectiveness of celecoxib vs diclofenac in the treatment of osteoarthritis in the UK an update to the NICE model using data from the CONDOR	Brereton (2012)	UK	OA	celecoxib+PPI vs diclofenac+PPI	not mentioned	NHS perspective	3-month time horizon was used in the base case. Lifetime horizon	3.5	QALY	drug costs, AE management costs
A cost effectiveness analysis of celecoxib compared with diclofenac in the treatment of pain in osteoarthritis OA	Brereton (2014)	Sweden	OA	celecoxib, diclofenac, celecoxib+PPI, diclofenac+PPI	not mentioned	healthcare system perspective	lifetime	3	QALY	Costs included in the model consisted of drug acquisition cost and treatment of AEs

within the Swedish health system										
The cost-effectiveness of meniscal repair versus partial meniscectomy A model-based projection for the United States	Brian (2016)	US	meniscal tears	meniscal repair vs. meniscectomy	baseline age assumed to be 37.7	Medicare third-party perspective	30-years horizon	3	QALYs	direct healthcare costs
The cost-effectiveness of dual mobility implants for primary total hip arthroplasty A computer-based cost-utility model	Brian (2017)	US	hip	Dual mobility implants with conventional bearings	65-year-old patients began with primary unilateral hip arthroplasty	societal perspective	lifetime	3	QALYs	direct costs
Cost-effectiveness Study of Celecoxib for Osteoarthritis in China	Bruce (2017)	China	OA	celecoxib vs. diclofenac +PPI	55-year-old OA population		lifetime	4.76	QALYs	medical costs
Cost-effectiveness of unicompartmental compared with total knee replacement a population based study using data from the National Joint Registry for England and Wales	Burn (2018)	England and Wales	knee	unicompartmental knee replacement with TKR	six age group of patients who could receive either a UKA or TKR (<60 years, 60-75 years, and >75 years)	health system perspective	lifetime	3.5	QALYs	medical costs

Efficiency of naproxenesomeprazole in association for osteoarthrosis treatment in Spain	Capel (2014)	Spain	OA	naproxen/esomeprazole compared to other NSAID with or without PPI (celecoxib/+PPI; diclofenac+PPI; etoricoxib/+PPI; ibuprofen+PPI; naproxen+PPI; paracetamol)	patient over 65 years of age with OA and increased GI risk, defined as a history of ulcer in the upper GI tract	NHS perspective	one year, cycles of 3 months	undiscounted	QALY	include only the costs associated with pharmacological treatment and management derived from clinical episodes.
Cost-effectiveness Analysis of Viscosupplementation versus Conventional Supportive Therapy for Knee Osteoarthritis in Colombia	Castro (2015)	Colombia	knee	viscosupplementation (hylan G-F 20) alone with conventional supportive therapy	patients aged between younger than 50 years and older than 80 years	third-party payer	treatment outcomes were simulated at different time horizons in the interval of 5 to 20 years	3	QALYs	drugs, diagnostic tests, procedures, and hospitalisation, other direct costs generated by medical services
Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor, in Switzerland	Chancellor (2001)	Switzerland	arthritis	6 treatments: celecoxib,NSAID alone, NSAID with PPI or H2RA or misoprostol or diclofenac	90% with OA and 10% with RA	public health insurers	6 months	not mentioned	GI events adverted	drug acquisition and GI adverse event management
A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip	Chang (1996)	US	hip	total hip arthroplasty vs. nonoperative strategies	not mentioned	societal perspective	lifetime	3	QALYs	direct medical costs

Cost effectiveness of patellofemoral versus total knee arthroplasty in younger patients	Chawla (2017)	US	isolated patellofemoral OA	Patellofemoral arthroplasty versus TKA	patients were aged 60 (base case) and 50 years	healthcare payer perspective	lifetime/ one year cycle length	3	QALYs	direct cost
Cost-effectiveness analysis for joint pain treatment in patients with osteoarthritis treated at the Instituto Mexicano del Seguro Social (IMSS)	Contreras-Hernández (2008)	Mexican	OA	celecoxib, nonselective NSAIDs (naproxen, diclofenac and piroxicam) and acetaminophen	knee and/or hip OA patients treated at Instituto Mexicano del Seguro Social (IMSS)	Mexican institutional perspective	6 months	undiscounted	number of patients with pain control and no adverse events per each 1000 patients	direct medical costs
Modeling the cost-effectiveness for cement-less and hybrid prosthesis in total hip replacement in emilia romagna, Italy	Di Tanna (2011)	Italy	hip	cement-less with hybrid prosthesis in THR	not mentioned	provider perspective	lifetime	3.5	revision free life year	acquisition cost
Meniscus Root Repair vs Meniscectomy or Nonoperative Management to Prevent Knee Osteoarthritis After Medial Meniscus	Faucett (2018)	US	Medial meniscus root tears	meniscus repair, meniscectomy, and nonoperative treatment	55-yearold patients presenting with medial meniscus root tears with no osteoarthritis at the time of treatment	third-party payer	30 years, monthly cycle	3	QALYs	Medicare reimbursement as a proxy for cost

Root Tears Clinical and Economic Effectiveness										
Effect of age on cost-effectiveness of unicompartmental knee arthroplasty compared with total knee arthroplasty in the U.S	Ghomrawi (2015)	US	knee	UKA vs TKA	patients forty-five through eight-five years of age	societal perspective	lifetime	3	QALYs	medical costs and rehabilitation costs
The Swedish ACCES model predicting the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis	Haglund (2000)	Sweden	arthritis	celecoxib	NA		not mentioned	5	event averted, life-year gained	
Cost-effectiveness analysis of intra-articular injections of a high molecular weight bioengineered hyaluronic acid for the treatment of osteoarthritis knee pain	Hatoum (2014)	US	knee	bioengineered hyaluronic acid vs conventional care	mean patient age was 61.7 years	payer perspective	52 weeks	not mentioned	QALY	BioHA injection costs, fee of physician visits and 2 courses of 3 injectable drug administrations
Cost-utility of metal-on-metal hip resurfacing compared to conventional total hip replacement in young active patients with osteoarthritis	Heintzberger (2013)	Canada	hip	MoM HRA vs. THR	patients with OA aged 50 years, seven age- or sex-specific subgroups were analyses	healthcare payer perspective	15 years	3	QALY	medical costs

Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis discrete-event simulation model	Higashi (2011)	Australian	OA	total hip and knee replacements compared with 'doing nothing'(continued non-surgical therapies without joint replacements)	patients aged 40 or over has at least one joint with moderate OA or worse	health system perspective	lifetime	3	DALY	all costs that fall on the health sector with the interventions were included, both in the government and private sectors
An economic evaluation of meloxicam 7.5 mg versus diclofenac 100 mg retard in the treatment of osteoarthritis in the UK A decision analysis model based on gastrointestinal complications	Jansen (1996)	UK	OA	meloxicam vs. diclofenac	NA	NHS perspective	30 days	not mentioned	NA	direct medical costs incurred with the initial treatment and associated with GI AE
Economic evaluation of meloxicam (7.5 mg) versus sustained release diclofenac (100 mg) treatment for osteoarthritis A cross-national assessment for France, Italy and the UK	Jansen (1997)	France, Italy and UK	OA	meloxicam vs. diclofenac	not mentioned	NHS in UK, French statutory health insurance, Italian NHS perspective	30 days	not mentioned	NA	resources utilisation related to consultations and hospitalisation, drug costs
Cost-effectiveness of antibiotic-impregnated bone	Justin (2009)	US	hip	antibiotic-impregnated bone cement for total hip	68 years of age patients	not mentioned	lifetime	3	QALYs	cost included procedure,

cement used in primary total hip arthroplasty				arthroplasty with cement without antibiotics						hospitalisation
The cost-effectiveness of acetaminophen, NSAIDs, and selective COX-2 inhibitors in the treatment of symptomatic knee osteoarthritis	Kamath (2003)	US	knee	rofecoxib and celecoxib compared with high-dose acetaminophen or ibuprofen	50+ who had radiographically identified knee OA (higher risk)	institutional/payer perspective	6 months	undiscounted	number of upper GI Aes averted, number of patients who achieved perceptible pain relief.	direct medical cost
A Fresh Perspective on a Familiar Problem Examining Disparities in Knee Osteoarthritis Using a Markov Model	Karmarkar (2017)	US	knee	10 types of treatments	NA	patients, employer, and society	40 years (lifetime)/ one year cycle	3	QALYS	measured as Medicare reimbursement rates, or collected from literature or made assumption
Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities	Katz (2016)	US	knee	NSAIDs and opioids as implemented to SOC	older patients with CVD and diabetes (with mean age 74)	not mentioned	lifetime	3	QALYs	medical cost
Do the potential benefits of metal-on-metal hip resurfacing justify the increased cost and risk of complications	Kevin (2010)	not mentioned	hip	metal-on-metal hip surfacing arthroplast with THA	men and women aged 50 years or older undergoing MoM HRA or THA for advanced OA of hip	healthcare system perspective	30 years, one year cycle	5	QALYS	medical costs

Estimating the Societal Benefits of THA After Accounting for Work Status and Productivity A Markov Model Approach	Koenig (2016)	US	hip	THA compared with nonoperative treatment	patients with OA of the hip	societal perspective	lifetime and one-year cycle	3	QALYs	direct medical costs
The cost-effectiveness of surgical treatment of medial unicompartmental knee osteoarthritis in younger patients a computer model-based evaluation	Konopka (2015)	US	knee	high tibial osteotomy, unicompartmental knee arthroplasty, and TKA	fifty to sixty-year-old patients with medial unicompartmental OA with varus deformity	societal perspective	lifetime	3	QALYs	medical costs
The Cost-Effectiveness of Total Hip Arthroplasty in Patients 80 Years of Age and Older	Kunkel (2017)	US	hip	THA compared with non-opreative management	patients ≥80 years of age with ESOA of the hip	societal perspective	lifetime	3	QALYs	medical costs and costs of long-term assisted living
Cost effectiveness of COX 2 selective inhibitors and traditional NSAIDs alone or in combination with a proton pump inhibitor for people with osteoarthritis	Latimer (2009)	England and Wales	OA	COX 2 selective inhibitors (delecoxib and etoricoxib) vs traditional NASIDs (diclofenac, ibuprofen, and naproxen). No treatment, paracetamol, and the addition of PPI to each NSAID were also considered.	patients aged 55, and an older cohort of patients (age 65)	healthcare payer perspective that of the NHS in England and Wales	base case: 3 months, lifetime horizon was also adopted	3.5	QALY	treating side effects, drug costs. costs of outpatient appointments and GP consultations
Value of information in the osteoarthritis setting cost effectiveness of COX-2 selective	Latimer (2011)	UK	OA	NSAIDs+PPI vs COX-2s +PPIs	a modelled cohort of age 55 years and a modelled cohort of age 65 years (representing low-	NHS payer perspective	lifetime duration and 3 months cycle length	3.5	QALYs	OA and Aes treatments costs

inhibitors, traditional NSAIDs and proton pump inhibitors					and high-risk patients)					
Cost-effectiveness of unicompartmental knee arthroplasty, high tibial osteotomy, and KineSpring(R) Knee Implant System for unicompartmental osteoarthritis of the knee	Li (2013)	not mentioned	knee	UKA, HTO and knee implant system	patients with unicompartmental OA of the knee at the age of 55 years or younger	societal perspective	10 years	3 for effect	QALYs	surgical costs and complication costs
Economic evaluation of tramadol paracetamol combination tablets for osteoarthritis pain in the Netherlands	Liedgens (2005)	Netherlands	OA	tramadol/paracetamol tablets, NSAIDs alone, NSAIDs plus PPIs, NSAIDs plus H2RAs	patients with OA	health insurance system (healthcare system)	6 months	undiscounted	NA	only direct medical costs
Cost-effectiveness of total knee arthroplasty in the United States patient risk and hospital volume	Losina (2009)	US	knee	No TKA performed, TKA performed in a low-volume hospital, TKA performed in a medium-volume hospital, and TKA performed in a high-volume hospital	persons 65 years or older with end-stage knee OA	societal perspective	lifetime, 1 year cycle	3	QALY	TKA-related costs and costs of living with end-stage knee OA
Pharmacologic regimens for knee osteoarthritis	Losina (2014)	US	knee	Pharmacologic regimens for knee osteoarthritis prevention	NA		lifetime	3	QALY	NA

prevention can they be cost-effective										
Model-based evaluation of cost-effectiveness of nerve growth factor inhibitors in knee osteoarthritis impact of drug cost, toxicity, and means of administration	Losina (2016)	US	knee	standard of care adding vs. without nerve growth factor inhibitors (Tanezumab as an example)	NA	not mentioned	lifetime	3	reduction in pain severity, risk of TKR, QALYs,	medical costs
Disease-modifying drugs for knee osteoarthritis can they be cost-effective	Losing (2013)	US	knee	standard of care (consists of four regimens: conservative pain management, corticosterioid injection, primary TKR, and revision TKA) vs standard of care +disease-modified drug	considered cohorts with a mean age of 53.5 years	health system perspective	lifetime	3	QALY	medical costs (drug costs, office visit costs)
Cost-effectiveness of generic celecoxib in knee osteoarthritis for average-risk patients a model-based evaluation	Losing(2018)	US	knee	generic celecoxib with naproxen with or without PPIs	knee OA patients (mean age of 65) without major comorbidities	healthcare sector perspective	lifetime	3	QALYs	drug costs and OA-related medical costs
An economic model of long-term use of celecoxib in patients with osteoarthritis	Loyd (2007)	US	OA	celecoxib compared with nsNSAIDs (diclofenac and naproxen)	population of 60-year-old OA patients with average risks of upper gastrointestinal complications	societal perspective	lifetime	3	QALY	drug costs, inpatient hospital costs, lengths of stay
The cost effectiveness of rofecoxib and celecoxib in patients with	Maetzel (2003)	Canada	arthritis	COX-2 (rofecoxib, celecoxib) vs NSNSAIDs (naproxen, ibuprofen and diclofenac)	RA and OA patients with an average age of 58 years who do not require low-dose	third-party payer	5 years	5	QALY	GI-related hospitalizatio n cost, ambulatory care costs for

osteoarthritis or rheumatoid arthritis					aspirin. High risk patients with UGI history					general GI diagnostic investigation, costs for all physician billings, drug costs
Cost utility modeling of early vs late total knee replacement in osteoarthritis patients	Mari (2016)	Frence	knee	early vs. late TKR	the cohort composition was defined by age to mirroe the French population distribution of OA patients	healthcare payer perspective	30 years (lifetime)	4	QALYs	direct medical costs
Incremental cost-effectiveness analysis comparing rofecoxib with nonselective NSAIDs in osteoarthritis Ontario Ministry of Health perspective	Marshall (2001)	Canada	OA	rofecoxib with nonselective NSAIDs (a mixed of ibuprofen, diclofenac and nabumetone)	OA patients aged >65 years who did not respond to paracetamol (acetaminophen) therapy	Ministry of Health perspective	1 year	undiscounted	number of perforations, ulcers and bleeds (PUBs) averted	direct medical costs
Cost effectiveness analysis of hemiarthroplasty and total shoulder arthroplasty	Mather (2010)	US	glenohumeral OA	total shoulder arthroplasty and hemiarthroplasty	64-year-old patients	societal perspective	lifetime, one year cycle	3	QALYs	NA
Economic evaluation of access to musculoskeletal care the case of waiting for total knee arthroplasty	Mather (2014)	not mentioned	knee	TKA without delay vs. waiting period with/without non-operative treatment	60-year-old patients with knee OA	societal perspective	lifetime/monthly cycle	3	QALY	medical cost
A cost-minimisation analysis comparing	McKell (1994)	UK	OA	piroxicam vs. ibuprofen	patients with mild OA of the superficial joints	not mentioned	3 months	not mentioned	NA	health service costs

topical versus systemic NSAIDs in the treatment of mild osteoarthritis of the superficial joints										
Health economic comparisons of rofecoxib versus conventional nonsteroidal antiinflammatory drugs for osteoarthritis in the United Kingdom	Moore (2001)	UK	OA	convetional NSAID vs. rofecoxib	patients with OA	NHS perspective	one years	undiscounted	years of life saved	direct medical costs
Economic evaluation of etoricoxib versus non-selective NSAIDs in the treatment of osteoarthritis and rheumatoid arthritis patients in the UK	Moore (2004)	UK	arthritis	etoricoxib versus non-selective NSAID (NSNSAID alone, NSNSAID+PPI, NSNSAID +H2RA, NSNSAID+ misoprostol)	patients with RA or OA in the UK	NHS perspective	1 year	1.5 for effect	QALY	the key cost items included costs of treatment, GP consultations, outpatient
Can Robot-Assisted Unicompartmental Knee Arthroplasty Be Cost-Effective A Markov Decision Analysis	Moschetti (2016)	not mentioned	knee	robit-assisstant UKA vs. tradiional UKA	low-demand patient population with unicompartmental end-stage OA and an average age of 65 years	societal perspective	lifetime	3	QALYs	medical costs
Cost-effectiveness analysis of early versus late total hip replacement in Italy	Mota (2013)	Italy	hip	early vs. delayed THR	six cohorts defined by sex and age (50-59 years, 60-74 years, and 75 years and older)	NHS perspective	lifetime	3	QALY	medical costs

The cost-effectiveness of celecoxib versus non-steroidal anti-inflammatory drugs plus proton-pump inhibitors in the treatment of osteoarthritis in Saudi Arabia	Nasef (2015)	Saudi Arabia	OA	celecoxib versus ns-NSAIDs, with and without PPI	patients with OA aged ≥65 years	A patient perspective	6 months, 2 years, and 5 years	3	QALYs	drug acquisition, treatment of Aes and Physician visits.
Cost-effectiveness of unicondylar versus total knee arthroplasty a Markov model analysis	Peersman (2014)	Belgian	knee	unicompartmental knee arthroplasty vs TKA	patient age (≥ 75 years, 65-75 years, 55-65 years, and <55 years)	healthcare payer perspective	lifetime/ yearly cycle	3 for cost, 1.5 for effect	QALYs	only direct medical costs
Economic Evaluation of Rofecoxib Versus Nonselective Nonsteroidal Anti-Inflammatory Drugs for the Treatment of Osteoarthritis	Pellissier (2001)	US	OA	rofecoxib versus nonselective NSAIDs (the cost is weighhted average by market share)	not mentioned	third-party payer	1 year	3 for effect	PUB avoided; years of life saved	direct medical costs
Cemented, cementless, and hybrid prostheses for total hip replacement cost effectiveness analysis	Pennington (2013)	UK	hip	three commonly used prosthesis for THR	patients who enter the model at the time oof the primary THR	NHS perspective	lifetime	3.5	QALY	medical costs

Cost-Effectiveness of Five Commonly Used Prosthesis Brands for Total Knee Replacement in the UK A Study Using the NJR Dataset	Pennington (2016)	UK	knee	five commonly used prosthesis for total knee replacement	patients with average pre-operative characteristics	health care perspective	lifetime	3.5	QALYs	surgical costs
Cost effectiveness of total hip arthroplasty in osteoarthritis comparison of devices with differing bearing surfaces and modes of fixation	Pulikottil-Jacob (2015)	England and Wales	hip	five commonly used combination of components in THA, including type of fixtion and bearing of surface	not mentioned	NHS and personal social services perspective	ten years and lifetime	3.5	QALYs	medical cost
Has Metal-On-Metal Resurfacing Been a Cost-Effective Intervention for Health Care Providers-A Registry Based Study	Pulikottil-Jacob (2016)	UK	hip	metal-on-metal hip resurfacing procedure vs. commonly employed THR	not mentioned	NHS and personal and socital perspective	10 years and lifetime	3.5	QALYs	medical costs
Cost-Effectiveness Analysis of unicompartmental knee arthroplasty and high tibial osteotomy for treatment of medial compartmental osteoarthritis	Richard (2010)	US	knee	unicompartmental knee arthroplasty vs high tibial osteotomy	cohort of 40-year-old patients with unicompartmental knee OA	payer perspective	lifetime, one year cycle	3	QALYs	costs of procedures (includes infection)

Societal and Economic Effect of Meniscus Scaffold Procedures for Irreparable Meniscus Injuries	Rongen (2016)	not mentioned	irreparable meniscus injury	meniscus scaffold vs. standard partial meniscectomy	patients who had an irreparable injury to the medial meniscus with mean age of 39 years	societal perspective	lifetime	4 for costs, 1.5 for effects	QALYs	direct costs
Arthroscopic meniscectomy for degenerative meniscal tears reduces knee pain but is not cost-effective in a routine health care setting	Rongen (2018)	not mentioned	knee	arthroscopic meniscectomy vs. no surgery	subjects with, or at risk of for, symptomatic knee OA	societal perspective	9 years	4 for costs, 1.5 for effects	QALYs	health care consumption
Cost-effectiveness of timely versus delayed primary total hip replacement in Germany A social health insurance perspective	Ruben (2017)	Germany	hip	timely total primary hip replacement, delayed total hip replacement, and non-surgical therapy	functionally independent serious OA	statutory health insurer	lifetime/ one year cycle length	5	QALYs	
The direct and indirect costs to society of treatment for endstage knee osteoarthritis	Ruiz (2013)	US	knee	TKAvs. Nonsurgical treatment	the population forty years of age or older	societal perspective	lifetime	3	QALY	medical costs
Assessing the cost-effectiveness of COX-2 specific inhibitors for arthritis in the Veterans Health Administration	Schaefer (2005)	US	arthritis	COX-2(rofecoxib and celecoxib) with noeselective NSAIDs	arthritis patients considered at higher risk of developing clinically significant upper GI: patients of any	veterans health administration	one year	not mentioned	clinically significant upper GI event avoided, QALY gained	only direct medical costs

					age with previous medical history of perforation/ulcer/bleed, patients 65 years and older					
Cost-effectiveness of unicompartmental and total knee arthroplasty in elderly low-demand patients. A Markov decision analysis	Slover (2006)	US	arthritis (knee)	unicompartmental knee arthroplasty vs total knee arthroplasty	seventy-eight-year-old patients with unicompartmental arthritis	payer perspective	lifetime	3	QALY	NA
Impact of hospital volume on the economic value of computer navigation for total knee replacement	Slover (2008)	US	knee	computer-assisted surgery vs. TKA	65-year-old patients with end stage arthritis of the knee	not mentioned	lifetime	3	QALY	primary and revision costs
Medial compartment knee osteoarthritis age-stratified cost-effectiveness of total knee arthroplasty, unicompartmental knee arthroplasty, and high tibial osteotomy	Smith (2017)	US	knee	3 strategies: opioid-sparing (OS), tramadol (T), and tramadol followed by oxycodone(T+O)	knee OA patients with a mean age 60 years without major comorbidities	societal perspective	lifetime	3	QALYs	OA-related and non-OA-related medical costs

Cost-Effectiveness of Tramadol and Oxycodone in the Treatment of Knee Osteoarthritis	Smith (2017)	UK	knee	total knee arthroplasty (TKA), unicompartmental knee arthroplasty (UKA), and high tibial osteotomy (HTO)	cohort of patients 40, 50, 60, and 70 years at the time of primary surgical intervention	perspective of the health services sector	ten post-operative years	3.5	QALYs	total costs incorporated length of hospital stay, implant costs, and cement mix differences
50 Cost-effectiveness analysis of total ankle arthroplasty	SooHoo (2004)	US	ankle arthritis	total ankle arthroplasty vs ankle fusion	target patient population with end-stage ankle osteoarthritis at the age of fifty-five years.	societal perspective	lifetime (25 years)	3	QALY	direct treatment costs of various procedures
Cost-effectiveness analysis of unicompartmental knee arthroplasty as an alternative to total knee arthroplasty for unicompartmental osteoarthritis	Soohoo (2006)	US	knee	unicompartmental knee arthroplasty vs total knee arthroplasty	a target population seeking treatment for unicompartmental arthritis at the age of sixty five years.	societal perspective	lifetime (the time horizon of this analysis encompasses the remaining eighteen years of life expectancy for this target population)	3	QALY	direct lifetime treatment costs
The cost-effectiveness of cyclooxygenase-2 selective inhibitors in the management of chronic arthritis	Spiegel (2003)	US	arthritis	coxibs (rofecoxib and celecoxib) with generic nonselective NSAID (naproxen)	60-year-old patients with OA or RA (high risk subgroup analysis)	third-party payer	lifetime	3	QALY	consider only direct health care costs

Minimizing complications from nonsteroidal antiinflammatory drugs cost-effectiveness of competing strategies in varying risk groups	Spiegel (2005)	US	arthritis	3 strategies: NSAID alone, NSAID with PPI, coxib alone (with and without aspirin were compared)	60-year-old patients with chronic arthritis	third-party payer	1 year	not mentioned	ulcer complication avoided, QALY	consider only direct health care costs
Use of the ACCES model to predict the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis in Norway	Svarvar (2000)	Norway	arthritis	celecoxib with nonselective NSAIDs with or without gastroprotective agents	NA	societal perspective	1 year	5	event averted; life-year gained	drug costs, resource utilization
Economic evaluation of nimesulide versus diclofenac in the treatment of osteoarthritis in France, Italy and Spain	Tarricone (2001)	France, Italy and Spain	OA	nimesulide vs. diclofenac	not mentioned	NHS perspective (healthcare system perspective)	2 weeks	undiscounted	NA	direct healthcare costs
Modelling therapeutic strategies in the treatment of osteoarthritis an economic evaluation of meloxicam versus diclofenac and piroxicam	Tavakoli (2003)	UK	OA	meloxicam(cox-2) vs diclofenac and piroxicam (nonselective NSAIDs)	not mentioned	NHS perspective	4 weeks	undiscounted	NA	only consider direct cost
Knee Joint Distraction Compared to Total	van der Woude (2016)	Dutch	knee	Knee Joint Distraction Compared to Total Knee Arthroplasty	the target population consisted of	health care perspective	20 years	4 for costs, 1.5 for effects	QALYs, costs perTKA	direct medical cost

Knee Arthroplasty for Treatment of End Stage Osteoarthritis Simulating Long-Term Outcomes and Cost-Effectiveness					patients with advanced, generalized knee OA indicated for TKA				saved, costs per revision operation saved, costs per 2nd revision/BSC saved	
Cost analysis of flavocoxid compared to naproxen for management of mild to moderate OA	Walton (2010)	not mentioned	OA	flavocoxid vs. naproxen	patients with mild to moderate OA who are over age 65	Medicare payer perspective	one year	not mentioned	NA	drug costs and costs of events
Modeling the economic and health consequences of managing chronic osteoarthritis pain with opioids in Germany comparison of extended-release oxycodone and OROS hydromorphone	Ward (2007)	Germany	OA	The Osmotic Controlled-Release Oral delivery System hydromorphone with equianalgesic dose of extended-release oxycodone	a cohort of 1000 individual patients with severe pain due to OA	health insurance system	one year	not mentioned	QALY	direct medical costs
Cost-utility analysis of duloxetine in osteoarthritis a US private payer perspective	Wielage (2013)	US	OA	duloxetine vs. post first-line (acetaminophen) oral treatments (celecoxib, naproxen, a combination of oxycodone/acetaminophen, oxycodone extended release, and tramadol and tapentadol)	cohort enters the model at 55 years of age with no previous GI or CV events. The study also performed two subgroup analysis: a normal-risk population 65 years	private payer perspective	lifetime	3	QALYs	direct medical costs

					and older, and a high-risk population that had earlier experienced CV and GI events.					
Cost effectiveness of duloxetine for osteoarthritis a Quebec societal perspective	Wielage (2014)	Canada	OA	duloxetine, NSAID (celecoxib, diclofenac, Naproxen), opioid (Hydromorphone, Oxycodone)	the modelled population consisted of patients with OA with chronic moderate to severe pain uncontrolled by acetaminophen. In the base case, the cohort began at age 55 years without a history of GI or CV events.	societal perspective	lifetime, using 3-month cycles for the first 3 years and annual cycles thereafter	5	QALYs	drug costs, costs of consultation, and costs for Aes
Development and validation of a new population-based simulation model of osteoarthritis in New Zealand	Wilson (2018)	New Zealand	knee	no specific treatment	NA		lifetime/ annual cycle	undiscounted	QALYs	
Cost-effectiveness of adjunct non-pharmacological interventions for osteoarthritis of the knee	Woods (2017)	UK	knee	adjunct non-pharmacological (appliances, electrotherapy, manual therapy, static magnets, heat treatment, usual care)	a general cohort of patients with knee OA (age >55 years)	NHS perspective and Personal Social Services	8 weeks	undiscounted	QALYs	Intervention costs comprised equipment costs and staff time
Cost-effectiveness of treatment strategies for osteoarthritis of the knee in Taiwan	Yen (2004)	Taiwan	knee	naproxen, celecoxib, and hyaluronan	60-year-old woman with symptomatic and radiological knee OA	societal perspective	26 weeks	undiscounted	QALYs	outpatient treatments, inpatient treatments for serious GI complications

Arthritis treatment in Hong Kong--cost analysis of celecoxib versus conventional NSAIDS, with or without gastroprotective agents	You (2002)	HongKong	arthritis	celecoxib vs. conventional NSAID	not mentioned	public health organisation perspective	6 months	not mentioned	NA	direct medical costs (costs of drugs and GI events)
An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada	Zabinski (2001)	Canada	arthritis	6 treatments: celecoxib,NSAID alone, NSAID with PPI or H2RA or misoprostol or diclofenac	arthritis patients aged ≥65 years	Ministry of Health perspective	6 months	not mentioned	NA	direct medical costs

Supplement 3.4 The extracted information from included studies (continuous table)

Title of the studies	Indirect cost	Currency	Modelling software	Model type	Health state	Model events	Sensitivity analysis	Reporting score	Quality group
Nabumetone compared with Ibuprofen and a weighted NSAID combination an economic evaluation	not mentioned	not mentioned	not mentioned	decision tree model	minor S/E, no S/E continue NSAID, and major S/E stop NSAID	minor and major GI side effects	sensitivity analyses were performed by modifying key probability variables over 95% CI	73%	2
Costs and effects of various analgesic treatments for patients with rheumatoid arthritis	production losses are disregard	2004 Euros	not mentioned	decision tree model	no Gi side effects, a certain GI side effect, death due to a serious GI event	GI Aes	PSA and univariate analyses	74%	2

and osteoarthritis in the Netherlands									
Economic benefit to society at large of total knee arthroplasty in younger patients a Markov analysis	Indirect costs (productivity losses) related to lost income due to work disability, absenteeism, decreased work hours, job change, or unemployment	2012 US dollars	Treeage	Markov model (patient level)	five health state: working with no operative treatment, working after primary TKA, not working after primary TKA, working after revision TKA, not working after revision TKA, and two absorbing states: a repeat revision TKA or death due to any cause.	revision surgery (two times), perioperative complications	one-way, two-way sensitivity analysis	90%	3
Cost-utility of celecoxib use in different treatment strategies for osteoarthritis and rheumatoid arthritis from the Quebec healthcare system perspective	not mentioned	2005 Canadian dollar	not mentioned	Markov decision-analytic model, cost-utility outcomes simulation model for OA and RA patients (COSMO)		Gi discomfort, complicated and uncomplicated ulcers, CV event, discontinuation due to Aes	multiple one-way sensitivity analyses	83%	3
Exercise, Manual Therapy, and Booster Sessions in Knee Osteoarthritis Cost-Effectiveness Analysis From a Multicenter Randomized Controlled Trial	direct non-medical cost (transportation to and from medical visits were included), not include lost productivity	2011 US dollars	Treeage	Markov model	poor/worsening function, good/improving function, good function after TKA, poor function after TKA, good function after	arthroscopy, TKA, the WOMAC index changed beyond the minimum clinically important difference	one-way, PSA	91%	3

					scope, poor function after scope, death				
The cost effectiveness of celecoxib vs diclofenac in the treatment of osteoarthritis in the UK an update to the NICE model using data from the CONDOR	not mentioned	2010/11 UK	not mentioned	NICE model (Markov model)	NA	GI and CV events	sensitivity analysis varying the treatment period, PSA	79%	3
A cost effectiveness analysis of celecoxib compared with diclofenac in the treatment of pain in osteoarthritis OA within the Swedish health system	the cost of unemployment because of illness were not considered	2012 US dollars	not mentioned	The NICE (Markov) model	14 possible health states consisting of no complications, GI symptoms/dyspepsia, symptomatic ulcer, post-symptomatic ulcer, complicated GI bleed, post-complicated GI bleed, MI, post-MI, stroke, post-stroke, HF, post-HF, post-treatment with no complications, and death	GI and CV events	one-way PSA	70%	2
The cost-effectiveness of meniscal repair versus partial meniscectomy A model-based projection for the United States	not included	2014 USD	not mentioned	Markov model	no OA, severe knee OA, post TKR, post-revision TKR	development of OA, TKR, revision TKR (two times)	one way and two-way analysis	91%	3
The cost-effectiveness of dual mobility implants for primary total hip arthroplasty A computer-based cost-utility model	lost-wage	2013 US dollar	Treeage	Markov model	primary THA, revision, repeat revision, chronic failed THA, dead	dislocation, revision, repeat revision (two revision)	one way and two-way deterministic analyses, PSA	87%	3

Cost-effectiveness Study of Celecoxib for Osteoarthritis in China	not included	US dollar	not mentioned	Markov model	no AES, Aes, death	GI, CV events	one way, PSA	65%	2
Cost-effectiveness of unicompartmental compared with total knee replacement a population based study using data from the National Joint Registry for England and Wales	not mentioned	2014 pound	not mentioned	Markov model	UKR/TKR, revision, re-revision (only two revision)	revision surgery (two revision)	PSA	96%	3
Efficiency of naproxenesomeprazole in association for osteoarthrosis treatment in Spain	not included	€, 2012	Microsoft excel	Markov model	eight health states: "without incident" state, GI-dyspepsia, symptomatic or complicated ulcer or ulcer, CV-myocardial infraction, stoke, or congestive heart failure, and death	GI and CV events, discontinuation, adherence	deterministic sensitivity analyses, PSA	78%	3
Cost-effectiveness Analysis of Viscosupplementation versus Conventional Supportive Therapy for Knee Osteoarthritis in Colombia	not mentioned	US dollar	microsoft Excel and Visual Basic macros	discrete-event simulation model		symptom improvement, no change in symptoms, worsening of symptoms, TKA	first-order Monte-Carlo simulation, second-order Monte-Carlo simulation, PSA	70%	2
Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor, in Switzerland	not included	Swiss francs	Treeage	decision tree, COMET	GI AE: no GI AE, GI discomfort, symptomatic ulcer, anaemia with occult bleeding, serious GI complication	GI AE: moderate to severe GI discomfort, symptomatic ulcer, anaemia with occult bleeding, and serious GI events requiring hospitalisation (with or without death).	multiple one-way sensitivity analyses, Monte Carlo simulation	73%	2

A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip	not mentioned	1991 dollars	Microsoft excel	stochastic decision tree model	health states were classified into four states by ACR functional status classification. Success surgery resulting in functional class 1, fair outcome resulting in functional class Ⅱ, short-term failure resulting in revision surgery within a year, and death due to perioperative mortality	aseptic failure, joint injection, injection revision, aseptic revision (three times revision)	worst-case analyses	79%	3
Cost effectiveness of patellofemoral versus total knee arthroplasty in younger patients	not mentioned	2015 United States dollars	Treeage	Markov model	primary PFA or TKA, post-operative state, implant failure, conversion to TKA, post-operative state, implant failure, revision TKA, post-operative state, implant failure	PFA, TKA, revision TKA (two times), medical complications, surgical complications	one-way deterministic and PSA	88%	3
Cost-effectiveness analysis for joint pain treatment in patients with osteoarthritis treated at the Instituto Mexicano del Seguro Social (IMSS)	not mentioned	2008 Sep Mexican pesos	Treeage	decision tree model	pain controlled, pain not controlled, without AE, with AE	peptic ulceration, GI bleeding, renal complications, CV event, and other minor AE	one-way and PSA	86%	3
Modeling the cost-effectiveness for cement-less and hybrid prosthesis	not included	not mentioned	Treeage	Markov model	THR implanted, revision A, successful revision A, revision B, successful	revision (two times), perioperative	sensitivity analysis	78%	3

in total hip replacement in emilia romagna, Italy					revision B, revision C, successful revision C, death after intervention, death any cause	mortality, death of any cause			
Meniscus Root Repair vs Meniscectomy or Nonoperative Management to Prevent Knee Osteoarthritis After Medial Meniscus Root Tears Clinical and Economic Effectiveness	not mentioned	2017US dollar	not mentioned	Markov model	Non OA state, OA, TKA, first TKA revision, second TKA revision, death	progression to OA, TKA, first TKA revision, second TKA revision, death (two revision)	sensitivity analyses varying all input parameters were conducted	78%	3
Effect of age on cost-effectiveness of unicompartmental knee arthroplasty compared with total knee arthroplasty in the U.S	not mentioned	2012 US dollars	Treeage	Markov model (cohort level)	Model health states for both procedures were full-benefit, post-surgery, limited-benefit, post-surgery, failed primary surgery, revision total knee arthroplasty, full-benefit post-revision, limited-benefit post-revision, failed revision, and death (WOMAC SCORE)	implant failure, revision surgery (more than one), surgical complication, rehabilitation	one-way sensitivity, PSA	88%	3
The Swedish ACCES model predicting the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis			not mentioned	decision tree model, Arthritis Cost Consequence Evaluation System (ACCES) pharmacoeconomic model				55%	2
Cost-effectiveness analysis of intra-articular injections of a high molecular weight bioengineered hyaluronic	not mentioned	2012 USD	Treeage	decision tree model	BioHA, continuation of baseline treatment, responder, non-responder		one way sensitivity, PSÁ	74%	2

acid for the treatment of osteoarthritis knee pain									
Cost-utility of metal-on-metal hip resurfacing compared to conventional total hip replacement in young active patients with osteoarthritis	societal costs were not included	2011 Canadian	Treeage	probabilistic Markov model	post primary surgical procedure, post-THA conversion, post-1st THA revision, post-2nd revision	surgical implication, revision surgery (two times)	PSA, subgroup analysis	96%	3
Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis discrete-event simulation model	unrelated healthcare costs due to extended life years of OA patients were included	AUD 2003	Microsoft Excel	discrete-event simulation model	NA	decision for JR, implant failure, revision implant	Monte Carlo simulation (PSA)	92%	3
An economic evaluation of meloxicam 7.5 mg versus diclofenac 100 mg retard in the treatment of osteoarthritis in the UK A decision analysis model based on gastrointestinal complications	not considered		Treeage	decision tree model		1.no GI events, 2.GI event not requiring treatment, 3. minor GI, ambulatory treatment required 4. ulcer, ambulatory treatment required, 5. ulcer requiring hospitalisation, 6. haemorrhage requiring hospitalisation, 7. perforation requiring hospitalisation	sensitivity analyses were performed by modifying key probability variables over 95% CI	76%	3

Economic evaluation of meloxicam (7.5 mg) versus sustained release diclofenac (100 mg) treatment for osteoarthritis A cross-national assessment for France, Italy and the UK	not mentioned	1995 currency	Treeage	decision tree model	1.no GI events, 2.GI event not requiring treatment, 3. minor GI, ambulatory treatment required 4. ulcer, ambulatory treatment required, 5.ulcer requiring hospitalisation, 6.haemorrhage requiring hospitalisation, 7.perforation requiring hospitalisation	GI event	sensitivity analyses were performed by modifying key probability variables over 95% CI	77%	3
Cost-effectiveness of antibiotic-impregnated bone cement used in primary total hip arthroplasty	surgeons' fee, costs for a rehabilitation stay, and lost wages were not included.	2002 US dollar	Treeage	Markov model	primary THA with or without antibiotic-impregnated bone cement, septic revision, aseptic revision, death	septic revision, aseptic revision (once time)	sensitivity analysis	78%	3
The cost-effectiveness of acetaminophen, NSAIDs, and selective COX-2 inhibitors in the treatment of symptomatic knee osteoarthritis	not included	2000 USD	Treeage	decision tree Model		clinical confirmed upper GI ulcer, PUB. Clinical suspicion of PUB	one-way, two-way and threshold analysis, PSA	87%	3
A Fresh Perspective on a Familiar Problem Examining Disparities in Knee Osteoarthritis Using a Markov Model	lower labor productivity derived from the Medical Expenditure Panel Survey		Microsoft Excel	Markov model		no knee OA, knee OA-K-L Grade 1&2, K-L Grade 3, K-L Grade 4, death		74%	2

Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities	not mentioned	2013 US dollar	not mentioned	Osteoarthritis Policy (OAPol) Model		lack of analgesic efficacy, major toxicity (CV and GI), voluntary discontinuation	one- and two-way, PSA	87%	3
Do the potential benefits of metal-on-metal hip resurfacing justify the increased cost and risk of complications	not mentioned	not mentioned	Treeage	Markov decision model	initial hip resurfacing, post-hip resurfacing, post-conversion to THA, post-major total revision THA, post-major partial revision THA, post-minor revision THA, death. Initial primary THA,	initial failure, subsequent failure requiring revision (two times)	one way, two-way, PSA	71%	2
Estimating the Societal Benefits of THA After Accounting for Work Status and Productivity A Markov Model Approach	employment status and worker earnings (based on methods used by Dall et al., used national health interview survey data to generate regression coefficients that described the relationship between physical functioning and economic outcomes)	2011US dollar	Treeage	Markov model (patient level)	non surgery, end-stage hip OA, more severe hip OA, initial post-THA, successful post-THA, post-first THA early revision, post-first THA late revision	primary THA, revision THA (two times), surgical complication (infection)	threshold analyses, Monte Carlo analysis (PSA)	75%	2
The cost-effectiveness of surgical treatment of medial unicompartmental knee osteoarthritis in	not mentioned	2012 U.S. dollars	Treeage	Markov model	primary HTO, optimal HTO, suboptimal HTO, UKA, optimal UKA, suboptimal UKA, TKA,	medical and surgical complications,	one-way and two-way deterministic sensitivity	70%	2

younger patients a computer model-based evaluation					optimal TKA, suboptimal TKA, revision TKA, optimal revision TKA, suboptimal revision TKA	revision surgery (two times)	analyses, PSA, EVPPI		
The Cost-Effectiveness of Total Hip Arthroplasty in Patients 80 Years of Age and Older	not included	2016 US dollars	Treeage	Markov model	end-stage OA, independent living, dependent living, THA, RTHA, FRTHA	once time revision	one-, two-, three-way sensitivity analysis	74%	2
Cost effectiveness of COX 2 selective inhibitors and traditional NSAIDs alone or in combination with a proton pump inhibitor for people with osteoarthritis	not mentioned	£2007-8	not mentioned	NICE model (Markov model)	The health states represent the most frequent and severe adverse events: dyspepsia; symptomatic ulcer; complicated gastrointestinal perforation, ulcer, or bleed; myocardial infarction; stroke; and heart failure. In addition, a patient can experience no adverse event, or death	GI and CV Aes	deterministic sensitivity analyses; PSA	88%	3
Value of information in the osteoarthritis setting cost effectiveness of COX-2 selective inhibitors, traditional NSAIDs and proton pump inhibitors	societal costs were not included	£2007-8 values	not mentioned	NICE model (Markov model)	health states included in the model represent the key AE associated with NSAIDs and COX-2 inhibitors, as well as a no complication and a post-treatment health state. AE included GI symptoms, symptomatic ulcer, serious GI bleed, stroke, MI, heart failure and death.	GI and CV events	PSA, EVPI, EVPPI	75%	2

Cost-effectiveness of unicompartmental knee arthroplasty, high tibial osteotomy, and KineSpring(R) Knee Implant System for unicompartmental osteoarthritis of the knee	not mentioned	2012 US dollars	Treeage	dicision tree model	baseline pre-treatment, 1 year post successful treatment, 2 years post successful treatment, baseline pre-conversion surgery, post-conversion surgery, complication	surgical complication, implant removal, conversion	not mentioned	61%	2
Economic evaluation of tramadol paracetamol combination tablets for osteoarthritis pain in the Netherlands	not included	2005 Euros	Treeage	decision tree model	Patients modelled as receiving tramadol/paracetamol could either have no adverse events or an adverse event. For patients modelled as receiving one of the comparator treatments, there were six possible final outcomes: no GI toxicity, GI distress, serious GI complications that they survived, serious GI complications that lead to death, symptomatic ulcer, anaemia with occult bleeding.	GI Aes, renal AE (in subsequent scenario)	univariate sensitivity analyses	82%	3
Cost-effectiveness of total knee arthroplasty in the United States patient risk and hospital volume	not mentioned	2006 US dollars	Treeage	Markov model (patient level)	10 health states: end-stage knee OA (pre-TKA); TKA; full-benefit post-TKA (successful TKA);limited-benefit post-TKA (unsuccessful TKA); failed TKA; revision TKA; full-benefit post revision TKA (successful	medical complication, revision surgery (multiple times)	used deterministic 1-way and 2-way sensitivity analysis, PSA	83%	3

					revision); limited-benefit post revision TKA (unsuccessful revision);failed revision TKA; death. (WOMAC SCORE)				
Pharmacologic regimens for knee osteoarthritis prevention can they be cost-effective		in 2012 USD	not mentioned	Osteoarthritis Policy (OAPol) Model	no OA, NO OA DMOADs, OA DMOADs, OA nonsurgical regimens, TKR/ post TKR	regimen minor and major toxicity, discontinuation		79%	3
Model-based evaluation of cost-effectiveness of nerve growth factor inhibitors in knee osteoarthritis impact of drug cost, toxicity, and means of administration	not mentioned	2014 USD	not mentioned	Osteoarthritis Policy (OAPol) Model	primary TKA with less efficacious, revision TKA with less efficacious, primary TKA, revision TKA	Tanezumab toxicity (discontinuation, complication)	two-way sensitivity analysis	79%	3
Disease-modifying drugs for knee osteoarthritis can they be cost-effective	not to model indirect costs	2010USD	not mentioned	OAPOL model		Each year, subjects may develop a comorbid condition, increase in BMI, progression in OA severity, and/or die. Progression of OA is defined as an increase by one K-L radiographic grade and is dependent on obesity status and sex.	two-way	88%	3
Cost-effectiveness of generic celecoxib in knee osteoarthritis for average-	not mentioned	2015 US dollar	not mentioned	OAPoL Model	NSAIDs with or without PPIs, corticosteroid injections, primary	major toxicity (CV and GI events), discontinuation due to lack of efficacy,	one-way and PSA	92%	3

risk patients a model-based evaluation					TKA, revision TKA, death.	voluntary discontinuation due to another reason or death			
An economic model of long-term use of celecoxib in patients with osteoarthritis	exclude indirect costs	drug prices as of February 2006. we employ only the estimated generic price in our model after mid 2013.	Microsoft excel	decision tree model	no UGI symptoms, NUD, symptomatic peptic ulcer, POB	NUD, symptomatic peptic ulcer, POB, CV events were tested in sensitivity analysis	one-way, univariate analysis	83%	3
The cost effectiveness of rofecoxib and celecoxib in patients with osteoarthritis or rheumatoid arthritis	not mentioned	1999 Canadian dollars	not mentioned	Markov model	arthritis without GI events, dyspepsia, symptomatic ulcer, complicated UGI event with medical management, complicated UGI event with surgical management, nonfatal MI, life post-MI	GI and CV events	single variable sensitivity analysis	88%	3
Cost utility modeling of early vs late total knee replacement in osteoarthritis patients	not included	EUROS	Treeage	Markov model	non-pharmacological treatment option, both non-pharmacological and pharmacological treatment option, pharmacological treatment option, surgical treatment, death	GI, CV and renal AE associated to NSAIDs and opioids, revision surgery	sensitivity analysis	91%	3
Incremental cost-effectiveness analysis comparing rofecoxib with nonselective NSAIDs in osteoarthritis Ontario Ministry of Health perspective	direct nonmedical and indirect costs were not addressed	1999 Canadian dollars	not mentioned	decision tree model	GI symptoms, No GI symptoms, major GI symptoms, minor GI symptoms, confirmed PUBs, investigated PUBs, hospitalisation, outpatient, surgery for	Gastrointestinal symptoms were defined as 'minor' or 'major'. Major GI symptoms were subclassified as an 'investigated PUB'	one-way and two-way sensitivity analysis	95%	3

					confirmed PUBs, no surgery.	or 'confirmed PUB'.			
Cost effectiveness analysis of hemiarthroplasty and total shoulder arthroplasty	not mentioned	2008 US dollar	Treeage	Markov model (cohort level)	primary procedure, initial post-procedure, well post-procedure, revision TSA, death	failure procedure, revision procedure (once time)	one-way, multivariate sensitivity analyses, Monte Cristo microsimulation	87%	3
Economic evaluation of access to musculoskeletal care the case of waiting for total knee arthroplasty	included indirect costs	not mentioned	Treeage	Markov decision model	end-stage knee OA, end-stage knee OA with treatment bridge, primary TKA, revision TKA, recovery from revision TKA, recovery from early TKA complication	two revision TKA, perioperative complication	one, two, three-way sensitivity analysis	78%	3
A cost-minimisation analysis comparing topical versus systemic NSAIDs in the treatment of mild osteoarthritis of the superficial joints	not mentioned	£1991-92 price	not mentioned	decision tree model	no ulcer, out-patient care, hospital in-patient medical care, hospital in-patient surgical care	GI events, silent ulceration was not considered	sensitivity analysis	50%	1
Health economic comparisons of rofecoxib versus conventional nonsteroidal antiinflammatory drugs for osteoarthritis in the United Kingdom	indirect cost or disability costs were not included	1999UK	not mentioned	decision tree model	no GI AE, serious GI AE, minor GI AE, hospitalisation for PUB, suspected PUB	serious and minor GI AE	sensitivity analysis	86%	3
Economic evaluation of etoricoxib versus non-selective NSAIDs in the treatment of osteoarthritis and rheumatoid arthritis patients in the UK	No indirect or disability costs were included in the model	2002 pounds	not mentioned	decision tree model	treated arthritis, major GI surgery, inpatient treatment for major GI problem, outpatient treatment for major GI problem, inpatient investigation for suspected PUB,	major GI problem, minor GI problem	one-way, PSA, analyses by risk group	88%	3

					outpatient investigation for suspected major GI event, minor GI problem requiring treatment, minor GI problem not requiring treatment, death				
Can Robot-Assisted Unicompartmental Knee Arthroplasty Be Cost-Effective A Markov Decision Analysis	not mentioned	2012 US dollars	Treeage	Markov model	well postoperative, revision, death	revision (once time)	one way and two way	74%	2
Cost-effectiveness analysis of early versus late total hip replacement in Italy	not mentioned	2010euros	not mentioned	Markov model	severe OA-functionally independent, functionally dependent, annual successful primary THR, preoperative revision hip replacement, successful RHR, preoperative states and revision states, annual successful delayed primary THR (ACR CLASS)	perioperative mortality, postoperative short-term pulmonary embolism and infection. Revision surgery (two times)	best-worst scenario, PSA	83%	3
The cost-effectiveness of celecoxib versus non-steroidal anti-inflammatory drugs plus proton-pump inhibitors in the treatment of osteoarthritis in Saudi Arabia	not mentioned	2013 USD	not mentioned	NICE model (Markov model)	adverse event, no adverse event	GI AE: dyspepsia, symptomatic ulcer, major bleeding, CV AE: stroke-post stroke/death, myocardial infarction-post MI/death, heart failure-post HF/death	PSA	91%	3

Cost-effectiveness of unicondylar versus total knee arthroplasty a Markov model analysis	disease-related impact on productivity were excluded.	2014 Euros	Treeage	Markov model	the model contained states for patients age (≥75 years, 65–75 years, 55–65 years, and <55 years) and outcome ('revision', 're-revision', and 'death')	two revision procedure	threshold analysis, deterministic and PSA	83%	3
Economic Evaluation of Rofecoxib Versus Nonselective Nonsteroidal Anti-Inflammatory Drugs for the Treatment of Osteoarthritis	no included	1998 US dollar	not mentioned	decision tree model	no GI problem, hospitalization given a PUB, inpatient investigation of suspected PUB, surgery given hospitalization.	GI problems (classified as serious (PUB-related) or minor (nuisance symptom-related))	extensive sensitivity analyses	86%	3
Cemented, cementless, and hybrid prostheses for total hip replacement cost effectiveness analysis	not mentioned	British pounds 2011-12 prices	Microsoft excel	Markov model	primary THR, one stage revision, two stage revision, revised THR, dead	revision (two times)	sensitivity analysis	96%	3
Cost-Effectiveness of Five Commonly Used Prosthesis Brands for Total Knee Replacement in the UK A Study Using the NJR Dataset	not mentioned	British pounds 2011-12 prices	Microsoft excel	probabilistic Markov model	primary TKR, revision surgery, dead	revision (not mentioned times)	sensitivity analysis	88%	3
Cost effectiveness of total hip arthroplasty in osteoarthritis comparison of devices with differing bearing surfaces and modes of fixation	not included	2012 British pounds	not mentioned	Markov model	successful primary THA, revision surgery, successful revision surgery, death	revision (more than once time)	sensitivity analysis	74%	2

Has Metal-On-Metal Resurfacing Been a Cost-Effective Intervention for Health Care Providers-A Registry Based Study	not mentioned	2014 British pound	not mentioned	semi-Markov model	successful primary THR or RS surgery, revision THR surgery, successful revision THR, dead	revision surgery (more than once)	PSA, scenario analysis	83%	3
Cost-Effectiveness Analysis of unicompartmental knee arthroplasty and high tibial osteotomy for treatment of medial compartmental osteoarthritis	not mentioned	2008 US dollar	Treeage	Markov model (cohort level)	initial post procedure state, successful HTO, successful UKA, TKA, revision TKA, death	surgical complication included non-union, hardware removal, revision surgery (two times), infection	one-way and multivariate sensitivity analyses, Monte Carlo microsimulation was conducted	83%	3
Societal and Economic Effect of Meniscus Scaffold Procedures for Irreparable Meniscus Injuries	indirect costs (absenteeism)	euros	Treeage	Markov model	4 acute states (partial meniscectomy, meniscus scaffold, TKA, and revision TKA and 9 chronic states (irreparable meniscus injury, radiographic knee OA, symptomatic knee OA, full benefit post TKA, limited benefit post TKA, failed TKA, full benefit revision TKA, limited benefit revision, failed revision).	revision surgery (one time)	probabilistic Monte Carlo simulation, deterministic 1-way probabilistic	83%	3
Arthroscopic meniscectomy for degenerative meniscal tears reduces knee pain but is not cost-effective in a routine health care setting	productivity loss	2015 Euros	Treeage	Markov model	knee OA, arthroscopic meniscectomy, knee arthroplasty surgery, failed surgery, revision surgery, death.	revision surgery	sensitivity analysis, deterministic sensitivity analysis	83%	3
Cost-effectiveness of timely versus delayed primary total hip	indirect costs and nursing care	Euros at 2013	not mentioned	Markov model	functionally independent, functionally dependent,	surgical complication (pulmonary	deterministic sensitivity	92%	3

replacement in Germany A social health insurance perspective	costs were not included				primary THR, success primary THR, revision THR, death (ACR CLASS)	embolous, infection, bleeding, dislocation, revision), revision (once time)	analyses; two-way,		
The direct and indirect costs to society of treatment for end stage knee osteoarthritis	included indirect costs	2009 dollar	not mentioned	Markov model	end-stage OA of the knee, current primary TKA and rehabilitation, full benefit after primary or revision TKA, limited benefit after primary or revision TKA, current revision TKA and rehabilitation, failed primary TKA, failed revision TKA, death. (WOMAC SCORE)	revision (one time), surgical complication	one way	79%	3
Assessing the cost-effectiveness of COX-2 specific inhibitors for arthritis in the Veterans Health Administration	not considered	2001 US dollar	not mentioned	decision tree model	No AE, AE	GI and CV events, renal toxicity	univariate and multivariate sensitivity analyses, threshold analyses, PSA	83%	3
Cost-effectiveness of unicompartmental and total knee arthroplasty in elderly low-demand patients. A Markov decision analysis	not mentioned	2005 US dollars	Treeage	Markov decision model	primary UKA, primary TKA, peri-op death, revision, well with UKA, well with revision, well with TKA, death	infection, single revision, death	one-way	83%	3

Impact of hospital volume on the economic value of computer navigation for total knee replacement	not mentioned	2007 US dollar	Treeage	Markov model	well surgery, revision surgery, dead	revision surgery (once time)	two way	74%	2
Medial compartment knee osteoarthritis age-stratified cost-effectiveness of total knee arthroplasty, unicompartmental knee arthroplasty, and high tibial osteotomy	not mentioned	2014 USD	not mentioned	Osteoarthritis Policy (OAPol) Model	opioid-sparing strategy: conservative therapy, primary TKA, revision TKA; opioid-based strategy: conservative therapy, tramadol, oxycodone, primary TKA, revision TKA	major toxicity (CV events and fractures), lack of efficacy or voluntary discontinuation	one-way, multiple way, and PSA	70%	2
Cost-Effectiveness of Tramadol and Oxycodone in the Treatment of Knee Osteoarthritis	not mentioned	2013 £	not mentioned	Markov model (model structure was based on NICE guidelines)	well with primary surgery, revision TKA, well with revision TKA	revision TKA (two times)	discrete and PSA	79%	3
50 Cost-effectiveness analysis of total ankle arthroplasty	indirect costs such as lost productivity were not included	1998 US dollar	Treeage	decision tree model	ankle fusion, ankle replacement, revision ankle replacement, hindfoot arthritis, nonunion, surgery and postoperative recovery, below-the-knee amputation, revision fusion after replacement	short-term and long-term surgical complication, revision surgery	univariate and multivariate sensitivity analyses	82%	3

Cost-effectiveness analysis of unicompartmental knee arthroplasty as an alternative to total knee arthroplasty for unicompartmental osteoarthritis	not mentioned	1998 United States dollars	Treeage	decision tree model	primary TKA, UKA, treatment of infection, surgery and postoperative recovery, death, revision TKA, resection knee arthroplasty	successful TKA, successful UKA, complication (infection and death), revision surgery (once time)	several key variables were selected for sensitivity analysis, multivariate sensitivity analysis	78%	3
The cost-effectiveness of cyclooxygenase-2 selective inhibitors in the management of chronic arthritis	not considered	2002 US dollars (not sure, costs obtained from the 2002 American Medical Association)	Treeage	decision tree model	no GI complication, severe dyspepsia, moderate dyspepsia, ulcer haemorrhage, complicated ulcer requiring surgery.	upper GI dyspeptic symptoms, ulcer complications, CV events (considered in sensitivity analysis)	one-way sensitivity analysis, probabilistic (M onte Carlo) simulation	88%	3
Minimizing complications from nonsteroidal antiinflammatory drugs cost-effectiveness of competing strategies in varying risk groups	not included	2002 USD	Treeage	decision tree model	moderate dyspepsia, severe dyspepsia, ulcer haemorrhage without surgery, ulcer haemorrhage or perforation with surgery, myocardial infarction, post myocardial infarction state for survivors, death from an ulcer complication or myocardial infarction.	GI and CV Aes	one way sensitivity	79%	3
Use of the ACCES model to predict the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis in Norway	not mentioned	1999 Norwegian krone	not mentioned	decision tree model, Arthritis Cost Consequence Evaluation System (ACCES) pharmacoeconomic model,		GI events (serious GI event, symptomatic ulcers, anaemia, GI discomfort)	one-way	68%	2
Economic evaluation of nimesulide versus diclofenac in the treatment	not included	1999 Euros	Treeage	decision tree model	no AE, gastric AE, intestinal AE	GI AE	two way	91%	3

of osteoarthritis in France, Italy and Spain									
Modelling therapeutic strategies in the treatment of osteoarthritis an economic evaluation of meloxicam versus diclofenac and piroxicam	no included	1998 UK (drug costs were in 2000 values)	not mentioned	decision tree model		GI events, non GI events (renal, hepatic, CV), discontinuation due to lack of efficacy	one-way, probabilistic (stochastic), Monte Carlo simulation	86%	3
Knee Joint Distraction Compared to Total Knee Arthroplasty for Treatment of End Stage Osteoarthritis Simulating Long-Term Outcomes and Cost-Effectiveness	not mentioned	2013 Euros	not mentioned	individual patient Markov (or Health state) model	post KJD, post TKA, post revision TKA, post 2nd revision TKA/best supportive care	revision TKA (two times)	PSA, deterministic sensitivity analyses	83%	3
Cost analysis of flavocoxid compared to naproxen for management of mild to moderate OA	not mentioned	2009 Dollar	Treeage	decision tree model	no GI events, mild GI events, severe GI event	use of PPI, mild GI event, severe GI event, ulcer, bleeding ulcer, ulcer perforation	one, two-way sensitivity analysis	86%	3
Modeling the economic and health consequences of managing chronic osteoarthritis pain with opioids in Germany comparison of extended-release oxycodone and OROS hydromorphone	not mentioned	2005 Euros	ARENA	discrete event simulation		Aes, pain recurrence, discontinue the opioid	univariate sensitivity analyses, PSA	87%	3
Cost-utility analysis of duloxetine in osteoarthritis a US private payer perspective	not included	2011 USD	not mentioned	discrete-state, time-dependent semi-Markov model based on NICE model	fourteen health states comprised the structure of the model: treatment without persistent AE, six during-AE states, six post-AE states and death.	dyspepsia, serious GI and CV events (transient and persistent event), discontinuation	one-way, PSA	92%	3

Cost effectiveness of duloxetine for osteoarthritis a Quebec societal perspective	costs for loss of productivity	2011 Canadian dollars	Microsoft excel	Markov model (extended NICE model)	health states were defined by persistent Aes and a subsequent post-AE state	GI and CV event, fractures, transient and persistent AES, titration, discontinuation	one-way, and PSA	75%	2
Development and validation of a new population-based simulation model of osteoarthritis in New Zealand	NA	NA	not mentioned	the New Zealand Management of Osteoarthritis (NZ-MOA) model, a discrete-time state-transition microsimulation model	NA	NA	NA	NA	NA
Cost-effectiveness of adjunct non-pharmacological interventions for osteoarthritis of the knee	not mentioned	NA	not mentioned	decision analytic model	NA	NA	PSA	78%	3
Cost-effectiveness of treatment strategies for osteoarthritis of the knee in Taiwan	time lost from work	2002 US dollars	Treeage	decision tree model	OA with improvement from treatment, OA without improvement from treatment, OA with conventional treatment, OA with hyaluronan injections, OA with injection pain from hyaluronan treatment, serious GI complications	injection pain, GI AE	one-way	83%	3
Arthritis treatment in Hong Kong--cost analysis of celecoxib versus conventional NSAIDS, with or without gastroprotective agents	not mentioned	HK	Treeage and Microsoft excel	decision tree model	no toxicity, GI discomfort, symptomatic ulcer, anaemia with occult bleeding, GI complication	no toxicity, GI discomfort, symptomatic ulcer, anaemia with occult bleeding, GI complication	one way	77%	3

An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada	indirect costs and direct non-medical costs are not considered	1998 Canadian dollar	not mentioned	decision tree model	GI AE: no GI toxicity, GI distress, serious GI complications which can result in death or survival, symptomatic ulcer, anaemia with occult bleeding	GI distress, serious GI complications requiring hospitalisation, symptomatic ulcer, anaemia with occult bleeding, and GI-related death	one-way sensitivity analysis	73%	2

Supplement 3.5 Studies included in the systematic review

1. Akehurst RL, Backhouse M, Emery P, Haslock I, Kirsch J, McCabe CJ, et al. Nabumetone compared with Ibuprofen and a weighted NSAID combination: an economic evaluation. Journal of Medical Economics. 2008;1(1-4):19-37. doi: 10.3111/199801019037.

2. Al MJ, Maniadakis N, Grijseels EW, Janssen M. Costs and effects of various analgesic treatments for patients with rheumatoid arthritis and osteoarthritis in the Netherlands. Value Health. 2008;11(4):589-99. Epub 2008/01/16. doi: 10.1111/j.1524-4733.2007.00303.x. PubMed PMID: 18194404.

3. Andrew Moore R, Phillips CJ, Pellissier JM, Kong SX. Health economic comparisons of rofecoxib versus conventional nonsteroidal antiinflammatory drugs for osteoarthritis in the United Kingdom. Journal of Medical Economics. 2008;4(1-4):1-17. doi: 10.3111/200104001017.

4. Barlow BT, McLawhorn AS, Westrich GH. The Cost-Effectiveness of Dual Mobility Implants for Primary Total Hip Arthroplasty: A Computer-Based Cost-Utility Model. J Bone Joint Surg Am. 2017;99(9):768-77. Epub 2017/05/04. doi: 10.2106/JBJS.16.00109. PubMed PMID: 28463921.

5. Bedair H, Cha TD, Hansen VJ. Economic benefit to society at large of total knee arthroplasty in younger patients: a Markov analysis. J Bone Joint Surg Am. 2014;96(2):119-26. Epub 2014/01/17. doi: 10.2106/JBJS.L.01736. PubMed PMID: 24430411.

6. Bessette L, Risebrough N, Mittmann N, Roussy JP, Ho J, Zlateva G. Cost-utility of celecoxib use in different treatment strategies for osteoarthritis and rheumatoid arthritis from the Quebec healthcare system perspective. J Med Econ. 2009;12(3):246-58. Epub 2009/09/12. doi: 10.3111/13696990903288970. PubMed PMID: 19743942.

7. Bove AM, Smith KJ, Bise CG, Fritz JM, Childs JD, Brennan GP, et al. Exercise, manual therapy, and booster sessions in knee osteoarthritis: cost-effectiveness analysis from a multicenter randomized controlled trial. Physical therapy. 2017;98(1):16-27.

8. Bozic KJ, Pui CM, Ludeman MJ, Vail TP, Silverstein MD. Do the potential benefits of metal-on-metal hip resurfacing justify the increased cost and risk of complications? Clin Orthop Relat Res. 2010;468(9):2301-12. Epub 2010/03/17. doi: 10.1007/s11999-010-1301-0. PubMed PMID: 20232182; PubMed Central PMCID: PMCPMC2914258.

9. Brereton N, Pennington B, Ekelund M, Akehurst R. A cost-effectiveness analysis of celecoxib compared with diclofenac in the treatment of pain in osteoarthritis (OA) within the Swedish health system using an adaptation of the NICE OA model. J Med Econ. 2014;17(9):677-84. Epub 2014/06/11. doi: 10.3111/13696998.2014.933111. PubMed PMID: 24914585.

10. Brereton N, Winn B, Akehurst R. The cost-effectiveness of celecoxib vs diclofenac in the treatment of osteoarthritis in the UK; an update to the NICE model using data from the CONDOR trial. J Med Econ. 2012;15(3):465-72. Epub 2012/01/21. doi: 10.3111/13696998.2012.659778. PubMed PMID: 22260652.

11. Brown CA, Watters TS, Mather Iii RC, Orlando LA, Bolognesi MP, Moorman Iii CT, et al. Cost-Effectiveness Analysis of Unicompartmental Knee Arthroplasty and High Tibial Osteotomy for Treatment of Medial Compartmental Osteoarthritis. The Duke Orthopaedic Journal. 2011;1:27-36. doi: 10.5005/jp-journals-10017-1005.

12. Burn E, Liddle AD, Hamilton TW, Judge A, Pandit HG, Murray DW, et al. Cost-effectiveness of unicompartmental compared with total knee replacement: a population-based study using data from the National Joint Registry for England and Wales. BMJ open. 2018;8(4):e020977.

13. Capel M, Tornero J, Zamorano JL, Oyagüez I, Casado MÁ, Sánchez-Covisa J, et al. Efficiency of naproxen/esomeprazole in association for osteoarthrosis treatment in Spain. Reumatología Clínica (English Edition). 2014;10(4):210-7.

14. Castro JC, Daza AM, Misas JD. Cost-effectiveness Analysis of Viscosupplementation versus Conventional Supportive Therapy for Knee Osteoarthritis in Colombia. Value Health Reg Issues. 2015;8:56-61. Epub 2015/12/01. doi: 10.1016/j.vhri.2015.03.018. PubMed PMID: 29698172.

15. Chancellor JV, Hunsche E, de Cruz E, Sarasin FP. Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor, in Switzerland. Pharmacoeconomics. 2001;19(1):59-75.

16. Chang RW, Pellissier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. Jama. 1996;275(11):858-65.

17. Chawla H, Nwachukwu B, Van Der List J, Eggman A, Pearle A, Ghomrawi H. Cost effectiveness of patellofemoral versus total knee arthroplasty in younger patients. The bone & joint journal. 2017;99(8):1028-36.

18. Contreras-Hernandez I, Mould-Quevedo JF, Torres-Gonzalez R, Goycochea-Robles MV, Pacheco-Dominguez RL, Sanchez-Garcia S, et al. Cost-effectiveness analysis for joint pain treatment in patients with osteoarthritis treated at the Instituto Mexicano del Seguro Social (IMSS): Comparison of nonsteroidal anti-inflammatory drugs (NSAIDs) vs. cyclooxygenase-2 selective inhibitors. Cost Eff Resour Alloc. 2008;6:21. Epub 2008/11/19. doi: 10.1186/1478-7547-6-21. PubMed PMID: 19014495; PubMed Central PMCID: PMCPMC2626578.

19. Cummins JS, Tomek IM, Kantor SR, Furnes O, Engesaeter LB, Finlayson SR. Cost-effectiveness of antibiotic-impregnated bone cement used in primary total hip arthroplasty. J Bone Joint Surg Am. 2009;91(3):634-41. Epub 2009/03/04. doi: 10.2106/JBJS.G.01029. PubMed PMID: 19255224.

20. Di Tanna GL, Ferro S, Cipriani F, Bordini B, Stea S, Toni A, et al. Modeling the cost-effectiveness for cement-less and hybrid prosthesis in total hip replacement in Emilia Romagna, Italy. J Surg Res. 2011;169(2):227-33. Epub 2010/01/26. doi: 10.1016/j.jss.2009.10.031. PubMed PMID: 20097368.

21. Faucett SC, Geisler BP, Chahla J, Krych AJ, Kurzweil PR, Garner AM, et al. Meniscus Root Repair vs Meniscectomy or Nonoperative Management to Prevent Knee Osteoarthritis After Medial Meniscus Root Tears: Clinical and Economic Effectiveness. Am J Sports Med. 2019;47(3):762-9. Epub 2018/03/09. doi: 10.1177/0363546518755754. PubMed PMID: 29517925.

22. Feeley BT, Liu S, Garner AM, Zhang AL, Pietzsch JB. The cost-effectiveness of meniscal repair versus partial meniscectomy: A model-based projection for the United States. The Knee. 2016;23(4):674-80. doi: 10.1016/j.knee.2016.03.006.

23. Ghomrawi HM, Eggman AA, Pearle AD. Effect of age on cost-effectiveness of unicompartmental knee arthroplasty compared with total knee arthroplasty in the U.S. J Bone Joint Surg Am. 2015;97(5):396-402. Epub 2015/03/06. doi: 10.2106/JBJS.N.00169. PubMed PMID: 25740030; PubMed Central PMCID: PMCPMC4344593.

24. Haglund U, Svarvar P. The Swedish ACCES model: predicting the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis. Rheumatology. 2000;39(suppl_2):51-6.

25. Hatoum HT, Fierlinger AL, Lin SJ, Altman RD. Cost-effectiveness analysis of intra-articular injections of a high molecular weight bioengineered hyaluronic acid for the treatment of osteoarthritis knee pain. J Med Econ. 2014;17(5):326-37. Epub 2014/03/15. doi: 10.3111/13696998.2014.902843. PubMed PMID: 24625229.

26. Heintzbergen S, Kulin NA, Ijzerman MJ, Steuten LM, Werle J, Khong H, et al. Cost-utility of metal-on-metal hip resurfacing compared to conventional total hip replacement in young active patients with osteoarthritis. Value Health. 2013;16(6):942-52. Epub 2013/09/18. doi: 10.1016/j.jval.2013.06.021. PubMed PMID: 24041344.

27. Higashi H, Barendregt JJ. Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis: discrete-event simulation model. PLoS One. 2011;6(9):e25403. Epub 2011/10/04. doi: 10.1371/journal.pone.0025403. PubMed PMID: 21966520; PubMed Central PMCID: PMCPMC3179521.

28. Jansen R, Burrell A, Nuijten M, Hardens M. An economic evaluation of meloxicam 7.5 mg versus diclofenac 100 mg retard in the treatment of osteoarthritis in the UK: a decision analysis model based on gastrointestinal complications. British Journal of Medical Economics. 1996;10:247-62.

29. Jansen R, Capri S, Nuijten M, Burrell A, Marini M, Hardens M. Economic evaluation of meloxicam (7.5 mg) versus sustained release diclofenac (100 mg) treatment for osteoarthritis: a cross-national assessment for France, Italy and the UK. British journal of medical economics. 1997;11:9-22.

30. Kamath CC, Kremers HM, Vanness DJ, O'Fallon WM, Cabanela RL, Gabriel SE. The cost-effectiveness of acetaminophen, NSAIDs, and selective COX-2 inhibitors in the treatment of symptomatic knee osteoarthritis. Value Health. 2003;6(2):144-57. Epub 2003/03/19. doi: 10.1046/j.1524-4733.2003.00215.x. PubMed PMID: 12641865.

31. Karmarkar TD, Maurer A, Parks ML, Mason T, Bejinez-Eastman A, Harrington M, et al. A fresh perspective on a familiar problem: examining disparities in knee osteoarthritis using a Markov model. Medical care. 2017;55(12):993.

32. Katz JN, Smith SR, Collins JE, Solomon DH, Jordan JM, Hunter DJ, et al. Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities. Osteoarthritis Cartilage. 2016;24(3):409-18. Epub 2015/11/04. doi: 10.1016/j.joca.2015.10.006. PubMed PMID: 26525846; PubMed Central PMCID: PMCPMC4761310.

33. Koenig L, Zhang Q, Austin MS, Demiralp B, Fehring TK, Feng C, et al. Estimating the Societal Benefits of THA After Accounting for Work Status and Productivity: A Markov Model Approach. Clin Orthop Relat Res. 2016;474(12):2645-54. Epub 2016/11/01. doi: 10.1007/s11999-016-5084-9. PubMed PMID: 27699631; PubMed Central PMCID: PMCPMC5085951.

34. Konopka JF, Gomoll AH, Thornhill TS, Katz JN, Losina E. The cost-effectiveness of surgical treatment of medial unicompartmental knee osteoarthritis in younger patients: a computer model-based evaluation. J Bone Joint Surg Am. 2015;97(10):807-17. Epub 2015/05/23. doi: 10.2106/JBJS.N.00925. PubMed PMID: 25995491; PubMed Central PMCID: PMCPMC4430101.

35. Kunkel ST, Sabatino MJ, Kang R, Jevsevar DS, Moschetti WE. The Cost-Effectiveness of Total Hip Arthroplasty in Patients 80 Years of Age and Older. J Arthroplasty. 2018;33(5):1359-67. Epub 2017/12/26. doi: 10.1016/j.arth.2017.11.063. PubMed PMID: 29276115.

36. Latimer N, Lord J, Grant RL, O'Mahony R, Dickson J, Conaghan PG. Value of Information in the Osteoarthritis Setting. Pharmacoeconomics. 2011;29(3):225-37.

37. Latimer N, Lord J, Grant RL, O'Mahony R, Dickson J, Conaghan PG, et al. Cost effectiveness of COX 2 selective inhibitors and traditional NSAIDs alone or in combination with a proton pump inhibitor for people with osteoarthritis. BMJ. 2009;339:b2538. Epub 2009/07/16. doi: 10.1136/bmj.b2538. PubMed PMID: 19602530; PubMed Central PMCID: PMCPMC2714674.

38. Li CS, Bhandari M. Cost-effectiveness of unicompartmental knee arthroplasty, high tibial osteotomy, and KineSpring® knee implant system for unicompartmental osteoarthritis of the knee. Journal of long-term effects of medical implants. 2013;23(2-03).

39. Liedgens H, Nuijten MJ, Nautrup BP. Economic evaluation of tramadol/paracetamol combination tablets for osteoarthritis pain in the Netherlands. Clinical drug investigation. 2005;25(12):785-802.

40. Losina E, Burbine SA, Suter LG, Hunter DJ, Solomon DH, Daigle ME, et al. Pharmacologic regimens for knee osteoarthritis prevention: can they be cost-effective? Osteoarthritis Cartilage. 2014;22(3):415-30. Epub 2014/02/04. doi: 10.1016/j.joca.2014.01.005. PubMed PMID: 24487044; PubMed Central PMCID: PMCPMC4006219.

41. Losina E, Daigle ME, Suter LG, Hunter DJ, Solomon DH, Walensky RP, et al. Disease-modifying drugs for knee osteoarthritis: can they be cost-effective? Osteoarthritis Cartilage. 2013;21(5):655-67. Epub 2013/02/06. doi: 10.1016/j.joca.2013.01.016. PubMed PMID: 23380251; PubMed Central PMCID: PMCPMC3670115.

42. Losina E, Michl G, Collins JE, Hunter DJ, Jordan JM, Yelin E, et al. Model-based evaluation of cost-effectiveness of nerve growth factor inhibitors in knee osteoarthritis: impact of drug cost, toxicity, and means of administration. Osteoarthritis Cartilage. 2016;24(5):776-85. Epub 2016/01/10. doi: 10.1016/j.joca.2015.12.011. PubMed PMID: 26746146; PubMed Central PMCID: PMCPMC4838505.

43. Losina E, Usiskin IM, Smith SR, Sullivan JK, Smith KC, Hunter DJ, et al. Cost-effectiveness of generic celecoxib in knee osteoarthritis for average-risk patients: a model-based evaluation. Osteoarthritis Cartilage. 2018;26(5):641-50. Epub 2018/02/27. doi: 10.1016/j.joca.2018.02.898. PubMed PMID: 29481917; PubMed Central PMCID: PMCPMC6334297.

44. Losina E, Walensky RP, Kessler CL, Emrani PS, Reichmann WM, Wright EA, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. Arch Intern Med. 2009;169(12):1113-21; discussion 21-2. Epub 2009/06/24. doi: 10.1001/archinternmed.2009.136. PubMed PMID: 19546411; PubMed Central PMCID: PMCPMC2731300.

45. Loyd M, Rublee D, Jacobs P. An economic model of long-term use of celecoxib in patients with osteoarthritis. BMC Gastroenterol. 2007;7:25. Epub 2007/07/06. doi: 10.1186/1471-230X-7-25. PubMed PMID: 17610716; PubMed Central PMCID: PMCPMC1925103.

46. Maetzel A, Krahn M, Naglie G. The cost effectiveness of rofecoxib and celecoxib in patients with osteoarthritis or rheumatoid arthritis. Arthritis Rheum. 2003;49(3):283-92. Epub 2003/06/10. doi: 10.1002/art.11121. PubMed PMID: 12794781.

47. Mari K, Degieux P, Mistretta F, Guillemin F, Richette P. Cost utility modeling of early vs late total knee replacement in osteoarthritis patients. Osteoarthritis Cartilage. 2016;24(12):2069-76. Epub 2016/08/06. doi: 10.1016/j.joca.2016.07.013. PubMed PMID: 27492465.

48. Marshall JK, Pellissier JM, Attard CL, Kong SX, Marentette MA. Incremental cost-effectiveness analysis comparing rofecoxib with nonselective NSAIDs in osteoarthritis. Pharmacoeconomics. 2001;19(10):1039-49.

49. Mather RC, 3rd, Hug KT, Orlando LA, Watters TS, Koenig L, Nunley RM, et al. Economic evaluation of access to musculoskeletal care: the case of waiting for total knee arthroplasty. BMC Musculoskelet Disord. 2014;15:22. Epub 2014/01/21. doi: 10.1186/1471-2474-15-22. PubMed PMID: 24438051; PubMed Central PMCID: PMCPMC3897923.

50. Mather RC, 3rd, Watters TS, Orlando LA, Bolognesi MP, Moorman CT, 3rd. Cost effectiveness analysis of hemiarthroplasty and total shoulder arthroplasty. J Shoulder Elbow Surg. 2010;19(3):325-34. Epub 2010/03/23. doi: 10.1016/j.jse.2009.11.057. PubMed PMID: 20303459.

51. McKell C, Stewart A. Cost-minimisation analysis comparing topical versus systematic NSAIDS in the treatment of mild osteoarthritis of the superficial joints. British Journal of Medical Economics. 1994;7:137-46.

52. Moore A, Phillips C, Hunsche E, Pellissier J, Crespi S. Economic evaluation of etoricoxib versus non-selective NSAIDs in the treatment of osteoarthritis and rheumatoid arthritis patients in the UK. Pharmacoeconomics. 2004;22(10):643-60.

53. Moschetti WE, Konopka JF, Rubash HE, Genuario JW. Can Robot-Assisted Unicompartmental Knee Arthroplasty Be Cost-Effective? A Markov Decision Analysis. J Arthroplasty. 2016;31(4):759-65. Epub 2015/12/27. doi: 10.1016/j.arth.2015.10.018. PubMed PMID: 26706836.

54. Mota RE. Cost-effectiveness analysis of early versus late total hip replacement in Italy. Value Health. 2013;16(2):267-79. Epub 2013/03/30. doi: 10.1016/j.jval.2012.10.020. PubMed PMID: 23538178.

55. Mujica-Mota RE, Watson LK, Tarricone R, Jager M. Cost-effectiveness of timely versus delayed primary total hip replacement in Germany: A social health insurance perspective. Orthop Rev (Pavia). 2017;9(3):7161. Epub 2017/10/27. doi: 10.4081/or.2017.7161. PubMed PMID: 29071040; PubMed Central PMCID: PMCPMC5641833.

56. Nasef SA, Shaaban AA, Mould-Quevedo J, Ismail TA. The cost-effectiveness of celecoxib versus non-steroidal anti-inflammatory drugs plus proton-pump inhibitors in the treatment of osteoarthritis in Saudi Arabia. Health Econ Rev. 2015;5(1):53. Epub 2015/06/11. doi: 10.1186/s13561-015-0053-7. PubMed PMID: 26061682; PubMed Central PMCID: PMCPMC4467807.

57. Peersman G, Jak W, Vandenlangenbergh T, Jans C, Cartier P, Fennema P. Cost-effectiveness of unicondylar versus total knee arthroplasty: a Markov model analysis. The Knee. 2014;21:S37-S42. doi: 10.1016/s0968-0160(14)50008-7.

58. Pellissier JM, Straus WL, Watson DJ, Kong SX, Harper SE. Economic evaluation of rofecoxib versus nonselective nonsteroidal anti-inflammatory drugs for the treatment of osteoarthritis. Clinical therapeutics. 2001;23(7):1061-79.

59. Pennington M, Grieve R, Black N, van der Meulen JH. Cost-Effectiveness of Five Commonly Used Prosthesis Brands for Total Knee Replacement in the UK: A Study Using the NJR Dataset. PLoS One. 2016;11(3):e0150074. Epub 2016/03/05. doi: 10.1371/journal.pone.0150074. PubMed PMID: 26943789; PubMed Central PMCID: PMCPMC4778929.

60. Pennington M, Grieve R, Sekhon JS, Gregg P, Black N, van der Meulen JH. Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis. BMJ. 2013;346:f1026. Epub 2013/03/01. doi: 10.1136/bmj.f1026. PubMed PMID: 23447338; PubMed Central PMCID: PMCPMC3583598.

61. Pulikottil-Jacob R, Connock M, Kandala N, Mistry H, Grove A, Freeman K, et al. Cost effectiveness of total hip arthroplasty in osteoarthritis: comparison of devices with differing bearing surfaces and modes of fixation. The bone & joint journal. 2015;97(4):449-57.

62. Pulikottil-Jacob R, Connock M, Kandala NB, Mistry H, Grove A, Freeman K, et al. Has Metal-On-Metal Resurfacing Been a Cost-Effective Intervention for Health Care Providers?-A Registry Based Study. PLoS One. 2016;11(11):e0165021. Epub 2016/11/02. doi: 10.1371/journal.pone.0165021. PubMed PMID: 27802289; PubMed Central PMCID: PMCPMC5089767.

63. Rongen JJ, Govers TM, Buma P, Grutters JP, Hannink G. Societal and Economic Effect of Meniscus Scaffold Procedures for Irreparable Meniscus Injuries. Am J Sports Med. 2016;44(7):1724-34. Epub 2016/05/10. doi: 10.1177/0363546516639290. PubMed PMID: 27159309.

64. Rongen JJ, Govers TM, Buma P, Rovers MM, Hannink G. Arthroscopic meniscectomy for degenerative meniscal tears reduces knee pain but is not cost-effective in a routine health care setting: a multi-center longitudinal observational study using data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2018;26(2):184-94. Epub 2017/09/14. doi: 10.1016/j.joca.2017.02.805. PubMed PMID: 28903016.

65. Ruiz D, Jr., Koenig L, Dall TM, Gallo P, Narzikul A, Parvizi J, et al. The direct and indirect costs to society of treatment for end-stage knee osteoarthritis. J Bone Joint Surg Am. 2013;95(16):1473-80. Epub 2013/08/24. doi: 10.2106/JBJS.L.01488. PubMed PMID: 23965697.

66. Schaefer M, DeLattre M, Gao X, Stephens J, Botteman M, Morreale A. Assessing the cost-effectiveness of COX-2 specific inhibitors for arthritis in the Veterans Health Administration. Current Medical Research and Opinion. 2004;21(1):47-60. doi: 10.1185/030079904x17974.

67. Slover J, Espehaug B, Havelin LI, Engesaeter LB, Furnes O, Tomek I, et al. Cost-effectiveness of unicompartmental and total knee arthroplasty in elderly low-demand patients: a Markov decision analysis. JBJS. 2006;88(11):2348-55.

68. Slover JD, Tosteson AN, Bozic KJ, Rubash HE, Malchau H. Impact of hospital volume on the economic value of computer navigation for total knee replacement. J Bone Joint Surg Am. 2008;90(7):1492-500. Epub 2008/07/03. doi: 10.2106/JBJS.G.00888. PubMed PMID: 18594098; PubMed Central PMCID: PMCPMC2657305.

69. Smith SR, Katz JN, Collins JE, Solomon DH, Jordan JM, Suter LG, et al. Cost-Effectiveness of Tramadol and Oxycodone in the Treatment of Knee Osteoarthritis. Arthritis

Care Res (Hoboken). 2017;69(2):234-42. Epub 2016/04/26. doi: 10.1002/acr.22916. PubMed PMID: 27111538; PubMed Central PMCID: PMCPMC5378156.

70. Smith WB, 2nd, Steinberg J, Scholtes S, McNamara IR. Medial compartment knee osteoarthritis: age-stratified cost-effectiveness of total knee arthroplasty, unicompartmental knee arthroplasty, and high tibial osteotomy. Knee Surg Sports Traumatol Arthrosc. 2017;25(3):924-33. Epub 2015/11/02. doi: 10.1007/s00167-015-3821-3. PubMed PMID: 26520646.

71. SooHoo NF, Kominski G. Cost-effectiveness analysis of total ankle arthroplasty. JBJS. 2004;86(11):2446-55.

72. Soohoo NF, Sharifi H, Kominski G, Lieberman JR. Cost-effectiveness analysis of unicompartmental knee arthroplasty as an alternative to total knee arthroplasty for unicompartmental osteoarthritis. JBJS. 2006;88(9):1975-82.

73. Spiegel BM, Chiou CF, Ofman JJ. Minimizing complications from nonsteroidal antiinflammatory drugs: cost-effectiveness of competing strategies in varying risk groups. Arthritis Rheum. 2005;53(2):185-97. Epub 2005/04/09. doi: 10.1002/art.21065. PubMed PMID: 15818647.

74. Spiegel BM, Targownik L, Dulai GS, Gralnek IM. The cost-effectiveness of cyclooxygenase-2 selective inhibitors in the management of chronic arthritis. Annals of internal medicine. 2003;138(10):795-806.

75. Svarvar P, Aly A. Use of the ACCES model to predict the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis in Norway. Rheumatology. 2000;39(suppl_2):43-50.

76. Tarricone R, Martelli E, Parazzini F, Darbà J, Le Pen C, Rovira J. Economic evaluation of nimesulide versus diclofenac in the treatment of osteoarthritis in France, Italy and Spain. Clinical Drug Investigation. 2001;21(7):453-64.

77. Tavakoli M. Modelling Therapeutic Strategies in the Treatment of Osteoarthritis. Pharmacoeconomics. 2003;21(6):443-54.

78. van der Woude JA, Nair SC, Custers RJ, van Laar JM, Kuchuck NO, Lafeber FP, et al. Knee Joint Distraction Compared to Total Knee Arthroplasty for Treatment of End Stage Osteoarthritis: Simulating Long-Term Outcomes and Cost-Effectiveness. PLoS One.

2016;11(5):e0155524. Epub 2016/05/14. doi: 10.1371/journal.pone.0155524. PubMed PMID: 27171268; PubMed Central PMCID: PMCPMC4865158.

79. Walton SM, Schumock GT, McLain DA. Cost analysis of flavocoxid compared to naproxen for management of mild to moderate OA. Curr Med Res Opin. 2010;26(9):2253-61. Epub 2010/08/10. doi: 10.1185/03007995.2010.505545. PubMed PMID: 20690891.

80. Wang CB, Xie X, Wesley F, Ronald C. Cost-effectiveness Study of Celecoxib for Osteoarthritis in China. China Pharmacy. 2017;28(5):586-90.

81. Ward A, Bozkaya D, Fleischmann J, Dubois D, Sabatowski R, Jaime Caro J. Modeling the economic and health consequences of managing chronic osteoarthritis pain with opioids in Germany: comparison of extended-release oxycodone and OROS hydromorphone. Current medical research and opinion. 2007;23(10):2333-45.

82. Wielage RC, Bansal M, Andrews JS, Klein RW, Happich M. Cost-utility analysis of duloxetine in osteoarthritis: a US private payer perspective. Appl Health Econ Health Policy. 2013;11(3):219-36. Epub 2013/04/26. doi: 10.1007/s40258-013-0031-3. PubMed PMID: 23616247.

83. Wielage RC, Patel AJ, Bansal M, Lee S, Klein RW, Happich M. Cost Effectiveness of Duloxetine for Osteoarthritis: A Quebec Societal Perspective. Arthritis Care & Research. 2014;66(5):702-8. doi: 10.1002/acr.22224.

84. Wilson R, Abbott JH. Development and validation of a new population-based simulation model of osteoarthritis in New Zealand. Osteoarthritis Cartilage. 2018;26(4):531-9. Epub 2018/01/15. doi: 10.1016/j.joca.2018.01.004. PubMed PMID: 29331740.

85. Woods B, Manca A, Weatherly H, Saramago P, Sideris E, Giannopoulou C, et al. Cost-effectiveness of adjunct non-pharmacological interventions for osteoarthritis of the knee. PLoS One. 2017;12(3):e0172749. Epub 2017/03/08. doi: 10.1371/journal.pone.0172749. PubMed PMID: 28267751; PubMed Central PMCID: PMCPMC5340388.

86. Yen Z-S, Lai M-S, Wang C-T, Chen L-S, Chen S-C, Chen W-J, et al. Cost-effectiveness of treatment strategies for osteoarthritis of the knee in Taiwan. The Journal of rheumatology. 2004;31(9):1797-803.

87. You J, Lee K, Chan T, Lau W, Chan F. Arthritis treatment in Hong Kong—cost analysis of celecoxib versus conventional NSAIDS, with or without gastroprotective agents. Alimentary pharmacology & therapeutics. 2002;16(12):2089-96.

88. Zabinski RA, Burke TA, Johnson J, Lavoie F, Fitzsimon C, Tretiak R, et al. An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada. Pharmacoeconomics. 2001;19(1):49-58.

Supplement 3.6 Interventions in the included studies

KNEE OA (n=37)		
pharmacological treatment (n=10)		
interventions	Frequency	References*
celecoxib vs. naproxen	1	43
opioid, tramadol, tramadol+oxycodone	1	70
Tanezumab	1	42
NSAIDs and opioids	1	32
Hylan G-F 20	2	14,25
disease modified drugs	1	41
naproxen, celecoxib, hyaluronan	1	86
rofecoxib and celecoxib vs. acetaminophen or ibuprofen	1	30
pharmacologic regimens	1	40
surgical treatment (n=23)		
interventions	Frequency	References*
meniscus repair, meniscectomy, and nonoperative treatment	4	21,22,63,64
unicompartmental knee arthroplasty (UKA) with total knee arthroplasty (TKA)	5	12,23,57,67,72
Patellofemoral arthroplasty vs. TKA	1	17
TKA, UKA, and high tibial osteotomy (HTO)	2	34,69
Knee Joint Distraction vs. TKA	1	78
early vs. late TKA	1	47
five commonly used prosthesis for TKA	1	59
robot assistant UKA vs. traditional UKA	1	53
TKA vs. nonoperative treatment	3	5,44,65
TKA without delay vs. waiting period with/without non-operative treatment	1	49
computer-assisted surgery vs. TKA	1	68
UKA vs. HTO	2	11,38
Others treatment (n=4)		

interventions	Frequency	References*
exercise, booster, manual therapy	1	7
adjunct non-pharmacological therapy	1	85
no specific treatment	1	84
10 types of treatment	1	31
HIP OA (n=13)		
surgical treatment (n=13)		
interventions	Frequency	References*
Dual mobility implants vs. conventional bearings	1	4
total hip arthroplasty (THA) vs. non-operative management	3	16,33,35
timely THR, delayed THR, and non-surgical therapy	2	54,55
metal-on-metal hip resurfacing procedure vs. commonly employed THR	3	8,26,62
five commonly used combination of components in THA, including type of fixation and bearing of surface	1	61
three commonly used prosthesis for THR	1	60
cement-less with hybrid prosthesis in THR	1	20
antibiotic-impregnated bone cement for THA vs. cement without antibiotics	1	19
OA (n=36)		
pharmacological treatment (n=35)		
interventions	Frequency	References*
cox-2 selective vs. non-selective NSAID, with and without PPI	8	9,10,13,36,37,56,75,80
cox-2 selective vs. non-selective NSAID (NSNSAID alone, NSNSAID+PPI, NSNSAID +H2RA, NSNSAID+ misoprostol)	4	2,15,52,88
cox-2 selective vs. non-selective NSAID	14	1,3,28,29,45,46,48,58,66, 74,76,77,79,87
duloxetine, NSAIDs (celecoxib, diclofenac, Naproxen), opioid (Hydromorphone, Oxycodone)	1	83
duloxetine vs. celecoxib, naproxen, a combination of oxycodone/acetaminophen, oxycodone extended release, and tramadol and tapentadol	1	82

celecoxib; traditional NSIDs switch to celecoxib; traditional NSAIDs+ PPI switch to celecoxib	1	6
celecoxib, nonselective NSAIDs (naproxen, diclofenac and piroxicam) and acetaminophen	1	18
The Osmotic Controlled-Release Oral delivery System hydromorphone vs. equianalgesic dose of extended-release oxycodone	1	81
tramadol/paracetamol, NSAIDs alone, NSAIDs+PPIs, NSAIDs+H2RAs	1	39
NSAID alone, NSAID with PPI, coxib alone (with and without aspirin were compared)	1	73
celecoxib	1	24
piroxicam vs. ibuprofen	1	51
surgical treatment (n=1)		
total hip and knee replacements compared with 'doing nothing'(continued non-surgical therapies without joint replacements)	1	27
Other joints OA (n=2)		
surgical treatment (n=2)		
total shoulder arthroplasty and hemiarthroplasty	1	50
total ankle arthroplasty vs ankle fusion	1	71
Note: *references were shown in Supplement 3.5		

Supplement 3.7

There were four commonly used osteoarthritis (OA) model structures : 1) OA policy (OAPoL) model which was adopted in 6 studies which is a Markov model to simulate the natural history of knee OA and predominantly used in the US; 2) the National Institute for Health and Care Excellence (NICE) model which was adopted in 7 studies which is a Markov model originally developed in the UK to compare Nonsteroidal anti-inflammatory drugs (NSAID)/ cyclooxygenase (COX)-2 inhibitor oral analgesics and subsequently extended to incorporate dose titration, discontinuation and adverse events (AEs) in addition to gastrointestinal (GI) and cardiovascular (CV) AEs; 3) a model developed by Fitzpatrick which was adopted in 4 studies. It is a Markov model aimed at evaluating hip OA surgical treatments; and 4) a decision-tree model developed by Burke which was adopted in 4 studies to compare NSAID/COX-2 inhibitor oral analgesics. Further specific details of each model are mentioned below. See also the table that is included at the end of this document to summarise these commonly used OA health economic evaluation model structures.

(1) The Osteoarthritis Policy (OAPol) Model

The OAPol Model is a state-transition, computer simulation model of the natural history of knee OA that runs on an annual cycle. “State transition” refers to the fact that the model characterizes each person’s history as a sequence of annual transitions from one health state to another. Annual transition probabilities used in the OAPol model, are derived from published data or secondary data analyses. Health states are chosen to describe the individual’s current health including the number of comorbidities, obesity status and knee OA status. They are designed to be predictive of comorbidities and mortality. The model defines four general health state categories: knee OA- and obesity-free, knee OA only, obesity only, knee OA and obesity. Throughout most of their lives, patients reside in one of these chronic states. Death can occur in any state. The OAPol Model utilizes the Kellgren-Lawrence (K-L) scale to define OA severity: K-L 0 (normal radiograph) is defined as ‘no OA,’ K-L 1 (questionable osteophytes) as ‘pre-radiographic OA,’ K-L 2 (definite osteophytes) as ‘early OA,’ K-L 3 (< 50% narrowing of knee joint space) as ’advanced OA,’ and K-L 4 (≥ 50% narrowing of joint space) as ’end-stage OA.’ Symptomatic knee OA is defined as the concomitant presence of radiographic knee OA and knee pain on most days.

The OAPol Model tracks subjects’ life courses until death. Over their life spans, subjects without knee OA are at risk for developing OA and subjects with the disease are at risk for

progressing to more advanced stages based on subjects' current K-L grade and obesity, as defined by body mass index (BMI).

Knee OA incidence and progression rates in the OAPol Model are stratified by obesity and sex. Incidence is further stratified by year of age and progression is further stratified by K-L grade. At the beginning of a simulation, subjects are assigned a K-L grade and symptom status, based on their age and BMI. During each model cycle (one year), subjects may develop knee OA if they are currently OA-free or progress by one K-L grade if they already have OA. For example, a subject with symptomatic early OA (K-L 2) at baseline surviving to the following year may be assigned one of two states: (1) symptomatic early OA (where they started), or (2) symptomatic advanced OA (K-L 3). Competing risks for mortality are accounted for in the OAPol Model by incorporating several major comorbidities including cardiovascular disease, diabetes, chronic obstructive pulmonary disorders and malignancies.

Source: *Holt HL, Katz JN, Reichmann WM, Gerlovin H, Wright EA, Hunter DJ, et al. Forecasting the burden of advanced knee osteoarthritis over a 10-year period in a cohort of 60–64-year-old US adults. Osteoarthritis and cartilage. 2011;19(1):44-50*

(2) National Institute of Health and Care Excellence (NICE) model

The model is in the form of a Markov model with a 3-month cycle length. The probability of moving between states is based on within-state decision trees which are informed by clinical evidence and expert opinion. The health states that make up the Markov model represent a range of possible adverse events.

The model seeks to compare the cost effectiveness of individual NSAIDs and COX-2 inhibitors for which sufficient adverse event data exists. Patients do not move between treatments in the model (apart from the addition of a PPI in some circumstances and switching to paracetamol following serious adverse events or at the end of the treatment period). This is a simplifying assumption which keeps the model manageable. Therefore, the model considers first-line NSAID or COX-2 inhibitor treatment.

The model can be split into two key components:

• Markov model health states

• Within state decision trees to determine type of adverse event (if any).

The possible health states considered in the model are as follows:

• no complications

• GI symptoms / dyspepsia

• symptomatic ulcer

• post-symptomatic ulcer

• complicated GI bleed

• post-complicated GI bleed

• myocardial infarction (MI)

• post MI

• stroke

• post Stroke

• heart failure (HF)

• post HF

• post treatment (given no serious adverse events during the treatment period).

Source: Conditions NCCfC, Excellence NIfC, editors. Osteoarthritis: national clinical guidelines for care and management in adults2008: Royal College of Physicians

(3) Model developed by Fitzpatrick and colleagues

This Markov type model is aimed at predicting the prognosis of patients who have undergone primary total hip arthroplasty. Following the operation, patients are assumed to enter one of the four distinct Markov states.

Successful primary: if patients survive the initial total hip replacement (THR) they move to this state.

Revision THR: patients move to this state if their hip replacement fails (e.g. due to infection or loosening) and they then require revision surgery. As some patients require more than one revision operation, it is possible for a patient to move into this state more than once. Patients only remain in this state for one cycle.

Successful revision: if patients survive revision surgery they progress to this state.

Death: patients can die and enter this state at any point in the model. Patients can enter this state due to death related to surgery or due to the underlying risk of death.

Source: Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, et al. Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses. Health technology assessment (Winchester, England). 1998;2(20):1-64

(4) Model developed by Burke

This model is based on a decision-tree framework and aims to compare different treatments for patients with OA and Rheumatoid Arthritis (RA). The model compares the costs and clinical events for patients on the following arthritis treatments: NSAID alone, celecoxib, NSAID and prophylactic PPIs (NSAID + PPI), NSAID and prophylactic H2RAs (NSAID + H2RA), NSAID and prophylactic misoprostol (NSAID + misoprostol), and a single-tablet formulation of diclofenac/ misoprostol.

The primary clinical outcomes measured in the model are GI distress, serious GI complications requiring hospitalisation, symptomatic ulcer, anaemia with occult bleeding, and GI-related death.

The model uses the following definitions for the clinical outcomes:

• GI distress – moderate to severe dyspepsia, abdominal pain, or nausea

• Serious GI complications – any of the following events requiring hospitalisation: gross GI bleed, perforation, gastric outlet obstruction, GI distress, ulcer, or anaemia

• Symptomatic ulcer – uncomplicated, symptomatic gastric or duodenal ulcer confirmed by endoscopy (ulcers detected via protocol-driven endoscopy were not included)

• Anaemia with occult bleeding – anaemia serious enough to require withdrawal from the trial, or classified by the investigator as ‘severe’

• GI-related death – mortality associated with a hospitalised GI adverse event; the mortality rate of serious GI complications resulting in hospitalisation was assumed to be 10%.

Source: Zabinski RA, Burke TA, Johnson J, Lavoie F, Fitzsimon C, Tretiak R, et al. An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada. PharmacoEconomics. 2001;19 Suppl 1:49-58.

Table. Summary of four identified commonly used OA health economic evaluation models

Model Structure	Studies (n)*	Publication years	Study settings	Model type	OA type of interest	OA treatment of interest	Key model features	Key Strengths	Key Weakness
OAPoL	6	2013, 2014, 2016, 2017, 2018	US	Markov	Knee OA	Simulate the nature history of knee OA	Knee OA incidence and progression rate are stratified by obesity and sex.	Considers inter-relationships among pain, obesity and comorbidities	Sex and obesity specific data on knee OA incidence and progression rate is not available for countries other than USA but is imperative to build the OAPoL model structure.
NICE	7	2009, 2011, 2012, 2013, 2014, 2015	UK, US, Canada, Sweden, Saudi Arabia	Markov	OA	NSAIDs/Cox-2	Consider GI and CV AEs	Considers GI, CV AES, and treatment discontinuation	Lack of adequate data on discontinuation
Burke	4	2001, 2002, 2008	Canada, Netherlands, Switzerland, Hong Kong	Decision-tree	OA	NSAIDs/Cox-2	Consider GI AEs	Easy to be adapted for different settings	Fails to consider other AEs, discontinuation, and is based on short time horizon
Fitzpatrick	4	2013, 2015, 2016, 2017	UK, Germany, Italy	Markov	Hip OA	Surgery	Hip OA surgery	Easy to be adapted for alternative settings and study objectives	Fails to consider the impacts of other important variables (for example, age) on surgery outcomes

*This is the number of included studies adopted by each of the four model structures.

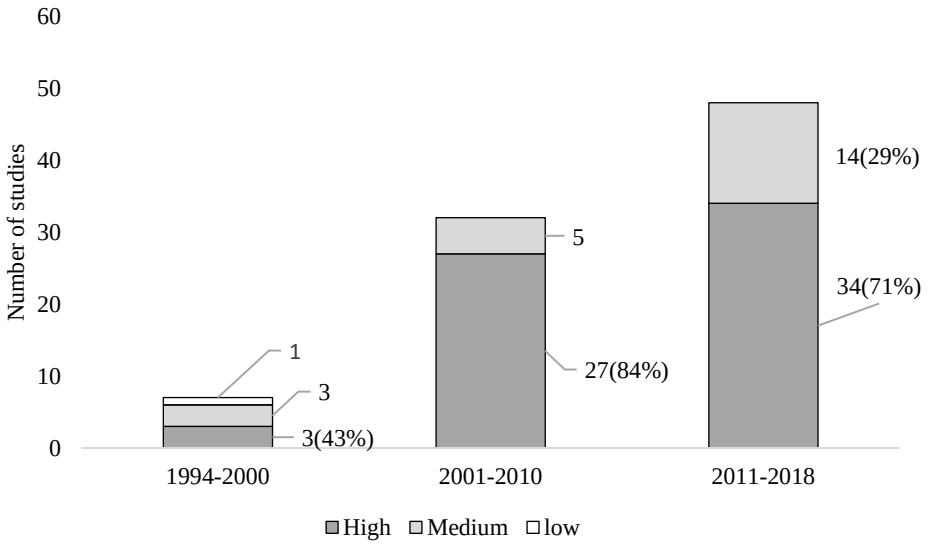
OAPoL: OA policy model; NICE: National Institute of Health and Care Excellence; GI: gastrointestinal; CV: cardiovascular; AEs: adverse events.

Supplement 3.8a The description of the reporting quality assessment scores of included studies

Published year	Mean	SD	Minimum	Maximum	Q1	Q2	Q3	Q4	95%CI
1994-2000 (n=7)	68%	0.12	50%	79%	61%	73%	77%	79%	58%-79%
2001-2010 (n=32)	83%	0.06	71%	95%	79%	83%	87%	95%	81%-85%
2011-2018 (n=48)	82%	0.09	61%	96%	75%	81%	88%	96%	79%-84%
Total (n=87)	81%	0.09	50%	96%	75%	83%	87%	96%	79%-83%

SD: Standard deviation, Q1: First quartile, Q2: Second quartile, Q3: Third quartile, Q4: Fourth quartile, 95%CI: 95% confident interval.

Supplement 3.8b The distribution of reporting quality of included studies by published time



Chapter 4: A systematic review and meta-analysis of health state utility values for osteoarthritis-related conditions

4.1 Preface

Chapter 4 presents Study 2 of this thesis which reports the results from a comprehensive systematic review and meta-analysis of OA-related HSUVs. The chapter identifies the areas where the current HSUV evidence is lacking and meta-analyses the OA-related HSUVs for people with OA with different affected joint sites undergoing different treatments, with the overall aim of generating a HSUVs database to guide input choices in future health economic models of OA.

The text in this chapter has been published in *Arthritis Care & Research* (Zhao, Ting, Tania Winzenberg, Barbara de Graaff, Dawn Aitken, Hasnat Ahmad, and Andrew J. Palmer. "A systematic review and meta-analysis of health state utility values for osteoarthritis-related conditions." Arthritis Care & Research (2020) (Supplement 4A).

4.2 Abstract

Background: Health state utility values (HSUVs) are a key input in health economic modelling but HSUVs of people with osteoarthritis (OA)-related conditions have not been systematically reviewed and meta-analysed.

Objective: To systematically review and meta-analyse the HSUVs for people with OA.

Methods: Searches within health economic/biomedical databases were performed to identify eligible studies reporting OA-related HSUVs. Data on study design, participant characteristics, affected OA joint sites, treatment type, HSUV elicitation method, considered health states, and the reported HSUVs were extracted. HSUVs for people with knee, hip and mixed OA in pre- and post-treatment populations were meta-analysed using random effects models.

Results: One-hundred and fifty-one studies were included in the systematic review, and 88 in meta-analyses. Of 151 studies, 56% were conducted in Europe, 75% were in people with knee and/or hip OA and 79% were based on the EQ-5D. The pooled mean (95% confidence interval [CI]) baseline HSUVs for knee OA core interventions, medication, injection and primary

surgery treatments were 0.64 (0.61–0.66), 0.56 (0.45−0.68), 0.58 (0.50–0.66) and 0.52 (0.49–0.55), respectively. These were 0.71 (0.59–0.84) for hip OA core interventions and 0.52 (0.49–0.56) for hip OA primary surgery. For all knee OA treatments and hip OA primary surgery, pooled HSUVs were significantly higher in the post- than the pre- treatment populations.

Conclusion: This study provides a comprehensive summary of OA-related HSUVs and generates a HSUVs database for people with different affected OA joint sites undergoing different treatments to guide HSUV choices in future health economic modelling of OA interventions.

4.3 Introduction

Osteoarthritis (OA) is one of the most common chronic joint diseases. It mostly affects knees, hips and small joints of hands. OA is characterised by joint pain, stiffness, swelling, loss of function and disability; which in turn, negatively impact individuals' health-related quality of life (HRQoL) (1) and incur a substantial socio-economic burden (2, 3). Currently, there is no cure for OA, but many treatments and approaches including lifestyle, medications, injections, and surgery are available to help relieve disease syptoms.

Health state utility value(s) [HSUV(s)] are typically used to reflect HRQoL and to calculate quality-adjusted life years (QALYs) - a preferred measure of clinical effectiveness in cost utility/effectiveness analyses (CUA/CEA) (4) . HSUVs measure the strength of a preference for a particular health state, represented as a number between 0 (death) and 1 (optimal health). Health states worse than death may exist, with negative HSUVs assigned (5). HSUVs can be obtained through several methods (6). Direct methods ask individuals to describe and assess health states and place weights on them, using valuation techniques such as the Standard Gamble (SG), Time Trade-off (TTO) and the rating scales (RS) (6). Indirect methods involve the use of preference-based multi-attribute utility instruments (MAUIs), where patients answer questions relating to multiple dimensions of their current health state and the responses are then scored using a value set obtained from respective general populations. Commonly used MAUIs include the EuroQoL-5-Dimension (EQ-5D) instruments, Health Utility Index (HUI), Short-Form-6-Dimension (SF-6D) and Assessment of Quality of Life (AQoL) instruments (7). Finally, mapping techniques are used to transform non-preferenced-based HRQoL measures into HSUVs.

As the stated preference data for a set of health states for an appropriate population is not always available, HSUVs obtained from the literature are widely used in economic evaluations (4). These HSUV estimates may differ from each other due to several factors, including differences in the utility elicitation techniques, MAUIs, the choice of respondent, sample size and quality of studies (4). With an increasingly growing literature of HSUVs, selection of which values to use in economic evaluations becomes challenging. The correct choice of HSUVs is important to accurately calculate QALYs and other CUA outcomes. To obtain the best estimate for a decision-analytic model from the literature, the methods of identification of the data should be systematic and transparent. To date, there is no systematic review and meta-analysis that summarizes estimates of OA-related HSUVs and evaluates the extent of differences between various sub-groups of patients based on affected OA joint sites, treatments, and utility measurements. Our systematic review and meta-analysis aim to generate a database of OA-related HSUVs to address this.

4.4 Material and Methods

4.4.1 Protocol Registration

The study protocol was registered on 17 April 2019 at PROSPERO (registration number: CRD42019129408; https://www.crd.york.ac.uk/Prospero/). Our systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (8).

4.4.2 Literature search

Based on previous recommendations (9), four databases were searched from their inception up to March 2019: Embase, Health Technology Assessment (HTA) Database, Medline and Scopus. This was supplemented by hand searching the bibliography lists of all included articles and relevant reviews. The search strategy was developed in consultation with co-authors based on the previous literature (10, 11). Supplement 4.1 provided the search strategy used for Embase which was also revised to suit other databases.

4.4.3 Screening criteria

Title/abstract screening and full-text screening were conducted in Covidence (12) (an online systematic review program to manage and facilitate the selection of studies) by two reviewers (TZ and HA) independently based on predefined criteria. Any disagreements were discussed

between the two reviewers, and a third reviewer (AP) was consulted in cases of no consensus. Studies were included if they: 1) involved humans; 2) reported OA related HSUVs estimates (excluding those based on mapping techniques); and 3) were published in English, Chinese or German. Conference abstracts were included when adequate data were available for extraction. We included HSUV studies as long as they were based on OA participants, regardless of the utilised OA diagnosis criteria/definition (i.e., self-reported, clinical or other). If the OA patients were part of a broader study population, we included studies reporting on a cohort with ≥80% OA representation. Health economic modelling studies based on HSUVs reported elsewhere and those based on systematic reviews or meta-analyses were excluded. Review reports, books and case reports were excluded.

4.4.4 Data extraction

A pre-defined Microsoft Excel spreadsheet was piloted to extract data from 20% of studies by the first author (TZ). Adjustments and improvements were made to the initial spreadsheet where necessary, and the improved spreadsheet was then used to extract data independently by TZ and HA. Discrepancies were resolved by consensus, and an additional reviewer (AP) was consulted to reach an agreement in cases of no consensus. The following data were extracted: authors' names, year of publication, study setting, study design (e.g.: trial, observational), sample size, characteristics of the patients (e.g.: age, sex, body mass index [BMI]), affected OA joint sites, treatment type, utility elicitation method, the health states considered, and the reported HSUVs [mean, standard deviation (SD)/standard error (SE), 95% confidence intervals (95% CIs), the median, minimum, maximum, quartile] (Supplement 4.2.1).

4.4.5 Meta-analyses

Based on data availability, the selection of studies for meta-analyses included: 1) studies related to knee, hip and mixed (including a variety of OA patients without specifying their affected OA joint site) OA; and 2) studies of core intervention, medication, intraarticular injection, and primary surgery treatments. We followed OA management guidelines (13) to group the included interventions under one of these four categories of treatment. The core intervention category included: exercise, weight management, and education/programs related to exercise and weight management. Medications included: all drugs used to decrease pain and improve function in patients with OA. Intraarticular injections included corticosteroids, viscosupplements, and blood-derived products. Finally, primary surgery included joint

resurfacing and primary joint replacement. Supplement 4.3 provides the full list of included interventions under each category of treatment for knee, hip and mixed OA from studies included in meta-analyses. Observational studies that did not include delivery of an intervention were excluded from the meta-analysis. HSUVs were summarized by key OA affected joint sites (knee, hip, and mixed OA) for baseline (pre-treatment) and at the most commonly available post-treatment time points (i.e.: 3, 6, 12 and 24 months). When more than one HSUVs study was based on the same data, the study with the highest number of participants was included in the meta-analyses. Sub-group meta-analyses by utility elicitation methods were also conducted, where possible.

The meta-analyses were programmed in STATA (STATA 15.1, StataCorp, College Station, Texas, USA), using the "metan" command that required mean and SD/SE as meta-analytical inputs (14). Therefore, when the mean values and SD/SE were not reported, we used 95% CIs, median, minimum, maximum, first quartile, and third quartile values to estimate these parameters (15-17). HSUVs at baseline in observational studies and in both control and intervention groups of trials were pooled (termed pre-treatment HSUVs). Post-treatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups), for each time points. Heterogeneity among the pooled studies was assessed using the I^2 statistic (where an I^2 >=50% indicated substantial heterogeneity) (17). To account for within-study and between-study heterogeneity, random effects models were estimated.

4.5 Results

4.5.1 Eligible Studies

Initially 7,621 potential references were identified (Figure 4.1). After we removed duplicates (n=4,358), 3,263 were left for title and abstract screening. We excluded 2,593 during title and abstract screening, leaving 670 for the full-text assessment. Of those, 522 were excluded due to not meeting the inclusion criteria (Figure 4.1). Three additional studies identified through reference hand-searching were subsequently included, resulting in a final total of 151 (including 7 abstracts) being included in the systematic review. Eighty-eight of these studies were included in meta-analyses (including 4 conference abstracts).

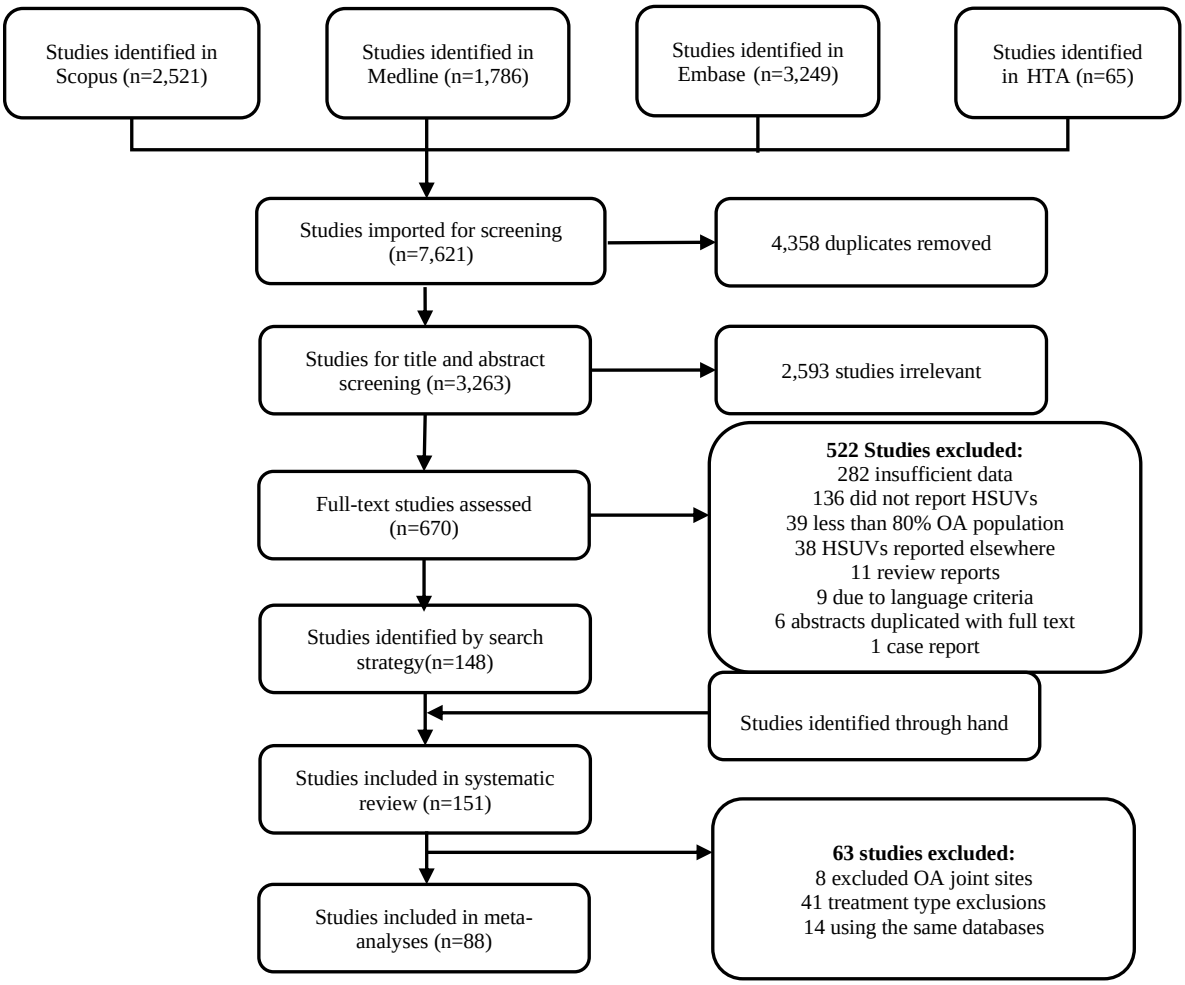


Figure 4.1. Flow chart results of study search based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses methodology.

Note: The exclusions by OA joint sites and treatment type were because of the small numbers of studies in these joint sites and treatments, which meant that meta-analysis was not feasible. Eight exclusions by OA joint sites involved 2 shoulder and 6 hand OA-related studies, 41 exclusions by treatment type involved studies of massage, foot insoles, brace, mud therapy, balneotherapy, spa therapy, revision surgery and observational studies that did not focus on any treatment.

4.5.2 Results of systematic review

Publication date, study setting and study design

The majority (n=131, 87%) of included studies were published after 2010 (Figure 4.2A). More than half (n=86, 57%) were conducted in Europe, followed by Asia (n=20, 13%) and the Americas (n=16, 11%). Four studies focused on Australians with OA, one study was conducted in multiple countries, and 24 studies did not report the study setting (Figure 4.2B). Fifty-eight (38%) included studies were trials, 65 (43%) were observational studies of interventions, and 28 (19%) were observational studies that did not have an intervention component.

Affected OA joint sites

Fifty-nine (39%) studies focused on knee OA and 41 (27%) focused on hip OA. Thirteen (9%) studies focused on both knee and hip OA and reported HSUVs separately. Two (1%) and six (4%) studies were focused on shoulder and hand OA, respectively. Thirty studies (20%) focused on mixed OA (Figure 4.2C).

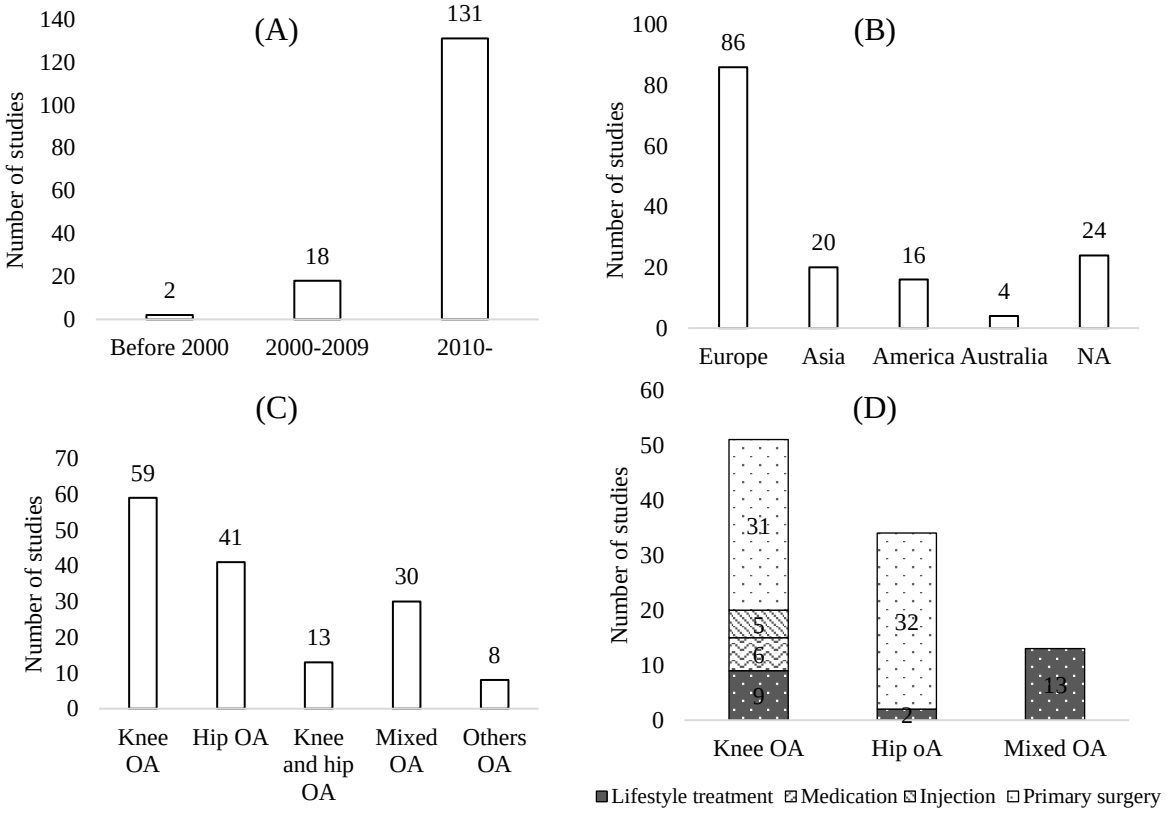


Figure 4.2. The distribution of included studies in systematic review (A) by years of publication; (B) by study setting; (C) by OA joint sites; and (D) the distribution of included studies in meta-analysis by OA joint sites and treatments.
Note: Knee and hip OA group in Figure 4.2C included studies reporting HSUVs for each type separately. Other OA included hand and shoulder OA studies.

Knee OA

Of the 72 knee OA related studies, 10 (14%) focused on core interventions, 6 (8%) and 5 (7%) focused on medication and injection treatments, respectively. Thirty-two (44%) studies focused on surgical treatments. Seven investigated other treatments such as massage, foot insoles, brace, and mud therapy. Twelve (17%) reporting the cross-sectional HSUVs of knee OA did not focus on any specific treatment (Table 4.1).

Hip OA

Of the 54 hip OA related studies, most (n=46, 85%) focused on surgical treatments. Two (4%) studies focused on core interventions, one investigated balneotherapy, and 5 (9%) reporting

the cross-sectional HSUVs of hip OA did not focus on any specific treatment. There were no studies reporting the HSUVs related to hip OA medication and injection treatments (Table 4.1).

Table 4.1 The number and percentage of studies included in systematic review and meta-analyses for each OA affected joint site and treatment

	Studies in systematic review	Studies in meta-analyses
Knee OA (n, %)		
Core intervention	10 (14)	9 (18)
Medication	6 (8)	6 (12)
Injection	5 (7)	5 (10)
Surgery	32 (44)	31 (61)
Other treatments	7 (10)	0 (0)
No treatments	12 (17)	0 (0)
Sub-total	**72**	**51**
Hip OA (n, %)		
Core intervention	2 (4)	2 (6)
Surgery	46 (85)	32 (94)
Other treatments	1 (2)	0 (0)
No treatments	5 (9)	0 (0)
Sub-total	**54**	**34**
Mixed OA (n, %)		
Core intervention	14 (50)	13 (100)
Other treatments	4 (13)	0 (0)
No treatments	12 (40)	0 (0)
Sub-total	**30**	**13**
Other joint sites of OA	8	0
Total	**164***	**98***

*Note: *Thirteen studies reporting knee and hip HSUVs separately have been counted in both the hip and knee OA groups and ten of them were included in the meta-analyses. Mixed OA included a variety of OA patients without specifying their OA type, Other joint site of OA included shoulder and hand OA studies. OA=osteoarthritis.*

Other types of OA

Two shoulder OA-related studies focused on surgical treatments. Among six hand OA-related studies, two reported the cross-sectional HSUVs of hand OA populations, and one study each focused on spa, mud, a core intervention and surgery treatment.

Mixed OA type

Of the 30 mixed OA related studies, 14 (50%) focused on core interventions, and 12 (40%) reported the cross-sectional HSUVs of an OA population without specifying any treatment type. Two studies focused on surgical treatments, one focused on medication and one focused on spa therapy (Table 4.1).

Health State Utility measures

Nine HSUV measures were used in the included studies, with most (n=120, 79%) studies using the EQ-5D, followed by the SF-6D (n=12, 8%), HUI2/3 (n=4, 3%), and Quality of Well-Being (QWB) (n=3, 2%). One study each used the AQoL-6D, and the 15D. Ten studies (7%) included more than one measure including the Paper Adaptive Test (PAT-5D-QoL), SG, and RS.

4.5.3 Results of meta-analysis

Studies included in meta-analyses

Fifty-one knee OA related studies (Figure 4.2D) qualified for meta-analyses. Nine, 6, 5 and 31 related to core interventions, medications, injections, and primary surgery, respectively (Supplements 4.2.2-4.2.5). Thirty-four hip OA related studies (Figure 4.2D) qualified for meta-analyses. Two and 32 related to core interventions and primary surgery, respectively (Supplements 4.2.6-4.2.7). Thirteen studies for mixed OA core interventions qualified for meta-analyses (Figure 4.2D, Supplement 4.2.8).

The post-treatment time points included in meta-analyses varied between different OA joint sites and treatments based on data availability (Table 4.2).

HSUVs of Knee OA

The pooled mean baseline (pre-treatment) HSUV of knee OA core interventions was 0.64 (number of HSUVs pooled [n]=19, 95% CI:0.61−0.66, I^2=99%). The pooled HSUVs measuring this at 3 months in post-intervention populations were higher (0.73, n=6, 95% CI:0.70−0.76, I^2= 91%). The pooled 6-month and 1-year HSUVs did not differ significantly from baseline (0.65, n=4, 95% CI:0.60−0.71, I^2= 97% at 6-month; 0.71, n=5, 95% CI:0.64−0.79, I^2=1 at 1-year, respectively). In the subgroup analyses, there were significant difference in HSUVs estimates between different MUIs at each time point. See Table 4.2 and Supplement Figure 4.1.

The pooled mean HSUV for knee OA medication treatment was significantly different at baseline (0.56, n=9, 95% CI: 0.45−0.68, I^2=1) than at 3-months follow-up (0.75, n=3, 95% CI:0.70−0.80, I^2=87%). All knee medication related HSUVs were based on the EQ-5D. See Table 4.2 and Supplement Figure 4.2.

The pooled HSUVs for knee OA (intraarticular) injections were similar at baseline (0.58, n=7, 95% CI:0.50−0.66, I^2=94%) and 1-year post-treatment (0.63, n=1, 95% CI:0.59−0.67). The

baseline HSUVs estimates significantly differed between EQ-5D and HUI3 measures. See Table 4.2, and Supplement Figure 4.3.

For knee OA primary surgeries, the pooled mean HSUV was 0.52 (n=55, 95% CI:0.49−0.55, I^2=99.7%) at baseline. A significant difference was found between HSUVs of baseline and various post-surgery time points: 6 months (0.71, n=21, 95% CI:0.69−0.74, I^2=95%); 1-year (0.77, n=18, 95% CI:0.73−0.81, I^2=99%); and 2-years (0.74, n=17, 95% CI:0.71−0.78, I^2=99%). Significant differences existed between different MAUIs at each time point. See Table 4.2 and Supplement Figure 4.4.

HSUVs of Hip OA

Only 2 studies focused on hip OA core interventions, HSUVs did not differ significantly between the baseline (0.71, n=3, 95% CI:0.59−0.84, I^2=99%), 3-months (0.72, n=2, 95% CI:0.59−0.84, I^2=98%), or 1-year (0.72, n=2, 95% CI:0.58−0.85, I^2=98%) post-interventions. All HSUVs were based on the EQ-5D. See Table 4.2 and Supplement Figure 4.5.

For hip OA primary surgery treatments, there was a significant difference between the pooled mean HSUVs of baseline (0.52, n=46, 95% CI:0.49−0.56, I^2=1) and post-surgery periods: 6 months (0.79, n=9, 95% CI:0.76−0.82, I^2=94%); 1-year (0.83, n=22, 95% CI:0.80−0.85, I^2=99%) and 2-years (0.84, n=11, 95% CI:0.80−0.87, I^2=98%). Significant differences existed between different MAUIs at each time point. See Table 4.2 and Supplement Figure 4.6.

HSUVs of Mixed OA

For mixed OA core interventions, there was a significant difference between the pooled mean HSUVs of baseline (0.61, n=27, 95% CI:0.59−0.64, I^2=99%) and 3-months post intervention (0.71, n=10, 95% CI:0.68−0.73, I^2=97%), and 1-year post intervention (0.69, n=12, 95% CI: 0.66−0.71, I^2=98%). The same trend was found for EQ-5D HSUVs but not for SF-6D. See Table 4.2 and Supplement Figure 4.7.

Table 4.2 The number of pooled HSUVs and representing population in meta-analyses and the pooled mean HSUVs

	Overall			EQ-5D			SF-6D		
	Number of HSUVs	Representing population	Mean (95%CI)	Number of HSUVs	Representing population	Mean (95%CI)	Number of HSUVs	Representing population	Mean (95%CI)
Knee OA									
Core intervention									
Baseline	19	31349	0.64 (0.61, 0.66)	12	30869	0.60 (0.57, 0.62)	2	119	0.65 (0.55, 0.75)
3-mo	6	30,380	0.73 (0.70, 0.76)	4	30,233	0.70 (0.69, 0.72)	0	0	--
6-mo	4	542	0.65 (0.60, 0.71)	4	542	0.65 (0.60, 0.71)	0	0	--
1-year	5	20,549	0.71 (0.64, 0.79)	3	20,402	0.64 (0.56, 0.72)	0	0	--
Medication									
Baseline	9	3,749	0.56 (0.45, 0.68)	9	3,749	0.56 (0.45, 0.68)	0	0	--
3-mo	3	249	0.75 (0.70, 0.80)	3	249	0.75 (0.70, 0.80)	0	0	--
Injection									
Baseline	7	473	0.58 (0.50, 0.66)	5	224	0.63 (0.56, 0.70)	0	0	--
1-year	1	122	0.63 (0.59, 0.67)	0	0		0	0	--
Primary surgery									
Baseline	55	53,434	0.52 (0.49, 0.55)	43	45,434	0.49 (0.44, 0.53)	10	7,797	0.62 (0.60, 0.64)
6-mo	21	4,260	0.71 (0.69, 0.74)	14	3,378	0.72 (0.70, 0.73)	5	719	0.71 (0.68, 0.75)
1-year	18	3,790	0.77 (0.73, 0.81)	12	2,179	0.78 (0.73, 0.84)	4	1,456	0.75 (0.72, 0.79)
2-year	17	15,160	0.74 (0.71, 0.78)	11	8,872	0.77 (0.73, 0.81)	5	6,270	0.71 (0.66, 0.76)
Hip OA									
Core intervention									
Baseline	3	13,773	0.71 (0.59, 0.84)	3	13,773	0.71 (0.59, 0.84)	0		--
3-mo	2	13,671	0.72 (0.59, 0.84)	2	13,671	0.72 (0.59, 0.84)	0		--
1-year	2	8,421	0.72 (0.58, 0.85)	2	8,421	0.72 (0.58, 0.85)	0		--

Primary surgery									
Baseline	46	59,846	0.52 (0.49, 0.56)	34	52,671	0.50 (0.46, 0.54)	6	6,791	0.58 (0.56, 0.60)
6-mo	9	3,922	0.79 (0.76, 0.82)	6	3,727	0.80 (0.77, 0.83)	0	0	--
1-year	22	42,788	0.83 (0.80, 0.85)	20	42,468	0.83 (0.80, 0.85)	1	224	0.80 (0.78, 0.82)
2-year	11	16,732	0.84 (0.80, 0.87)	7	10,228	0.88 (0.85, 0.90)	4	6,504	0.78 (0.75, 0.81)
Mixed OA									
Core intervention									
Baseline	27	9,644	0.61 (0.59, 0.64)	18	5,672	0.58 (0.53, 0.62)	5	3,576	0.69 (0.66, 0.72)
3-mo	10	6,926	0.71 (0.68, 0.73)	6	3,587	0.67 (0.63, 0.71)	2	3,192	0.73 (0.65, 0.82)
1-year	12	7,305	0.69 (0.66, 0.71)	6	3,819	0.65 (0.60, 0.71)	4	3,339	0.71 (0.65, 0.77)

Note: Pooling at all time points includes observational data on intervention and at baseline includes both control and active treatment groups from trials, but at follow-up includes only data from active treatment groups from trials. OA=osteoarthritis, HSUVs=health state utility values, 95%CI=95% confidence interval.

4.6 Discussion

This is the first wide-ranging systematic review of OA-related HSUVs and meta-analyses on HSUVs for people with different OA affected joint sites before and after various treatments. Our systematic review identified important areas where the current evidence is lacking, namely under-represented geographical locations/ethnicities, affected OA joint sites, treatment options and HSUVs based on more sensitive MAUIs. Our meta-analyses provide a HSUV database for alternative pre- and post- OA treatments that could offer a variety of HSUV inputs for future cost utility models of OA-related conditions. HSUVs associated with four key treatment categories (core interventions, medication, injection, and surgery) often differed, as expected, pre- and post-treatment. Furthermore, we found significant inter-MAUI differences in the mean HSUVs, which is as expected from alternative descriptive systems and utility algorithms. As such this review provides important information that could be used by health economists and policy makers to determine the cost-effectiveness of various OA treatments and long-term disease outcomes using modelling techniques.

Our systematic review identified numerous gaps in the data on OA-related HSUVs, including geographical locations/ethnicities, affected OA joint sites, treatment options and HSUVs based on more sensitive MAUIs. We found that more than half (57%) of included studies were conducted in Europe, and none in Africa. Because HSUVs should ideally be based on local populations preferences, the generalizability of our results to under-represented (e.g.: African and Asian) populations may therefore be limited. Seventy-six percent of included studies focused on knee and hip OA, while other joint sites (e.g.: shoulder and hand) attracted limited attention. Whilst these results align well with the higher clinical impact, prevalence and societal burden of knee and hip OA (18-20), the increasing prevalence and disease burden of hand and shoulder OA as a result of population aging (21, 22) mandates further primary studies investigating the HSUVs of these joint sites.

The HSUVs that we have meta-analysed differed as expected between alternative OA joint sites, treatments, HSUVs measures and time points. We found a mean HSUVs difference of +0.09 units in knee OA patients using core interventions between baseline and 3-months post-intervention, and this difference exceeds the minimal clinically important difference for all the MAUIs reported in previous studies (from +0.04 units [EQ-5D] to 0.08 units [AQoL-8D]) (23-27). Our findings are consistent with the RCT evidence showing the short-lived effects of knee OA core interventions (28, 29). Other possible explanations include: the limited number of

core intervention studies with a follow-up period of greater than 3 months (and hence, wider 95% CIs for our 6 months and 1-year post-treatment HSUVs), and a likely reduction in the core intervention adherence in the long-term (30, 31).

Most (83.3%) studies of knee OA medication treatments had relatively shorter follow up periods (3-months), with only one study with a follow-up period of >3months. Consistent with RCT evidence of effectiveness of medication treatments (32, 33), the pooled HSUV of studies with follow-up at 3 months post-medication treatment was significantly higher than the pooled HSUV of studies with baseline measures. As we did not have enough data on long-term HSUVs in patients using OA medications, we leave this on the agenda for future research when long-term data becomes available. We found similar HSUVs at baseline and 1-year follow-up for knee OA injection treatments. However, these results should be carefully interpreted and used in economic modelling as being derived from only a limited number of studies (n=5 at baseline and n=1 at 1-year follow-up). HSUVs of knee OA patients recorded the largest difference (+0.25 units) between baseline and 1-year post- primary surgery and it remained relatively stable to 2-years post- primary surgery. These findings are once again consistent with the previous evidence of the effectiveness of knee surgery, suggesting that HSUVs record a significant improvement within 1-year of knee surgery, and this change in HSUVs is sustained for years (34).

Surgery was the commonest treatment in hip OA HSUV studies (85%). HSUVs in patients with primary hip OA surgery were significantly higher at 6-months post-surgery than at baseline and remained improved over the long term. The difference between pooled HSUV before and after surgery over 1-year was smaller in knee OA primary surgery (+0.25 units) than hip OA (+0.31 units). These findings align well with previous research (35) advocating a relatively higher efficacy of hip OA joint surgery. Only 2 studies (both based on the EQ-5D) investigated HSUVs in hip OA core intervention patients, which aligned well with the previous findings of the dearth of studies measuring the HSUVs in patients using hip OA core interventions (36, 37). No studies on hip OA medication and injection treatments were identified in our review as expected (38, 39), hence, no meta-analysis for these treatments was possible. We recommend future studies to investigate HSUVs in patients using medications and injections, subject to the availability of better long-term observed data.

The HSUVs for mixed OA core interventions showed the same trend observed for knee OA, with a significant difference (+0.10 units) between baseline and 3-months post-intervention

HSUVs. This aligns with the existing findings of short-term benefits associated with OA core interventions (29). A small number of studies of medication treatment (n=1) for mixed OA did not allow us to generate HSUV estimates 'before' and 'after' medication treatments for use in health-economic modelling. Future primary HSUVs studies in this area should therefore be imperative in bridging this evidence gap.

The EQ-5D was the most commonly used (79%) MAUI in the included studies, with little to no representation from other more detailed MAUIs (e.g.: AQoL-8D) that can more fully capture and assess the complex physical and psychosocial health aspects of OA patients (23, 40). Our MAUI-specific (sub-group) analysis revealed significant differences between HSUVs based on alternative MAUIs (EQ-5D and SF-6D, for example), which is as expected from the MAUIs that are far from identical in terms of their descriptive systems and measurement scales (41). As the key objective of our review was not to explore the extent of agreement between alternative MAUIs, we leave the head-to-head comparison of HSUVs obtained through alternative MAUIs on the agenda for future research. Moreover, there is no consensus on the choice of MAUI to be used in measuring HSUVs of OA patients (41, 42). Many health care decision-making bodies around the world, such as the National Institute for Health and Care Excellence, require the use of generic MAUIs (e.g., the EQ-5D, SF-6D, HUI3 and the AQoL-8D). While three generic MAUIs (i.e., EQ-5D, HUI, SF-6D) are each cited in at least 10 country guidelines (43), the EQ-5D is most widely used HSU elicitation instrument in all disease areas including OA. However, this instrument does not fully capture and assess all aspects of people with OA's HRQoL, particularly those related to their psychosocial health. Whilst the commonly used generic MAUIs are far from identical, they are equipped with alternative ability and preferential sensitivity to capturing and assessing the complex physical and psychosocial health aspects. Importantly, in health economic research there is an important distinction between generic measures of HRQoL and those that are more specific to a condition, dimension or population. While the use of preference-based condition-specific measures is becoming popular and comes with its unique strengths (44), our suggestion being health economists is to identify and use a generic MAUI that that can capture and assess all important aspects of OA patients' quality of life to improve healthcare resource allocation decision making process, enabling consistency and inter-disease area HRQoL comparability. Future research should therefore endeavour to identify and use generic MAUI(s) with comprehensive descriptive system to capture and assess all important (physical and psychological) health

aspects of OA patients' HRQoL to improve healthcare resource allocation decision making process, enabling consistency and inter-disease area HRQoL comparability.

When the baseline HSUVs for various treatments were compared, the mean baseline HSUVs for knee and hip OA patients using core interventions were significantly higher than those using surgery treatments, which is likely to be due to the specified selection criteria for RCTs. Due to the recommended stepwise approach for OA treatments (45), patients are more likely to receive core interventions at earlier stages of their OA (with better HRQoL) and surgery treatments at more severe stages of their OA (with relatively worse HRQoL), which can also explain this pattern. This result reinforces the need to use different HSUVs in modelling for treatments used at different disease stages.

The strength of this study is that this is the first comprehensive review and meta-analysis of all types of OA-related preference-based HSUVs by OA affected joint sites, OA treatments, and utility elicitation methods. It provided a HSUV database for alternative pre- and post- OA treatments that could offer a variety of HSUV inputs to future cost utility models of OA-related conditions and identified important areas where there are evidence gaps in these estimates to inform future research needs. Our study has several limitations. It is important that the differences in HSUVs at different time points are not interpreted as true pre-post change or direct evidence of intervention effectiveness as the data do not examine differences in change in HSUVs between controls and intervention groups over time and the data included in pooling at each time point comes from different studies. Heterogeneity of the included studies due to the differences in terms of their study design, settings and HSUV elicitation techniques can affect the interpretation of generated HSUVs. For example, the studies included in our systematic review and meta-analyses are both RCTs and observational studies, so our estimates of state specific HSUVs may suffer from some degree of bias as what economic models require are estimates of how HSUVs change with treatment and its associated impact on patients' HRQoL (44). However, the results from meta-analyses are recommended to be used as inputs in health economic models compared with those from a single RCTs (usually derived from constrained samples due to restricted selection criteria) as meta-analytic HSUV estimates lead to improved estimation quality and consistency (46). While we acknowledge that the HSUVs based on RCT-based meta-analyses may be closer to the reality, the estimates generated in our research are the best available data source for economic modelling to date. Future primary RCT studies assessing the HSU impact of various treatments for people with different affected OA joint sites are therefore recommended, so a database of more precise

HSU inputs of various treatments can be generated to obtain the direct estimates of intervention effects. Whilst we have conducted subgroup analyses where possible to highlight some possible sources of heterogeneity, we had limited capacity to explain and account for all sources of heterogeneity. The random effects model in our meta-analyses aims to account for heterogeneity but may have consequences for the precision of model estimates (47). Therefore, in modelling, as well as the pooled mean sensitivity analyses, consideration of the potential imprecision of our estimates is important. A further limitation is that due to the paucity of available studies it was not possible to conduct meta-analyses for all treatments of hip, knee and other OA joint sites or group the treatment types in a more detailed way or to perform meta-regression to account for more than one potential effect modifying variable at a time. Quite a few potentially eligible CUA/CEA reports were not able to be included as they did not adequately report the required (pre-and/or-post-treatment) HSUVs (48-52), despite clear reporting guidelines that recommend these are reported (53, 54). We recommend that future CUA/CEA studies refer to these guidelines to help improve the availability of this important data. Also, the exploration of long-term HSUVs of patients using different OA treatments was mostly not possible. Finally, due to the paucity of data, we could not generate the estimates of HSUVs associated with alternative therapy adherence levels and medication adverse event types.

Conclusions

Our systematic review found that studies of OA related HSUVs are of wide variety, and differ from each other in terms of their setting, design, focused OA joint sites, utility measurement technique, generalisability and others. The HSUVs that we have generated will be useful in conducting future health economic modelling for people suffering from various OA-related conditions. Our results should however be interpreted with caution as being derived from a relatively small number of heterogeneous studies. More research is needed to investigate changes in HSUVs of OA patients for longer follow-up periods.

Authors' contributions

TZ, BdeG, HA, DA, TW and AP conceived and designed the methods, and critically reviewed/edited the manuscript. TZ and HA conducted literature search, screening and data extraction. TZ conducted the meta-analyses and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank Mr Petr Otahal (Statistical Officer, Menzies Institute for Medical Research, University of Tasmania) for his help in conducting the meta-analyses.

4.7 Reference

1. Wilson R, Blakely T, Abbott JH. Radiographic knee osteoarthritis impacts multiple dimensions of health-related quality of life: data from the Osteoarthritis Initiative. Rheumatology (Oxford). 2018;57(5):891-9.

2. Abbott JH, Usiskin IM, Wilson R, Hansen P, Losina E. The quality-of-life burden of knee osteoarthritis in New Zealand adults: A model-based evaluation. PLoS One. 2017;12(10):e0185676.

3. Henderson JV, Harrison CM, Britt HC, Bayram CF, Miller GC. Prevalence, causes, severity, impact, and management of chronic pain in Australian general practice patients. Pain Med. 2013;14(9):1346-61.

4. Brazier J, Green C. A systematic review of health state utility values for osteoporosis-related conditions. Osteoporosis International. 2002;13(10):768-76.

5. Ahmad H, Taylor BV, van der Mei I, Colman S, O'Leary BA, Breslin M, et al. The impact of multiple sclerosis severity on health state utility values: evidence from Australia. Multiple Sclerosis Journal. 2017;23(8):1157-66.

6. Meregaglia M, Cairns J. A systematic literature review of health state utility values in head and neck cancer. Health Qual Life Outcomes. 2017;15(1):174.

7. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Annals of medicine. 2001;33(5):358-70.

8. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.

9. Arber M, Glanville J, Isojarvi J, Baragula E, Edwards M, Shaw A, et al. Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care. 2018;34(6):547-54.

10. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. Value in Health. 2013;16(4):686-95.

11. Arber M, Garcia S, Veale T, Edwards M, Shaw A, Glanville JM. Performance of Ovid Medline search filters to identify health state utility studies. Int J Technol Assess Health Care. 2017;33(4):472-80.

12. Better systematic reivew management C, accessed: https://www.covidence.org/home, dated 29 January 2019.

13. The Royal Australian College of General Practitioners. Guideline for the management of knee and hip osteoarthritis. 2nd edn. East Melbourne, Vic: RACGP, 2018.

14. Harris RJ, Deeks JJ, Altman DG, Bradburn MJ, Harbord RM, Sterne JA. Metan: fixed- and random-effects meta-analysis. The Stata Journal. 2008;8(1):3-28.

15. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014;14:135.

16. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res. 2018;27(6):1785-805.

17. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons; 2011.

18. Osteoarthritis Research Society International. Osteoarthritis: a serious disease. https://www.oarsi.org/sites/default/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf.

19. Mahir L, Belhaj K, Zahi S, Azanmasso H, Lmidmani F, El Fatimi A. Impact of knee osteoarthritis on the quality of life. Annals of physical and rehabilitation medicine. 2016;59:e159.

20. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clinical and Experimental Research. 2005;17(4):255-63.

21. Millett PJ, Gobezie R, Boykin RE. Shoulder osteoarthritis: diagnosis and management. Am Fam Physician. 2008;78(5):605-11.

22. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Annals of the rheumatic diseases. 2011;70(9):1581-6.

23. Campbell JA, Palmer AJ, Venn A, Sharman M, Otahal P, Neil A. A Head-to-Head Comparison of the EQ-5D-5L and AQoL-8D Multi-Attribute Utility Instruments in Patients Who Have Previously Undergone Bariatric Surgery. The Patient - Patient-Centered Outcomes Research. 2016;9(4):311-22.

24. Campbell JA, Hensher M, Neil A, Venn A, Wilkinson S, Palmer AJ. An exploratory study of long-term publicly waitlisted bariatric surgery patients' quality of life before and 1 year after bariatric surgery, and considerations for healthcare planners. PharmacoEconomics-open. 2018;2(1):63-76.

25. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. Journal of clinical epidemiology. 2003;56(5):395-407.

26. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Australian and New Zealand journal of public health. 2005;29(2):136-42.

27. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health and quality of life outcomes. 2003;1(1):4.

28. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. Br J Sports Med. 2015;49(24):1554-7.

29. Pisters MF, Veenhof C, Van Meeteren NL, Ostelo RW, De Bakker DH, Schellevis FG, et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. Arthritis Care & Research. 2007;57(7):1245-53.

30. van Gool CH, Penninx BWJH, Kempen GIJM, Rejeski WJ, Miller GD, van Eijk JTM, et al. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. Arthritis Care & Research. 2005;53(1):24-32.

31. Marks R, Allegrante JP. Chronic Osteoarthritis and Adherence to Exercise: A Review of the Literature. Journal of Aging and Physical Activity. 2005;13(4):434-60.

32. Lin J, Zhang W, Jones A, Doherty M. Efficacy of topical non-steroidal anti-inflammatory drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials. Bmj. 2004;329(7461):324.

33. Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV, et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology. 2000;39(10):1095-101.

34. Nilsdotter A-K, Toksvig-Larsen S, Roos E. A 5 year prospective study of patient-relevant outcomes after total knee replacement. Osteoarthritis and cartilage. 2009;17(5):601-6.

35. Bachmeier CJ, March LM, Cross MJ, Lapsley HM, Tribe KL, Courtenay BG, et al. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. Osteoarthritis Cartilage. 2001;9(2):137-46.

36. Murphy SL, Robinson-Lane SG, Niemiec SLS. Knee and hip osteoarthritis management: a review of current and emerging non-pharmacological approaches. Current Treatment Options in Rheumatology. 2016;2(4):296-311.

37. Lee PYF RS, Chandrappa MH, Othman A, Jury C, et al. . Modern Non-Pharmacological and Non-Surgical Treatments for Hip Pain. J Arthritis 2018;7: 262(DOI: 10.4172/2167-7921.1000262).

38. Myers J, Wielage RC, Han B, Price K, Gahn J, Paget MA, et al. The efficacy of duloxetine, non-steroidal anti-inflammatory drugs, and opioids in osteoarthritis: a systematic literature review and meta-analysis. BMC Musculoskelet Disord. 2014;15:76.

39. Ayral X. Injections in the treatment of osteoarthritis. Best practice & research Clinical rheumatology. 2001;15(4):609-26.

40. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology. 1999;38(9):807-13.

41. Zhang F, Yang Y, Huang T, Zhang Y, Zhao L, Li S. Is there a difference between EQ-5D and SF-6D in the clinical setting? a comparative study on the quality of life measured by AIMS2-SF, EQ-5D and SF-6D scales for osteoarthritis patients. Int J Rheum Dis. 2018;21(6):1185-92.

42. NIfHaC E. Guide to the methods of technology appraisal 2013. London: NICE. 2013.

43. Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. The European Journal of Health Economics. 2020;21:1245-57.

44. Brazier J. Valuing health states for use in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):769-79.

45. Sinusas K. Osteoarthritis: diagnosis and treatment. American family physician. 2012;85(1).

46. Karnon J. Heath state utility values for cost-effectiveness models. Springer; 2017.

47. Borenstein M, Hedges L, Rothstein H. Meta-analysis: Fixed effect vs. random effects. Meta-analysis com. 2007.

48. Abbott J, Wilson R, Pinto D, Chapple C, Wright A, Team MT. Incremental clinical effectiveness and cost effectiveness of providing supervised physiotherapy in addition to usual medical care in patients with osteoarthritis of the hip or knee: 2-year results of the MOA randomised controlled trial. Osteoarthritis and cartilage. 2019;27(3):424-34.

49. Bulthuis Y, Mohammad S, Braakman-Jansen LM, Drossaers-Bakker KW, van de Laar MA. Cost-effectiveness of intensive exercise therapy directly following hospital discharge in patients with arthritis: Results of a randomized controlled clinical trial. Arthritis care & research. 2008;59(2):247-54.

50. Tan SS, Teirlinck CH, Dekker J, Goossens LM, Bohnen AM, Verhaar JA, et al. Cost-utility of exercise therapy in patients with hip osteoarthritis in primary care. Osteoarthritis and cartilage. 2016;24(4):581-8.

51. Marra CA, Grubisic M, Cibere J, Grindrod KA, Woolcott JC, Gastonguay L, et al. Cost‐Utility Analysis of a Multidisciplinary Strategy to Manage Osteoarthritis of the Knee: Economic Evaluation of a Cluster Randomized Controlled Trial Study. Arthritis care & research. 2014;66(6):810-6.

52. Buitendyk M, Codding C, Torrance GW. Economic evaluation of controlled-release oxycodone vs oxycodone-acetaminophen for osteoarthritis pain of the hip or knee. The American journal of managed care. 2006;12:205-14.

53. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. Bmj. 2013;346:f1049.

54. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report. Value in Health. 2015;18(2):161-72.

4.8 Supplements

Supplement 4A: the publication of chapter 4 "A systematic review and meta-analysis of health state utility values for osteoarthritis-related conditions"

MS. TING ZHAO (Orcid ID : 0000-0002-2145-1930)

MISS DAWN AITKEN (Orcid ID : 0000-0001-5685-7634)

Article type : Original Article

Running head: Systematic review and meta-analysis of OA-related HSUVs

A systematic review and meta-analysis of health state utility values for osteoarthritis-related conditions

Ting Zhao[1], M.S.; Tania Winzenberg[1], PhD, Professor; Barbara de Graaff[1], PhD, Research fellow; Dawn Aitken[1], PhD, Research fellow; Hasnat Ahmad[1], MPhil; Andrew J. Palmer[1,2], MBBS, Professor

1. *Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia*
2. *Centre for Health Policy, School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia*

Correspondence to: 1. Andrew J. Palmer, Menzies Research Institute Tasmania, University of Tasmania, Medical Science 1 Building, 17 Liverpool St, Hobart, Tasmania 7000, Australia, E-mail: Andrew.palmer@utas.edu.au, Tel: +61 3 6226 7729; 2. Ting Zhao, Menzies Research Institute Tasmania, University of Tasmania, Medical Science 1 Building, 17 Liverpool St, Hobart, Tasmania 7000, Australia, E-mail: ting.zhao@utas.edu.au.

No financial support for this study, all authors declare that they have no competing interests.

Word count for the manuscript: 3865

ABSTRACT

Background: Health state utility values (HSUVs) are a key input in health economic modelling but HSUVs of people with osteoarthritis (OA)-related conditions have not been systematically reviewed and meta-analysed.

Objective: To systematically review and meta-analyse the HSUVs for people with OA.

Methods: Searches within health economic/biomedical databases were performed to identify eligible studies reporting OA-related HSUVs. Data on study design, participant characteristics, affected OA joint sites, treatment type, HSUV elicitation method, considered health states, and the reported HSUVs were extracted. HSUVs for people with knee, hip and mixed OA in pre- and post-treatment populations were meta-analysed using random effects models.

Results: One-hundred and fifty-one studies were included in the systematic review, and 88 in meta-analyses. Of 151 studies, 56% were conducted in Europe, 75% were in people with knee and/or hip OA and 79% were based on the EQ-5D. The pooled mean (95% confidence interval [CI]) baseline HSUVs for knee OA core interventions, medication, injection and primary surgery treatments were 0.64 (0.61–0.66), 0.56 (0.45−0.68), 0.58 (0.50–0.66) and 0.52 (0.49–0.55), respectively. These were 0.71 (0.59–0.84) for hip OA core interventions and 0.52 (0.49–0.56) for hip OA primary surgery. For all knee OA treatments and hip OA primary surgery, pooled HSUVs were significantly higher in the post- than the pre- treatment populations.

Conclusion: This study provides a comprehensive summary of OA-related HSUVs and generates a HSUVs database for people with different affected OA joint sites undergoing different treatments to guide HSUV choices in future health economic modelling of OA interventions.

Key Words: osteoarthritis; health-economics; HSUVs, HRQoL, Quality of Life

Significance and Innovations

- This is the first study comprehensively reviewing OA-related HSUVs.
- It identified important areas where the current evidence is lacking, including under-represented geographical locations/ethnicities, OA joint sites, treatment options and MAUIs.
- This study is first to meta-analyse the OA-related HSUVs for different affected joint sites before and after various treatments.
- It generates a HSUVs database for OA patients with different affected joint sites undergoing different treatments to guide HSUVs choices in future health economic modelling of OA interventions.

Osteoarthritis (OA) is one of the most common chronic joint diseases. It mostly affects knees, hips and small joints of hands. OA is characterised by joint pain, stiffness, swelling, loss of function and disability; which in turn, negatively impact individuals' health-related quality of life (HRQoL) (1) and incur a substantial socio-economic burden (2, 3). Currently, there is no cure for OA, but many treatments and approaches including lifestyle, medications, injections, and surgery are available to help relieve disease syptoms.

Health state utility value(s) [HSUV(s)] are typically used to reflect HRQoL and to calculate quality-adjusted life years (QALYs) - a preferred measure of clinical effectiveness in cost utility/effectiveness analyses (CUA/CEA) (4) . HSUVs measure the strength of a preference for a particular health state, represented as a number between 0 (death) and 1 (optimal health). Health states worse than death may exist, with negative HSUVs assigned (5). HSUVs can be obtained through several methods (6). Direct methods ask individuals to describe and assess health states and place weights on them, using valuation techniques such as the Standard Gamble (SG), Time Trade-off (TTO) and the rating scales (RS) (6). Indirect methods involve the use of preference-based multi-attribute utility instruments (MAUIs), where patients answer questions relating to multiple dimensions of their current health state and the responses are then scored using a value set obtained from respective general populations. Commonly used MAUIs include the EuroQoL-5-Dimension (EQ-5D) instruments, Health Utility Index (HUI), Short-Form-6-Dimension (SF-6D) and Assessment of Quality of Life (AQoL) instruments (7). Finally, mapping techniques are used to transform non-preferenced-based HRQoL measures into HSUVs.

As the stated preference data for a set of health states for an appropriate population is not always available, HSUVs obtained from the literature are widely used in economic evaluations (4). These HSUV estimates may differ from each other due to several factors, including differences in the utility elicitation techniques, MAUIs, the choice of respondent, sample size and quality of studies (4). With an increasingly growing literature of HSUVs, selection of which values to use in economic evaluations becomes challenging. The correct choice of HSUVs is important to accurately calculate QALYs and other CUA outcomes. To obtain the best estimate for a decision-analytic model from the literature, the methods of identification of the data should be systematic and transparent. To date, there is no systematic review and meta-analysis that summarizes estimates of OA-related HSUVs and evaluates the extent of differences between various sub-groups of patients based on affected OA joint sites, treatments, and utility measurements. Our

systematic review and meta-analysis aim to generate a database of OA-related HSUVs to address this.

Material and Methods

Protocol Registration

The study protocol was registered on 17 April 2019 at PROSPERO (registration number: CRD42019129408; https://www.crd.york.ac.uk/Prospero/). Our systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (8).

Literature search

Based on previous recommendations (9), four databases were searched from their inception up to March 2019: Embase, Health Technology Assessment (HTA) Database, Medline and Scopus. This was supplemented by hand searching the bibliography lists of all included articles and relevant reviews. The search strategy was developed in consultation with co-authors based on the previous literature (10, 11). Supplement 1 provided the search strategy used for Embase which was also revised to suit other databases.

Screening criteria

Title/abstract screening and full-text screening were conducted in Covidence (12) (an online systematic review program to manage and facilitate the selection of studies) by two reviewers (TZ and HA) independently based on predefined criteria. Any disagreements were discussed between the two reviewers, and a third reviewer (AP) was consulted in cases of no consensus. Studies were included if they: 1) involved humans; 2) reported OA related HSUVs estimates (excluding those based on mapping techniques); and 3) were published in English, Chinese or German. Conference abstracts were included when adequate data were available for extraction. If the OA patients were part of a broader study population, we included studies reporting on a cohort with ≥80% OA representation. Health economic modelling studies based on HSUVs reported elsewhere and those based on systematic reviews or meta-analyses were excluded. Review reports, books and case reports were excluded.

Data extraction

A pre-defined Microsoft Excel spreadsheet was piloted to extract data from 20% of studies by the first author (TZ). Adjustments and improvements were made to the initial spreadsheet where necessary, and the improved spreadsheet was then used to extract data independently by TZ and HA. Discrepancies were resolved by consensus, and an additional reviewer (AP) was consulted to reach an agreement in cases of no consensus. The following data were extracted: authors' names, year of publication, study setting, study design (e.g.: trial, observational), sample size, characteristics of the patients (e.g.: age, sex , body mass index [BMI]), affected OA joint sites, treatment type, utility elicitation method, the health states considered, and the reported HSUVs [mean, standard deviation (SD)/standard error (SE), 95% confidence intervals (95% CIs), the median, minimum, maximum, quartile] (Supplement 2.1).

Meta-analyses

Based on data availability, the selection of studies for meta-analyses included: 1) studies related to knee, hip and mixed (including a variety of OA patients without specifying their affected OA joint site) OA; and 2) studies of core intervention, medication, intraarticular injection, and primary surgery treatments. We followed OA management guidelines (13) to group the included interventions under one of these four categories of treatment. The core intervention category included: exercise, weight management, and education/programs related to exercise and weight management. Medications included: all drugs used to decrease pain and improve function in patients with OA. Intraarticular injections included corticosteroids, viscosupplements, and blood-derived products. Finally, primary surgery included joint resurfacing and primary joint replacement. Supplement 3 provides the full list of included interventions under each category of treatment for knee, hip and mixed OA from studies included in meta-analyses. Observational studies that did not include delivery of an intervention were excluded from the meta-analysis. HSUVs were summarized by key OA affected joint sites (knee, hip, and mixed OA) for baseline (pre-treatment) and at the most commonly available post-treatment time points (i.e.: 3, 6, 12 and 24 months). When more than one HSUVs study was based on the same data, the study with the highest number of participants was included in the meta-analyses. Sub-group meta-analyses by utility elicitation methods were also conducted, where possible.

The meta-analyses were programmed in STATA (STATA 15.1, StataCorp, College Station, Texas, USA), using the "metan" command that required mean and SD/SE as meta-analytical inputs (14). Therefore, when the mean values and SD/SE were not reported, we used 95% CIs, median, minimum, maximum, first quartile, and third quartile values to estimate these parameters (15-17). HSUVs at baseline in observational studies and in both control and intervention groups of trials were pooled (termed pre-treatment HSUVs). Post-treatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups), for each time points. Heterogeneity among the pooled studies was assessed using the I^2 statistic (where an I^2 >=50% indicated substantial heterogeneity) (17). To account for within-study and between-study heterogeneity, random effects models were estimated.

Results

Eligible Studies

Initially 7,621 potential references were identified (Figure 1). After we removed duplicates (n=4,358), 3,263 were left for title and abstract screening. We excluded 2,593 during title and abstract screening, leaving 670 for the full-text assessment. Of those, 522 were excluded due to not meeting the inclusion criteria (Figure 1). Three additional studies identified through reference hand-searching were subsequently included, resulting in a final total of 151 (including 7 abstracts) being included in the systematic review. Eighty-eight of these studies were included in meta-analyses (including 4 conference abstracts).

Results of systematic review

Publication date, study setting and study design

The majority (n=131, 87%) of included studies were published after 2010 (Figure 2A). More than half (n=86, 57%) were conducted in Europe, followed by Asia (n=20, 13%) and the Americas (n=16, 11%). Four studies focused on Australians with OA, one study was conducted in multiple countries, and 24 studies did not report the study setting (Figure 2B). Fifty-eight (38%) included studies were trials, 65 (43%) were observational studies of interventions, and 28 (19%) were observational studies that did not have an intervention component.

Affected OA joint sites

Fifty-nine (39%) studies focused on knee OA and 41 (27%) focused on hip OA. Thirteen (9%) studies focused on both knee and hip OA and reported HSUVs separately. Two (1%) and six (4%) studies were focused on shoulder and hand OA, respectively. Thirty studies (20%) focused on mixed OA (Figure 2C).

Knee OA

Of the 72 knee OA related studies, 10 (14%) focused on core interventions, 6 (8%) and 5 (7%) focused on medication and injection treatments, respectively. Thirty-two (44%) studies focused on surgical treatments. Seven investigated other treatments such as massage, foot insoles, brace, and mud therapy. Twelve (17%) reporting the cross-sectional HSUVs of knee OA did not focus on any specific treatment (Table 1).

Hip OA

Of the 54 hip OA related studies, most (n=46, 85%) focused on surgical treatments. Two (4%) studies focused on core interventions, one investigated balneotherapy, and 5 (9%) reporting the cross-sectional HSUVs of hip OA did not focus on any specific treatment. There were no studies reporting the HSUVs related to hip OA medication and injection treatments (Table 1).

Other types of OA

Two shoulder OA-related studies focused on surgical treatments. Among six hand OA-related studies, two reported the cross-sectional HSUVs of hand OA populations, and one study each focused on spa, mud, a core intervention and surgery treatment.

Mixed OA type

Of the 30 mixed OA related studies, 14 (50%) focused on core interventions, and 12 (40%) reported the cross-sectional HSUVs of an OA population without specifying any treatment type. Two studies focused on surgical treatments, one focused on medication and one focused on spa therapy (Table 1).

Health State Utility measures

Nine HSUV measures were used in the included studies, with most (n=120, 79%) studies using the EQ-5D, followed by the SF-6D (n=12, 8%), HUI2/3 (n=4, 3%), and Quality of Well-Being (QWB) (n=3, 2%). One study each used the AQoL-6D, and the 15D. Ten studies (7%) included more than one measure including the Paper Adaptive Test (PAT-5D-QoL), SG, and RS.

Results of meta-analysis

Studies included in meta-analyses

Fifty-one knee OA related studies (Figure 2D) qualified for meta-analyses. Nine, 6, 5 and 31 related to core interventions, medications, injections, and primary surgery, respectively (Supplements 2.2-2.5). Thirty-four hip OA related studies (Figure 2D) qualified for meta-analyses. Two and 32 related to core interventions and primary surgery, respectively (Supplements 2.6-2.7). Thirteen studies for mixed OA core interventions qualified for meta-analyses (Figure 2D, Supplement 2.8).

The post-treatment time points included in meta-analyses varied between different OA joint sites and treatments based on data availability (Table 2).

HSUVs of Knee OA

The pooled mean baseline (pre-treatment) HSUV of knee OA core interventions was 0.64 (number of HSUVs pooled [n]=19, 95% CI:0.61−0.66, I^2=99%). The pooled HSUVs measuring this at 3 months in post-intervention populations were higher (0.73, n=6, 95% CI:0.70−0.76, I^2= 91%). The pooled 6-month and 1-year HSUVs did not differ significantly from baseline (0.65, n=4, 95% CI:0.60−0.71, I^2= 97% at 6-month; 0.71, n=5, 95% CI:0.64−0.79, I^2=1 at 1-year, respectively). In the subgroup analyses, there were significant difference in HSUVs estimates between different MUIs at each time point. See Table 2 and Supplement Figure 1.

The pooled mean HSUV for knee OA medication treatment was significantly different at baseline (0.56, n=9, 95% CI: 0.45−0.68, I^2=1) than at 3-months follow-up (0.75, n=3, 95% CI:0.70−0.80, I^2=87%). All knee medication related HSUVs were based on the EQ-5D. See Table 2 and Supplement Figure 2.

The pooled HSUVs for knee OA (intraarticular) injections were similar at baseline (0.58, n=7, 95% CI:0.50−0.66, I^2=94%) and 1-year post-treatment (0.63, n=1, 95% CI:0.59−0.67). The baseline HSUVs estimates significantly differed between EQ-5D and HUI3 measures. See Table 2, and Supplement Figure 3.

For knee OA primary surgeries, the pooled mean HSUV was 0.52 (n=55, 95% CI:0.49−0.55, I^2=99.7%) at baseline. A significant difference was found between HSUVs of baseline and various post-surgery time points: 6 months (0.71, n=21, 95% CI:0.69−0.74, I^2=95%); 1-year (0.77, n=18, 95% CI:0.73−0.81, I^2=99%); and 2-years (0.74, n=17, 95% CI:0.71−0.78, I^2=99%). Significant differences existed between different MAUIs at each time point. See Table 2 and Supplement Figure 4.

HSUVs of Hip OA

Only 2 studies focused on hip OA core interventions, HSUVs did not differ significantly between the baseline (0.71, n=3, 95% CI:0.59−0.84, I^2=99%), 3-months (0.72, n=2, 95% CI:0.59−0.84, I^2=98%), or 1-year (0.72, n=2, 95% CI:0.58−0.85, I^2=98%) post-interventions. All HSUVs were based on the EQ-5D. See Table 2 and Supplement Figure 5.

For hip OA primary surgery treatments, there was a significant difference between the pooled mean HSUVs of baseline (0.52, n=46, 95% CI:0.49−0.56, I^2=1) and post-surgery periods: 6 months (0.79, n=9, 95% CI:0.76−0.82, I^2=94%); 1-year (0.83, n=22, 95% CI:0.80−0.85, I^2=99%) and 2-years (0.84, n=11, 95% CI:0.80−0.87, I^2=98%). Significant differences existed between different MAUIs at each time point. See Table 2 and Supplement Figure 6.

HSUVs of Mixed OA

For mixed OA core interventions, there was a significant difference between the pooled mean HSUVs of baseline (0.61, n=27, 95% CI:0.59−0.64, I^2=99%) and 3-months post intervention (0.71, n=10, 95% CI:0.68−0.73, I^2=97%), and 1-year post intervention (0.69, n=12, 95% CI: 0.66−0.71, I^2=98%). The same trend was found for EQ-5D HSUVs but not for SF-6D. See Table 2 and Supplement Figure 7.

Discussion

This is the first wide-ranging systematic review of OA-related HSUVs and meta-analyses on HSUVs for people with different OA affected joint sites before and after various treatments. Our systematic review identified important areas where the current evidence is lacking, namely under-represented geographical locations/ethnicities, affected OA joint sites, treatment options and HSUVs based on more sensitive MAUIs. Our meta-analyses provide a HSUV database for alternative pre- and post- OA treatments that could offer a variety of HSUV inputs for future cost utility models of OA-related conditions. HSUVs associated with four key treatment categories (core interventions, medication, injection, and surgery) often differed, as expected, pre- and post-treatment. Furthermore, we found significant inter-MAUI differences in the mean HSUVs, which is as expected from alternative descriptive systems and utility algorithms. As such this review provides important information that could be used by health economists and policy makers to determine the cost-effectiveness of various OA treatments and long-term disease outcomes using modelling techniques.

Our systematic review identified numerous gaps in the data on OA-related HSUVs, including geographical locations/ethnicities, affected OA joint sites, treatment options and HSUVs based on more sensitive MAUIs. We found that more than half (57%) of included studies were conducted in Europe, and none in Africa. Because HSUVs should ideally be based on local populations preferences, the generalizability of our results to under-represented (e.g.: African and Asian) populations may therefore be limited. Seventy-six percent of included studies focused on knee and hip OA, while other joint sites (e.g.: shoulder and hand) attracted limited attention. Whilst these results align well with the higher clinical impact, prevalence and societal burden of knee and hip OA (18-20), the increasing prevalence and disease burden of hand and shoulder OA as a result of population aging (21, 22) mandates further primary studies investigating the HSUVs of these joint sites.

The HSUVs that we have meta-analysed differed as expected between alternative OA joint sites, treatments, HSUVs measures and time points. We found a mean HSUVs difference of +0.09 units in knee OA patients using core interventions between baseline and 3-months post-intervention, and this difference exceeds the minimal clinically important difference for all the MAUIs reported in previous studies (from +0.04 units [EQ-5D] to 0.08 units [AQoL-8D]) (23-27). Our findings are consistent with the RCT evidence showing the short-lived effects of knee OA core interventions

(28, 29). Other possible explanations include: the limited number of core intervention studies with a follow-up period of greater than 3 months (and hence, wider 95% CIs for our 6 months and 1-year post-treatment HSUVs), and a likely reduction in the core intervention adherence in the long-term (30, 31).

Most (83.3%) studies of knee OA medication treatments had relatively shorter follow up periods (3-months), with only one study with a follow-up period of >3months. Consistent with RCT evidence of effectiveness of medication treatments (32, 33), the pooled HSUV of studies with follow-up at 3 months post-medication treatment was significantly higher than the pooled HSUV of studies with baseline measures. As we did not have enough data on long-term HSUVs in patients using OA medications, we leave this on the agenda for future research when long-term data becomes available. We found similar HSUVs at baseline and 1-year follow-up for knee OA injection treatments. However, these results should be carefully interpreted and used in economic modelling as being derived from only a limited number of studies (n=5 at baseline and n=1 at 1-year follow-up). HSUVs of knee OA patients recorded the largest difference (+0.25 units) between baseline and 1-year post- primary surgery and it remained relatively stable to 2-years post-primary surgery. These findings are once again consistent with the previous evidence of the effectiveness of knee surgery, suggesting that HSUVs record a significant improvement within 1-year of knee surgery, and this change in HSUVs is sustained for years (34).

Surgery was the commonest treatment in hip OA HSUV studies (85%). HSUVs in patients with primary hip OA surgery were significantly higher at 6-months post-surgery than at baseline and remained improved over the long term. The difference between pooled HSUV before and after surgery over 1-year was smaller in knee OA primary surgery (+0.25 units) than hip OA (+0.31 units). These findings align well with previous research (35) advocating a relatively higher efficacy of hip OA joint surgery. Only 2 studies (both based on the EQ-5D) investigated HSUVs in hip OA core intervention patients, which aligned well with the previous findings of the dearth of studies measuring the HSUVs in patients using hip OA core interventions (36, 37). No studies on hip OA medication and injection treatments were identified in our review as expected (38, 39), hence, no meta-analysis for these treatments was possible. We recommend future studies to investigate HSUVs in patients using medications and injections, subject to the availability of better long-term observed data.

The HSUVs for mixed OA core interventions showed the same trend observed for knee OA, with a significant difference (+0.10 units) between baseline and 3-months post-intervention HSUVs. This aligns with the existing findings of short-term benefits associated with OA core interventions (29). A small number of studies of medication treatment (n=1) for mixed OA did not allow us to generate HSUV estimates 'before' and 'after' medication treatments for use in health-economic modelling. Future primary HSUVs studies in this area should therefore be imperative in bridging this evidence gap.

The EQ-5D was the most commonly used (79%) MAUI in the included studies, with little to no representation from other more detailed MAUIs (e.g.: AQoL-8D) that can more fully capture and assess the complex physical and psychosocial health aspects of OA patients (23, 40). Our MAUI-specific (sub-group) analysis revealed significant differences between HSUVs based on alternative MAUIs (EQ-5D and SF-6D, for example), which is as expected from the MAUIs that are far from identical in terms of their descriptive systems and measurement scales (41). As the key objective of our review was not to explore the extent of agreement between alternative MAUIs, we leave the head-to-head comparison of HSUVs obtained through alternative MAUIs on the agenda for future research. Moreover, there is no consensus on the choice of MAUI to be used in measuring HSUVs of OA patients (41, 42). Future research should also endeavour to identify MAUI (s) that could be preferentially recommended for OA patients.

When the baseline HSUVs for various treatments were compared, the mean baseline HSUVs for knee and hip OA patients using core interventions were significantly higher than those using surgery treatments, which is likely to be due to the specified selection criteria for RCTs. Due to the recommended stepwise approach for OA treatments (43), patients are more likely to receive core interventions at earlier stages of their OA (with better HRQoL) and surgery treatments at more severe stages of their OA (with relatively worse HRQoL), which can also explain this pattern. This result reinforces the need to use different HSUVs in modelling for treatments used at different disease stages.

The strength of this study is that this is the first comprehensive review and meta-analysis of all types of OA-related preference-based HSUVs by OA affected joint sites, OA treatments, and utility elicitation methods. It provided a HSUV database for alternative pre- and post- OA treatments that could offer a variety of HSUV inputs to future cost utility models of OA-related

conditions and identified important areas where there are evidence gaps in these estimates to inform future research needs. Our study has several limitations. It is important that the differences in HSUVs at different time points are not interpreted as true pre-post change or direct evidence of intervention effectiveness as the data do not examine differences in change in HSUVs between controls and intervention groups over time and the data included in pooling at each time point comes from different studies. Heterogeneity of the included studies due to the differences in terms of their study design, settings and HSUV elicitation techniques can affect the interpretation of generated HSUVs. Whilst we have conducted subgroup analyses where possible to highlight some possible sources of heterogeneity, we had limited capacity to explain and account for all sources of heterogeneity. The random effects model in our meta-analyses aims to account for heterogeneity but may have consequences for the precision of model estimates (44). Therefore, in modelling, as well as the pooled mean sensitivity analyses, consideration of the potential imprecision of our estimates is important. A further limitation is that due to the paucity of available studies it was not possible to conduct meta-analyses for all treatments of hip, knee and other OA joint sites or group the treatment types in a more detailed way or to perform meta-regression to account for more than one potential effect modifying variable at a time. Quite a few potentially eligible CUA/CEA reports were not able to be included as they did not adequately report the required (pre-and/or-post-treatment) HSUVs (45, 46), despite clear reporting guidelines that recommend these are reported (47, 48). We recommend that future CUA/CEA studies refer to these guidelines to help improve the availability of this important data. Also, the exploration of long-term HSUVs of patients using different OA treatments was mostly not possible. Finally, due to the paucity of data, we could not generate the estimates of HSUVs associated with alternative therapy adherence levels and medication adverse event types.

Conclusions

Our systematic review found that studies of OA related HSUVs are of wide variety, and differ from each other in terms of their setting, design, focused OA joint sites, utility measurement technique, generalisability and others. The HSUVs that we have generated will be useful in conducting future health economic modelling for people suffering from various OA-related conditions. Our results should however be interpreted with caution as being derived from a relatively small number of heterogeneous studies. More research is needed to investigate changes in HSUVs of OA patients for longer follow-up periods.

Authors' contributions

TZ, BdeG, HA, DA, TW and AP conceived and designed the methods, and critically reviewed/edited the manuscript. TZ and HA conducted literature search, screening and data extraction. TZ conducted the meta-analyses and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank Mr Petr Otahal (Statistical Officer, Menzies Institute for Medical Research, University of Tasmania) for his help in conducting the meta-analyses.

Reference

1. Wilson R, Blakely T, Abbott JH. Radiographic knee osteoarthritis impacts multiple dimensions of health-related quality of life: data from the Osteoarthritis Initiative. Rheumatology (Oxford). 2018;57(5):891-9.
2. Abbott JH, Usiskin IM, Wilson R, Hansen P, Losina E. The quality-of-life burden of knee osteoarthritis in New Zealand adults: A model-based evaluation. PLoS One. 2017;12(10):e0185676.
3. Henderson JV, Harrison CM, Britt HC, Bayram CF, Miller GC. Prevalence, causes, severity, impact, and management of chronic pain in Australian general practice patients. Pain Med. 2013;14(9):1346-61.
4. Brazier J, Green C. A systematic review of health state utility values for osteoporosis-related conditions. Osteoporosis International. 2002;13(10):768-76.
5. Ahmad H, Taylor BV, van der Mei I, Colman S, O'Leary BA, Breslin M, et al. The impact of multiple sclerosis severity on health state utility values: evidence from Australia. Multiple Sclerosis Journal. 2017;23(8):1157-66.
6. Meregaglia M, Cairns J. A systematic literature review of health state utility values in head and neck cancer. Health Qual Life Outcomes. 2017;15(1):174.
7. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Annals of medicine. 2001;33(5):358-70.
8. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.
9. Arber M, Glanville J, Isojarvi J, Baragula E, Edwards M, Shaw A, et al. Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care. 2018;34(6):547-54.
10. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. Value in Health. 2013;16(4):686-95.
11. Arber M, Garcia S, Veale T, Edwards M, Shaw A, Glanville JM. Performance of Ovid Medline search filters to identify health state utility studies. Int J Technol Assess Health Care. 2017;33(4):472-80.

12. Better systematic reivew management C, accessed: https://www.covidence.org/home, dated 29 January 2019.
13. The Royal Australian College of General Practitioners. Guideline for the management of knee and hip osteoarthritis. 2nd edn. East Melbourne, Vic: RACGP, 2018.
14. Harris RJ, Deeks JJ, Altman DG, Bradburn MJ, Harbord RM, Sterne JA. Metan: fixed-and random-effects meta-analysis. The Stata Journal. 2008;8(1):3-28.
15. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014;14:135.
16. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res. 2018;27(6):1785-805.
17. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons; 2011.
18. Osteoarthritis Research Society International. Osteoarthritis: a serious disease. https://www.oarsi.org/sites/default/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf.
19. Mahir L, Belhaj K, Zahi S, Azanmasso H, Lmidmani F, El Fatimi A. Impact of knee osteoarthritis on the quality of life. Annals of physical and rehabilitation medicine. 2016;59:e159.
20. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clinical and Experimental Research. 2005;17(4):255-63.
21. Millett PJ, Gobezie R, Boykin RE. Shoulder osteoarthritis: diagnosis and management. Am Fam Physician. 2008;78(5):605-11.
22. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Annals of the rheumatic diseases. 2011;70(9):1581-6.
23. Campbell JA, Palmer AJ, Venn A, Sharman M, Otahal P, Neil A. A Head-to-Head Comparison of the EQ-5D-5L and AQoL-8D Multi-Attribute Utility Instruments in Patients Who Have Previously Undergone Bariatric Surgery. The Patient - Patient-Centered Outcomes Research. 2016;9(4):311-22.

24. Campbell JA, Hensher M, Neil A, Venn A, Wilkinson S, Palmer AJ. An exploratory study of long-term publicly waitlisted bariatric surgery patients' quality of life before and 1 year after bariatric surgery, and considerations for healthcare planners. PharmacoEconomics-open. 2018;2(1):63-76.
25. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. Journal of clinical epidemiology. 2003;56(5):395-407.
26. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Australian and New Zealand journal of public health. 2005;29(2):136-42.
27. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health and quality of life outcomes. 2003;1(1):4.
28. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. Br J Sports Med. 2015;49(24):1554-7.
29. Pisters MF, Veenhof C, Van Meeteren NL, Ostelo RW, De Bakker DH, Schellevis FG, et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. Arthritis Care & Research. 2007;57(7):1245-53.
30. van Gool CH, Penninx BWJH, Kempen GIJM, Rejeski WJ, Miller GD, van Eijk JTM, et al. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. Arthritis Care & Research. 2005;53(1):24-32.
31. Marks R, Allegrante JP. Chronic Osteoarthritis and Adherence to Exercise: A Review of the Literature. Journal of Aging and Physical Activity. 2005;13(4):434-60.
32. Lin J, Zhang W, Jones A, Doherty M. Efficacy of topical non-steroidal anti-inflammatory drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials. Bmj. 2004;329(7461):324.
33. Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV, et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology. 2000;39(10):1095-101.
34. Nilsdotter A-K, Toksvig-Larsen S, Roos E. A 5 year prospective study of patient-relevant outcomes after total knee replacement. Osteoarthritis and cartilage. 2009;17(5):601-6.

35. Bachmeier CJ, March LM, Cross MJ, Lapsley HM, Tribe KL, Courtenay BG, et al. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. Osteoarthritis Cartilage. 2001;9(2):137-46.
36. Murphy SL, Robinson-Lane SG, Niemiec SLS. Knee and hip osteoarthritis management: a review of current and emerging non-pharmacological approaches. Current Treatment Options in Rheumatology. 2016;2(4):296-311.
37. Lee PYF RS, Chandrappa MH, Othman A, Jury C, et al. . Modern Non-Pharmacological and Non-Surgical Treatments for Hip Pain. J Arthritis 2018;7: 262(DOI: 10.4172/2167-7921.1000262).
38. Myers J, Wielage RC, Han B, Price K, Gahn J, Paget MA, et al. The efficacy of duloxetine, non-steroidal anti-inflammatory drugs, and opioids in osteoarthritis: a systematic literature review and meta-analysis. BMC Musculoskelet Disord. 2014;15:76.
39. Ayral X. Injections in the treatment of osteoarthritis. Best practice & research Clinical rheumatology. 2001;15(4):609-26.
40. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology. 1999;38(9):807-13.
41. Zhang F, Yang Y, Huang T, Zhang Y, Zhao L, Li S. Is there a difference between EQ-5D and SF-6D in the clinical setting? a comparative study on the quality of life measured by AIMS2-SF, EQ-5D and SF-6D scales for osteoarthritis patients. Int J Rheum Dis. 2018;21(6):1185-92.
42. NIfHaC E. Guide to the methods of technology appraisal 2013. London: NICE. 2013.
43. Sinusas K. Osteoarthritis: diagnosis and treatment. American family physician. 2012;85(1).
44. Borenstein M, Hedges L, Rothstein H. Meta-analysis: Fixed effect vs. random effects. Meta-analysis com. 2007.
45. Abbott J, Wilson R, Pinto D, Chapple C, Wright A, Team MT. Incremental clinical effectiveness and cost effectiveness of providing supervised physiotherapy in addition to usual medical care in patients with osteoarthritis of the hip or knee: 2-year results of the MOA randomised controlled trial. Osteoarthritis and cartilage. 2019;27(3):424-34.
46. Bulthuis Y, Mohammad S, Braakman-Jansen LM, Drossaers-Bakker KW, van de Laar MA. Cost-effectiveness of intensive exercise therapy directly following hospital discharge in patients with arthritis: Results of a randomized controlled clinical trial. Arthritis care & research. 2008;59(2):247-54.

47. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. Bmj. 2013;346:f1049.

48. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report. Value in Health. 2015;18(2):161-72.

Table 1 The number and percentage of studies included in systematic review and meta-analyses for each OA affected joint site and treatment

	Studies in systematic review	Studies in meta-analyses
Knee OA (n, %)		
Core intervention	10 (14)	9 (18)
Medication	6 (8)	6 (12)
Injection	5 (7)	5 (10)
Surgery	32 (44)	31 (61)
Other treatments	7 (10)	0 (0)
No treatments	12 (17)	0 (0)
Sub-total	**72**	**51**
Hip OA (n, %)		
Core intervention	2 (4)	2 (6)
Surgery	46 (85)	32 (94)
Other treatments	1 (2)	0 (0)
No treatments	5 (9)	0 (0)
Sub-total	**54**	**34**
Mixed OA (n, %)		
Core intervention	14 (50)	13 (100)
Other treatments	4 (13)	0 (0)
No treatments	12 (40)	0 (0)
Sub-total	**30**	**13**
Other joint sites of OA	8	0
Total	**164***	**98***

*Note: *Thirteen studies reporting knee and hip HSUVs separately have been counted in both the hip and knee OA groups and ten of them were included in the meta-analyses. Mixed OA included a variety of OA patients without specifying their OA type, Other joint site of OA included shoulder and hand OA studies. OA=osteoarthritis.*

Table 2 The number of pooled HSUVs and representing population in meta-analyses and the pooled mean HSUVs

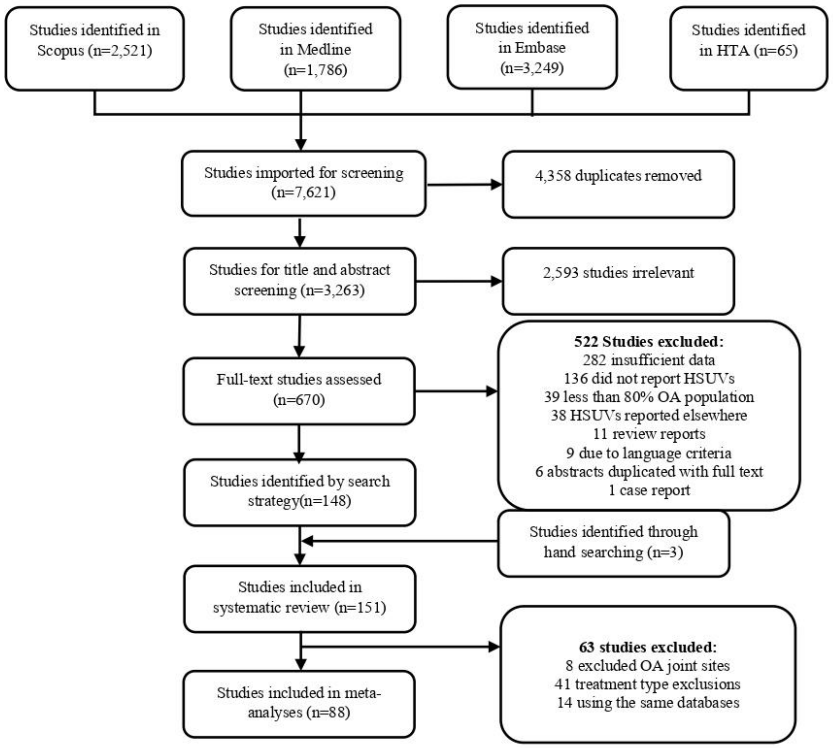
	Overall			EQ-5D			SF-6D		
	Number of HSUVs	Representing population	Mean (95%CI)	Number of HSUVs	Representing population	Mean (95%CI)	Number of HSUVs	Representing population	Mean (95%CI)
Knee OA									
Core intervention									
Baseline	19	31349	0.64 (0.61, 0.66)	12	30869	0.60 (0.57, 0.62)	2	119	0.65 (0.55, 0.75)
3-mo	6	30,380	0.73 (0.70, 0.76)	4	30,233	0.70 (0.69, 0.72)	0	0	--
6-mo	4	542	0.65 (0.60, 0.71)	4	542	0.65 (0.60, 0.71)	0	0	--
1-year	5	20,549	0.71 (0.64, 0.79)	3	20,402	0.64 (0.56, 0.72)	0	0	--
Medication									
Baseline	9	3,749	0.56 (0.45, 0.68)	9	3,749	0.56 (0.45, 0.68)	0	0	--
3-mo	3	249	0.75 (0.70, 0.80)	3	249	0.75 (0.70, 0.80)	0	0	--
Injection									
Baseline	7	473	0.58 (0.50, 0.66)	5	224	0.63 (0.56, 0.70)	0	0	--
1-year	1	122	0.63 (0.59, 0.67)	0	0		0	0	--
Primary surgery									
Baseline	55	53,434	0.52 (0.49, 0.55)	43	45,434	0.49 (0.44, 0.53)	10	7,797	0.62 (0.60, 0.64)
6-mo	21	4,260	0.71 (0.69, 0.74)	14	3,378	0.72 (0.70, 0.73)	5	719	0.71 (0.68, 0.75)
1-year	18	3,790	0.77 (0.73, 0.81)	12	2,179	0.78 (0.73, 0.84)	4	1,456	0.75 (0.72, 0.79)
2-year	17	15,160	0.74 (0.71, 0.78)	11	8,872	0.77 (0.73, 0.81)	5	6,270	0.71 (0.66, 0.76)
Hip OA									
Core intervention									
Baseline	3	13,773	0.71 (0.59, 0.84)	3	13,773	0.71 (0.59, 0.84)	0		--
3-mo	2	13,671	0.72 (0.59, 0.84)	2	13,671	0.72 (0.59, 0.84)	0		--
1-year	2	8,421	0.72 (0.58, 0.85)	2	8,421	0.72 (0.58, 0.85)	0		--
Primary surgery									
Baseline	46	59,846	0.52 (0.49, 0.56)	34	52,671	0.50 (0.46, 0.54)	6	6,791	0.58 (0.56, 0.60)
6-mo	9	3,922	0.79 (0.76, 0.82)	6	3,727	0.80 (0.77, 0.83)	0	0	--
1-year	22	42,788	0.83 (0.80, 0.85)	20	42,468	0.83 (0.80, 0.85)	1	224	0.80 (0.78, 0.82)
2-year	11	16,732	0.84 (0.80, 0.87)	7	10,228	0.88 (0.85, 0.90)	4	6,504	0.78 (0.75, 0.81)
Mixed OA									
Core intervention									
Baseline	27	9,644	0.61 (0.59, 0.64)	18	5,672	0.58 (0.53, 0.62)	5	3,576	0.69 (0.66, 0.72)
3-mo	10	6,926	0.71 (0.68, 0.73)	6	3,587	0.67 (0.63, 0.71)	2	3,192	0.73 (0.65, 0.82)
1-year	12	7,305	0.69 (0.66, 0.71)	6	3,819	0.65 (0.60, 0.71)	4	3,339	0.71 (0.65, 0.77)

Note: Pooling at all time points includes observational data on intervention and at baseline includes both control and active treatment groups from trials, but at follow-up includes only data from active treatment groups from trials. OA=osteoarthritis, HSUVs=health state utility values, 95%CI=95% confidence interval.

Accepted Article

This article is protected by copyright. All rights reserved

Figure 1. Flow chart results of study search based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses methodology.



Note: The exclusions by OA joint sites and treatment type were because of the small numbers of studies in these joint sites and treatments, which meant that meta-analysis was not feasible. Eight exclusions by OA joint sites involved 2 shoulder and 6 hand OA-related studies, 41 exclusions by treatment type involved studies of massage, foot insoles, brace, mud therapy, balneotherapy, spa therapy, revision surgery and observational studies that did not focus on any treatment.

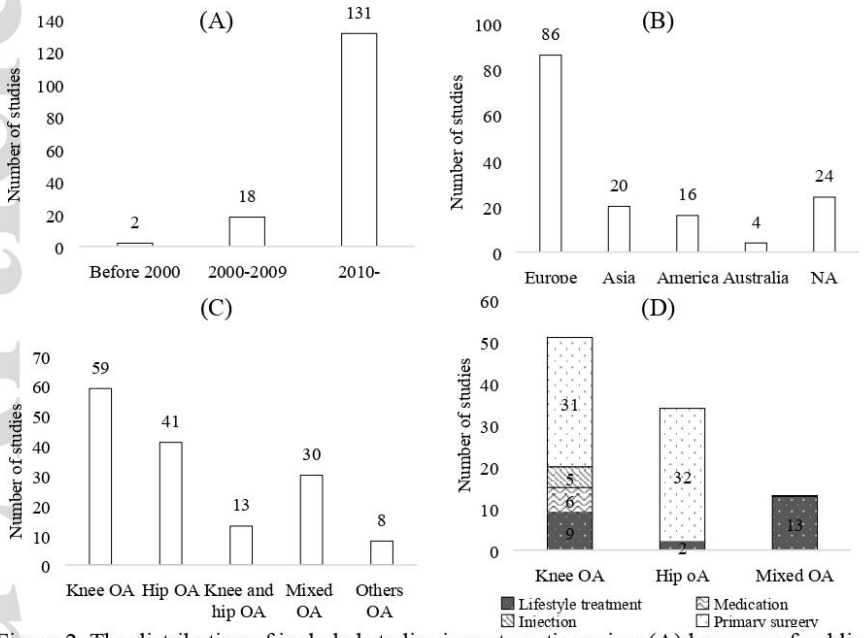


Figure 2. The distribution of included studies in systematic review (A) by years of publication; (B) by study setting; (C) by OA joint sites; and (D) the distribution of included studies in meta-analysis by OA joint sites and treatments. *Note: Knee and hip OA group in Figure 2C included studies reporting HSUVs for each type separately. Other OA included hand and shoulder OA studies.*

Supplement 1: The search strategy in Embase.

Supplement 2.1: Characteristics of studies included in systematic review

Supplement 2.2: Details of HSUVs included in the meta-analyses for knee OA core interventions

Supplement 2.3: Details of HSUVs included in the meta-analyses for knee OA medication treatments

Supplement 2.4: Details of HSUVs included in the meta-analyses for knee OA injection treatments

Supplement 2.5: Details of HSUVs included in the meta-analyses for knee OA primary surgery treatments

Supplement 2.6: Details of HSUVs included in the meta-analyses for hip OA core interventions

Supplement 2.7: Details of HSUVs included in the meta-analyses for hip OA primary surgery treatments

Supplement 2.8: Details of HSUVs included in the meta-analyses for mixed OA core interventions

Supplement 3: The list of included interventions under each category of treatments for knee, hip and mixed OA from studies included in meta-analyses

Supplement Figure 1: The forest plots of the meta-analyses for knee OA core intervention.

Supplement Figure 2: The forest plots of the meta-analyses for knee OA medication treatment.

Supplement Figure 3: The forest plots of the meta-analyses for knee OA injection treatment.

Supplement Figure 4: The forest plots of the meta-analyses for knee OA primary surgery treatment.

Supplement Figure 5: The forest plots of the meta-analyses for hip OA core intervention.

Supplement Figure 6: The forest plots of the meta-analyses for hip OA primary surgery treatment.

Supplement Figure 7: The forest plots of the meta-analyses for mixed OA core intervention.

Supplement 4.1: The search strategy in Embase.

1. Quality-Adjusted Life Years/

2. (quality adjusted or adjusted life year$).ti,ab,kf.

3. (qaly$ or qald$ or qale$ or qtime$).ti,ab,kf.

4. (illness state$1 or health state$1).ti,ab,kf.

5. (hui or hui1 or hui2 or hui3).ti,ab,kf.

6. (multiattribute$ or multi attribute$).ti,ab,kf.

7. (utility adj3 (score$1 or valu$ or health$ or cost$ or measur$ or disease$ or mean or gain or gains or index$)).ti,ab,kf.

8. utilities.ti,ab,kf.

9. (eq-5d or eq5d or eq-5 or eq5 or euroqual or euro qual or euro qual5d or euroqual5d or euro qol or euroqol or euro qol5d or euroqol5d or euro quol or euroquol or euro quol5d or euroquol5d or eur qol or eurqol or eur qol5d or eur qol5d or eur?qul or eur?qul5d or euro$ quality of life or european qol).ti,ab,kf.

10. (euro$ adj3 (5 d or 5d or 5 dimension$ or 5dimension$ or 5 domain$ or 5domain$)).ti,ab,kf.

11. (sf36$ or sf 36$ or sf thirtysix or sf thirty six).ti,ab,kf.

12. (time trade off$1 or time tradeoff$1 or tto or timetradeoff$1).ti,ab,kf.

13. quality of life/ and ((quality of life or qol) adj (score$1 or measure$1)).ti,ab,kf.

14. quality of life/ and ec.fs.

15. quality of life/ and (health adj3 status).ti,ab,kf.

16. (quality of life or qol).ti,ab,kf. and cost-benefit analysis/

17. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or shortform six or short form six).ti,ab,kf.

18. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17

19. osteoarthritis.ti,ab,kf.

20. osteoarthritis/

21. 19 or 20

22. 18 and 21

Supplement 4.2.1 Characteristics of Studies Included in Systematic Review (N=151)

No.	Title	List of authors	Year published	Country	Study design	Target population	Sample size	Mean age (SD), % Female (baseline)	Treatment	HSUVs assessment timepoints	HSUV Elicitation method	Health state utility values (HSUVs) mean (SD)	Abstract only (yes/no)	Included in meta-analysis (yes/no)
1	A comparative prospective cohort health economic analysis comparing ankle fusion, isolated great toe fusion and hallux valgus surgery	Loveday, D. T et al.	2018	UK	observational	mixed OA	63	**Ankle Fusion** 68; 50%, **MTPJ Fusion** 62; 55%, **Hallux Valgus Surgery** 56; 94%	Ankle fusion, MTPJ fusion, Hallux valgus surgery	pre-operation post-operation	EQ-5D-5L	**AF** pre 0.30 (95% CI 0.43-0.17) post 0.66 (95% CI 0.77-0.55) **MF** pre 0.45 (95% CI 0.52-0.38) post 0.83 (95% CI 0.90-0.76) **HV** pre 0.71 (95% CI 0.74-0.68) post 0.82 (95% CI 0.88-0.76)	no	no
2	A clinical and radiographic 13 year follow up study of 138 Charnley hip arthroplasties in patients 50 70 years old Comparison of university hospital	Hulleberg, G. et al.	2008	Norway	observational	hip OA	138	Median age (range), % female **Clinical Evaluation** 65 (50-70); 75% **Charnley Category A+B** 64 (51-70); 70% **Charnley Category C** 67 (50-70); 81%	THR	13-years after THR	EQ-5D	**Clinical Evaluation** 0.75 (0.24), **Charnley Category A+B** 0.88 (0.14), **Charnley C** 0.63 (0.25)	no	yes
3	An innovative care model coordinated by a physical therapist and nurse practitioner for osteoarthritis of the hip and knee in specialist care a prospective study	Voorn, V. M et al.	2013	Netherlands	observational	mixed OA	87	71 (9.8); 63%	An innovative care model coordinated by a physical therapist and nurse practitioner	baseline 10 weeks after	EQ-5D-3L	**Baseline** 0.48 (0.30) **10 weeks** 0.59 (0.29)	no	yes

4	A comparison of surgical approaches for primary hip arthroplasty A cohort study of patient reported outcome measures (PROMs) and early revision using linked national databases	Jameson, S. S. et al.	2014	England and Wales	observational	hip OA	3881	**Cemented Posterior** 72.6 (8.1); 68.2% **Cemented Lateral** 73.2 (7.2); 64.1% **Cementless Posterior** 63.2(9.9); 46.2% **Cementless Lateral** 64.3(8.9); 41.9%	posterior vs. lateral surgical approaches	pre-operation post-operation	EQ-5D	**Preop:** cemented posterior: 0.393 (0.307), cemented lateral: 0.341(0.313), cementless posterior: 0.390(0.316), cementless lateral: 0.377(0.318), **Postop**: median (range) cemented posterior: 0.815 (−0.003-1) cemented lateral: 0.760 (−0.016-1) cementless posterior: 0.883 (−0.074-1) cementless lateral: 0.812 (−0.077-1)	no	no
5	An exploratory study of response shift in health-related quality of life and utility assessment among patients with osteoarthritis	Zhang, X. H. et al.	2012	Singapore	observational	knee OA	74	**mean age (range),** 68 (63-76); 81.1%	TKR	baseline 6 months post 18 months post	SF-6D, EQ-5D	median (IQR) **baseline**: SF-6D 0.61 (0.58-0.68); EQ-5D 0.69 (0.17-0.73) **6 months post**: SF-6D 0.69 (0.63-0.72); **18 months post**: SF-6D 0.77 (0.66–0.90); EQ-5D: 0.87 (0.71–1.00)	no	yes
6	Comparative outcomes and cost-utility following surgical treatment of focal lumbar spinal stenosis compared with osteoarthritis of the hip or knee	Rampersaud, Y. R. et al.	2014	Canada	observational	knee OA, hip OA	99 hip OA, 99 knee OA	**hip OA: mean age (range)** 63.0 (40-84); 60% **knee OA: mean age (range)** 64.6 (43-83); 60%	THA, TKA	baseline 5 years post	SF-6D	**THA**: baseline 0.522 (0.222), 5th-year 0.750 (0.161) **TKA**: baseline 0.549 (0.0831), 5th-year 0.776 (0.154)	no	yes
7	Patient expectations and health-related quality of life outcomes following total joint replacement	Gonzalez Saenz De Tejada, M et al.	2010	Spain	observational	mixed OA	881	68.28 (9.85); 51.1%	TJR	baseline	EQ-5D	0.378 (0.341)	no	no

8	Comparison of lifetime incremental cost utility ratios of surgery relative to failed medical management for the treatment of hip, knee and spine osteoarthritis modelled	Tso, P. et al.	2012	Canada	observational	knee OA, hip OA	99 hip OA, 99 knee OA	**hip OA: mean age (range)** 63.0 (40-84); 60% **knee OA: mean age (range)** 64.6 (43-83); 60%	THA, TKA	baseline 2 years post	SF-6D	**THA:** baseline 0.522 (0.222), 2-year 0.715 (0.315) **TKA**: baseline 0.549 (0.0831), 2-year 0.642 (0.148)	no	yes
9	A simple visual analog scale for pain is as responsive as the WOMAC, the SF-36, and the EQ-5D in measuring outcomes of revision hip arthroplasty	Zampelis, V. et al.	2014	NA	observational	hip OA	45	mean age (range): 74 (62-86); 44%	RHA	pre revision 2 years after revision	EQ-5D	**pre**: 0.35 (0.31) **2-year post**: 0.74 (0.17)	no	no
10	Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis reliability, validity and responsiveness	Bilbao, A. et al.	2018	Spain	observational	mixed OA	758	69.78 (10.57); 61.87%	NA	baseline	EQ-5D-5L	0.53 (0.29)	no	no
11	Comparing the validity and responsiveness of the EQ-5D-5L to the Oxford hip and knee scores and SF-12 in osteoarthritis patients 1 year following total joint replacement	Conner-Spady, B. L. et al.	2018	Canada	observational	knee OA, hip OA	537 (hip OA 269 & knee OA 268)	**Hip OA**: 63.98 (10.87); 52% **knee OA:** 64.21 (9.67); 61%	TJR	baseline	EQ-5D-5L	hip OA: 0.35 (0.25); knee OA: 0.39 (0.27)	no	yes

12	Cost effectiveness and quality of life in knee arthroplasty	Lavernia, C. J. et al.	1997	US	observational	knee OA	100	mean age: male 62, female 64 70%	knee arthroplasty surgery	preop. 3 months postop. 6 months postop. 1-year post 2-year post	QWB	**preop**: 0.56 (0.05) **3 months**: 0.56 (0.08) **6 months**: 0.57 (0.09) **1 year**: 0.59 (0.08) **2 year**: 0.59 (0.07)	no	yes
13	Age- and health-related quality of life after total hip replacement	Gordon, M. et al.	2014	Sweden	observational	hip OA	27,245	69 (10); 57%	THR	preop 1-year postop	EQ-5D	**pre** 0.73 (0.11) **1-year post** 0.88 (0.11)	no	no
14	Cost-Effectiveness of five commonly used prosthesis brands for total knee replacement in the UK A study using the NJR Dataset	Pennington, M. et al.	2016	UK	observational	knee OA	53126	**PFC Sigma:** 69.9 (7.3), male 45% **AGC Biomet:** 70.5 (7.3), male 44% **Nexgen:** 69.8 (7.3), male 45% **Genesis 2:** 70.1 (7.3), male 44% **Triathlon:** 69.9 (7.3), male 43%	five different prosthesis brands for TKR	pre-op	EQ-5D-3L	PFC Sigma 0.38 (0.31) AGC Biomet 0.41 (0.31) Nexgen 0.39 (0.31) Genesis 2 0.41 (0.31) Triathlon 0.40 (0.31)	no	yes
15	An assessment of the discriminative ability of the EQ-5Dindex, SF-6D, and EQ VAS, using sociodemographic factors and clinical conditions	Barton, G. R. et al.	2008	UK	observational (no treatment)	mixed OA	220	Not available	NA	cross-sectional	EQ-5D SF-6D	EQ-5D: 0.607 (0.269) SF-6D: 0.675 (0.150)	no	no

16	Cost-utility analysis and economic burden of knee osteoarthritis treatment the analysis from the real clinical practice	Turajane, T. et al.	2012	Thailand	observational	knee OA	1319	**traditional NSAIDs:** 64.73 (10.10); male 29.07% **Celecoxib:** 62.27 (9.39); male 30.26%	traditional NSAIDs vs. celecoxib	baseline 6 months	EQ-5D	**baseline**: traditional NSAIDs -0.079 (0.19), Celecoxib -0.110 (0.18) **6 months**: traditional NSAIDs 0.595 (0.12), Celecoxib 0.602 (0.12)	no	yes
17	Determining Cost-Effectiveness of Total Hip and Knee Arthroplasty Using the Short Form-6D Utility Measure	Elmallah, R. K. et al.	2017	NA	observational	knee OA, hip OA	844 for TKA, 224 for THA	**TKA:** mean age (range) 65 (39-80); men 50% **THA:** mean age (range) 69 (44-88)	TKA, THA	preop 1-year postop	SF-6D	**preop**: TKA 0.62 (0.102), THA 0.614 (0.131) **1 year**: TKA 0.77 (0.113), THA 0.799 (0.134)	no	yes
18	Factors associated with health-related quality of life in Korean older workers	Hong, S. et al.	2015	Korea	observational (no treatment)	mixed OA	249	female 76%	NA	general health	EQ-5D	male: 0.920 (0.092) female: 0.860 (0.160)	no	no
19	Association between changes in global femoral offset after total hip arthroplasty and function, quality of life, and abductor muscle strength	Mahmood, S. S. et al.	2016	Sweden	observational	hip OA	222 (decreased femoral offset 71 & restored femoral offset 73 & increased femoral offset	**decreased FO group:** mean age 71; 46% **restored FO group:** mean age 68; 48% **increased FO group:** mean age 65; 50%	THA	preop 12-15 months post-op	EQ-5D	**decreased FO**: pre 0.44 (0.26); post 0.82 (0.19) **restored FO**: pre 0.43 (0.22); post 0.86 (0.17) **increased FO**: pre 0.51 (0.66); post 0.86 (0.19)	no	yes

20	Does patella resurfacing really matter Pain and function in 972 patients after primary total knee arthroplasty an observational study from the Norwegian Arthroplasty Register	Lygre, S. H. L. et al.	2010	Norway	observational	knee OA	972	76 (7.7); men 29%	TKA	preop.	EQ-5D	0.46 (0.22)	no	yes
21	Early Clinically Relevant Improvement in Quality of Life and Clinical Outcomes 1 Year Postsurgery in Patients with Knee and Hip Joint Arthroplasties	Neuprez, A. et al.	2018	Belgium	observational	knee OA, hip OA	626 (knee OA 280 & hip OA 346)	**knee OA:** 66.74 (8.95); male 45.71% **hip OA:** 64.64 (10.85), male 41.04%	hip arthroplasty, knee arthroplasty	preop. 3 months postop. 6 months postop. 12 months post	EQ-5D	**Knee OA** preop: 0.46 (0.23), 3 months: 0.66 (0.20), 6 months: 0.69 (0.19); 12 months: 0.67 (0.22) **hip OA** preop: 0.44 (0.23), 3 months: 0.71 (0.24), 6 months: 0.74 (0.25), 12 months: 0.71 (0.27)	no	yes
22	Effectiveness of hip or knee replacement surgery in terms of quality adjusted life years and costs	Räsänen, P. et al.	2007	Finland	observational	knee OA, hip OA	223 (primary THA 96 & revision hip replacement 24 & primary TKA 103)	**primary THA patients:** 63 (12); 63% **revision hip replacement:** 69 (10); 50% **primary TKA patients:** 69 (11); 75%	THA, TKA, revision hip replacement	baseline 6 months after operation 12 months after operation	15D	**baseline**: primary THA 0.805 (0.084), revision THA 0.805 (0.084), primary TKA 0.807 (0.093) **6 months** : primary THA 0.868 (0.092), revision THA 0.841 (0.088), primary TKA 0.830 (0.109) **12 months**: primary THA 0.858 (0.115), revision THA 0.823 (0.097), primary TKA 0.841 (0.108)	no	yes

23	Body Mass Index Class Is Independently Associated with Health-Related Quality of Life After Primary Total Hip Arthroplasty An Institutional Registry-Based Study	McLawhorn, A. S. et al.	2017	NA	observational	hip OA	2733	65.51 (0.19); male 45.8%	primary unilateral THA	Pre-op 2-year post	EQ-5D-3L	mean (SE) **baseline**: 0.64 (0.004), **2-year**: 0.89 (0.003)	no	yes
24	Cement in Cement Revision of the Femoral Component Using a Collarless Triple Taper A Midterm Clinical and Radiographic Assessment	Stefanovich-Lawbuary, N. S. et al.	2014	NA	observational	hip OA	44	72 (12.97); 55%	revision surgery of femoral component using collarless triple taper	after revision surgery	EQ-5D	median (IQR) 0.814 (0.587-0.883)	no	no
25	Cemented, cementless, and hybrid prostheses for total hip replacement Cost effectiveness analysis	Pennington, M. et al.	2013	England	observational	hip OA	30203 (cemented prothesis 11955 & cementless prothesis	**cemented prosthesis:** 72.4(6.7); 64.9% **cementless prosthesis:** 67.8(7.2); 55.4% **hybrid prosthesis:** 70.4(7.2); 62.0%	three types of THR: cemented, cementless, hybrid	preoperative	EQ-5D-3L	cemented: 0.34(0.32) cementless: 0.36(0.32) hybrid: 0.34(0.32)	no	no
26	Changes in the WOMAC, EuroQol and Japanese lifestyle measurements among patients undergoing total hip arthroplasty	Fujita, K. et al.	2009	Japan	observational	hip OA	451	60.6 (10); 84.5%	THA	preoperative 6 weeks post 6 months posts	EQ-5D	**pre**: 0.56 (0.13), **6 weeks post**: 0.74 (0.16), **6 months pos**t: 0.79 (0.17)	no	yes

27	Outcome of total hip arthroplasty, but not of total knee arthroplasty, is related to the preoperative radiographic severity of osteoarthritis	Tilbury, C. et al.	2016	Netherlands	observational	knee OA, hip OA	573 (THA KL grade 0-2: 77, THA KL grade 3-4: 225; TKA KL grade 0-2: 74, TKA KL	**THA patients KL grade 0-2:** 66 (11); 58% **THA patients KL grade 3-4:** 68 (9); 59% **TKA patients KL grade 0-2:** 56 (76); 76% **TKA patients KL grade 3-4:** 67 (10); 69%	THA/TKA	preop	EQ-5D	THA KL grade 0-2: 0.6 (0.2) THA KL grade 3-4: 0.6 (0.3) TKA KL grade 0-2: 0.6 (0.3) TKA KL grade 3-4: 0.6 (0.2)	no	yes
28	Clinical and radiological results 7 years after Copeland shoulder resurfacing arthroplasty in patients with primary glenohumeral osteoarthritis	Verstraelen, F. U. et al.	2018	NA	observational	shoulder OA	27	mean age (range): 67.7 (50.2-85.1); 55.5%	cementless stemless resurfacing arthroplasty	7 year follow up	EQ-5D	0.8 (0.1)	no	no
29	Evaluation of 1031 primary titanium nitride coated mobile bearing total knee arthroplasties in an orthopedic clinic	Breugem, S. J. M. et al.	2017	NA	observational	knee OA	910	mean (range): 65.4 (36-94) male 37.1%	primary titanium nitride coated mobile bearing TKA	1 year after TKA	EQ-5D	median (IQR) non-revised: 0.84 (0.78-1) revised: 0.76 (0.35-0.78)	no	yes
30	Comparing Short Form 6D, Standard Gamble, and Health Utilities Index Mark 2 and Mark 3 utility scores Results from total hip arthroplasty patients	Feeny, D. et al.	2004	Canada	observational	hip OA	86	mean age 69; male 54%	THA	pre-surgery post-surgery	SG, HUI2, HUI3, SF-6D	**pre-surgery**: SG 0.61 (0.33); SF-6D 0.59 (0.10); HUI2 0.55 (0.20); HUI3 0.49 (0.21) **post-surgery**: SG 0.76 (0.25); SF-6D 0.69 (0.11); HUI2 0.76 (0.15); HUI3 0.72 (0.18)	no	yes

31	Outcomes of unicompartmental knee arthroplasty after aseptic revision to total knee arthroplasty	Leta, T. H. et al.	2016	Norway	observational	knee OA	277 (TKA-TKA 150 & UKA-TKA 127)	TKA -TKA: female 77% UKA-TKA: female 61%	revision TKA	pre revision post revision	EQ-5D	**pre:** TKA-TKA 0.44 (0.23) UKA-TKA 0.41 (0.21) **post:** TKA-TKA 0.63 (0.24) UKA-TKA 0.63 (0.24)	no	no
32	Long term health impact of playing professional football in the United Kingdom	Turner, A. P. et al.	2000	UK	observational (no treatment)	mixed OA	138	mean age (SD): 56.1 (11.8)	NA	cross-sectional	EQ-5D	OA: 0.58 (0.31)	no	no
33	Consistency in patient-reported outcomes after total hip replacement	Bengtsson, A. et al.	2017	Sweden	observational	hip OA	15755	mean age 67, 58%	THA	preop. 1-year postop. 6-year postop.	EQ-5D	**preop.:** men 0.47 (0.30), women 0.38 (0.31) **1-year postop:** men 0.82 (0.21), women 0.77 (0.23) **6-year postop:** men 0.79 (0.25), women 0.72 (0.28)	no	no
34	Patient-reported outcomes after total and unicompartmental knee arthroplasty A study of 14 076 matched patients from the national joint registry for EngLand and Wales	Liddle, A. D. et al.	2015	England and Wales	observational	knee OA	14076 (TKA 10557 & UKA 3519)	**TKA:** 64.4 (8.6); male 53.0% **UKA:** 664.4 (8.2); male 53.1%	TKA vs. UKA	preop.	EQ-5D	TKA 0.482 (0.291), UKA 0.481 (0.289)	no	yes

35	Quality of Life and Cost-Effectiveness 1 Year After Total Hip Arthroplasty	Lavernia, C. J. et al.	2011	US	observational	knee OA, hip OA	12782 (primary THA 5463 &hip resurfacing 843 &primary TKA 5398 & UKA 240)	**primary THA:** 64.2 (11.5); 55.6% **hip resurfacing:** 51.2 (7.6); 21.5% **primary TKA:** 66.9 (9.5); 61.2% **UKA:** 63.7 (10.2); 60.4%	primary THA, hip resurfacing, TKA, UKA	preop. 2-year postop 5-year postop	EQ-5D, SF-6D	**Hip** preop: EQ-5D THA 0.63 (0.19), resurfacing 0.68 (0.16); SF-6D THA 0.6 (0.11), resurfacing 0.63 (0.10) 2-year post: EQ-5D THA 0.88 (0.14), resurfacing 0.93 (0.11); SF-6D THA 0.79 (0.13), resurfacing 0.82 (0.12) 5-year postop: SF-6D THA 0.81 (0.13), resurfacing 0.84 (0.12) **Knee** preop: EQ-5D TKA 0.68 (0.17) UKA 0.71 (0.14); SF-6D TKA 0.63 (0.11), UKA 0.65 (0.11) 2-year post: EQ-5D TKA 0.84 (0.15), UKA 0.87 (0.13); SF-6D TKA 0.76 (0.13), UKA 0.8 (0.13) 5-year postop: SF-6D TKA 0.79 (0.13), UKA 0.8 (0.14)	no	yes
36	Predicting the Long-Term Gains in Health-Related Quality of Life After Total Knee Arthroplasty	Schilling, C. G. et al.	2017	NA	observational	knee OA	488	71.0 (8.5); 33%	TKA	preop. 1-year post 7-year post	SF-6D	**preop:** 0.57 (0.10) **1-year post**: 0.71 (0.15) **7-year post**: 0.69 (0.15)	no	yes
37	Is there a difference between EQ-5D and SF-6D in the clinical setting	Zhang, F. et al.	2018	China	observational (no treatment)	mixed OA	100	male 34%	NA	cross-sectional	EQ-5D, SF-6D	EQ-5D 0.429 (0.313), SF-6D 0.675 (0.980)	no	no

38	Different patient-reported outcomes in immigrants and patients born in Sweden	Krupic, F. et al.	2014	Sweden	observational	hip OA	18791 (born in Sweden 17340 & Nordic countries 819 & Europe 523 &outside Europe 109)	**born in Sweden**: mean age 68.5 (95%CI 68.4–68.6); male 43% **Nordic countries:** mean age 68.0 (95%CI 67.4–68.6); male 31% **Europe:** mean age 66.7 (95%CI 65.8–67.7); male 42% **outside Europe**: mean age 61.3 (95%CI 58.4–64.3); male 56%	THA	preop. 1 year postop.	EQ-5D	**preop**.: Sweden 0.40 (0.39–0.41), Nordic countries 0.36 (0.34–0.38), Europe 0.33 (0.30–0.36), Outside Europe 0.29 (0.23–0.36) **1 year postop**.: Sweden 0.78 (0.77–0.78) , Nordic countries 0.74 (0.72–0.76), Europe 0.70 (0.67–0.72), Outside Europe 0.64 (0.58–0.70)	no	no
39	Quality of life benefits of knee arthroplasty for osteoarthritis	Neuprez, A. et al.	2014	NA	observational	knee OA	279	66.8 (8.9); male 46%	TKA	preop 3 months post 6 months post	EQ-5D	**preop**: 0.46 (0.23) **3 months post**: 0.66 (0.20) **6 months post**: 0.68 (0.19)	yes	yes
40	The effects of adherence to non-Steroidal anti-inflammatory drugs and factors influencing drug adherence in patients with knee osteoarthritis	Park, K. K. et al.	2016	Korea	observational	knee OA	1334	74.3 (5.4); male 79.8%	NSAIDs	baseline 3 weeks	EQ-5D	**baseline**: 0.71 (0.2) **3 weeks**: 0.76 (0.1)	no	yes

41	Joint protection and hand exercises for hand osteoarthritis an economic evaluation comparing methods for the analysis of factorial trials	Oppong, R. et al.	2015	UK	trial	hand OA	257	66 (9.1); 66%	joint protection only, hand exercises only, and joint protection plus hand exercises compared with leaflet and advice	baseline 3 months post 6 months post 12 months post	EQ-5D	**baseline**: leaflet advice 0.623 (0.26), joint protection 0.646 (0.25), hand exercise 0.645 (0.21), joint protection plus exercise 0.659 (0.26) **3 months**: leaflet advice 0.665 (0.24), joint protection 0.682 (0.17), hand exercise 0.660 (0.22), joint protection plus exercise 0.676 (0.24) **6 months**: leaflet advice 0.658 (0.25), joint protection 0.635 (0.25), hand exercise 0.692 (0.18), joint protection plus exercise 0.672 (0.24) **12 months**: leaflet advice 0.634 (0.22), joint protection 0.684 (0.19), hand exercise 0.708 (0.18), joint protection plus exercise 0.659 (0.27)	no	no
42	Osteoarthritis affects health-related quality of life in Korean adults with chronic diseases	Yang, J. H. et al.	2017	Korea	observational (no treatment)	mixed OA	13395	Not available	NA	cross-sectional	EQ-5D	50-59 years: 0.858 (0.134) 60-69 years: 0.798 (0.201) 70-79 years: 0.734 (0.24) ≥80: 0.697 (0.257) men: 0.816 (0.19) women: 0.768 (0.221)	no	no
43	Does surgeon experience affect patient reported outcomes 1 year after primary total hip arthroplasty	Jolbäck, P. et al.	2018	Sweden	observational	hip OA	6713	69 (10); male 42%	THA	preop. 1-year postop.	EQ-5D	**preop**. 0.42 (0.31) **1-year postop**. 0.77 (0.24)	no	no

44	The effects of age on patient-reported outcome measures in total knee replacements	Williams, D. P. et al.	2013	UK	observational	knee OA	2456	71.4 (9.1); men 39.2%	TKR	preop 6 months 2-year	EQ-5D	**preop**: < 55 years 0.36 (0.32) 55 to 64 years 0.37 (0.33) 65 to 74 years 0.43 (0.32) 75 to 84 years 0.45 (0.31) ≥ 85 years 0.42 (0.31) **6 months**: < 55 years 0.69 (0.30) 55 to 64 years 0.73 (0.24) 65 to 74 years 0.74 (0.24) 75 to 84 years 0.74 (0.22) ≥ 85 years 0.70 (0.26) **2 year**: < 55 years 0.76 (0.24) 55 to 64 years 0.73 (0.29) 65 to 74 years 0.76 (0.27) 75 to 84 years 0.74 (0.24) ≥ 85 years 0.66 (0.27)	no	yes
45	Total or partial knee replacement Cost-utility analysis in patients with knee osteoarthritis based on a 2-year observational study	Xie, F. et al.	2010	Singapore	observational	knee OA	533 (THA 431 & UKA 102)	**TKA**: 66.8 (7.6); 80% **UKA**: 63.3 (9.3); 75%	TKA vs. UKA	preop 6 months post 2-year post	SF-6D	Mean (95%CI): **Preop**: TKA 0.647 (0.639–0.655) UKA 0.658 (0.635–0.680) **6 months post**: TKA 0.684 (0.674–0.693) UKA 0.668 (0.646–0.688) **2-year post**: TKA 0.674 (0.664–0.683) UKA 0.681 (0.658–0.700)	no	yes
46	Unloading knee brace is a cost-effective method to bridge and delay surgery in unicompartmental knee arthritis	Lee, P. Y. F. et al.	2017	UK	observational	knee OA	63	50.9 (9.7); male 41.3%	unloading knee brace	baseline post wearing	EQ-5D	**baseline**: 0.012 (0.194) **post wearing**: 0.432 (0.291)	no	no
47	Does 3-Dimensional In Vivo Component Rotation Affect Clinical Outcomes in Unicompartmental Knee Arthroplasty	Liow, M. H. L. et al.	2016	NA	observational	knee OA	58	63.7 (9.2); male 62%	3-dimensional in vivo UKA component axial rotation	preop postop (mean 49.2 months)	EQ-5D	**preop:** 0.692 (0.234) **postop**: 0.826 (0.190)	no	yes

48	Rheumatoid arthritis does not increase risk of short-term adverse events after total knee arthroplasty A retrospective case-control study	LoVerde, Z. J. et al.	2015	NA	observational	knee OA	318	63.8 (11.2); 87.7%	TKA	baseline	EQ-5D	0.65 (0.18)	no	yes
49	Education attainment is associated with patient-reported outcomes Findings from the Swedish hip arthroplasty register	Greene, M. E et al.	2014	Sweden	observational	hip OA	11464	64 (7.9); 54%	THA	preop. 1-year postop.	EQ-5D	**preop.** 0.42 (0.31) **1-year postop**. 0.79 (0.23)	no	no
50	After patients are diagnosed with knee osteoarthritis, what do they do	Grindrod, K. A et al.	2010	Canada	observational (no treatment)	knee OA	190	63.1 (9.0); 64%	diagnosis	before diagnosis (baseline)	PAT-5D-QOL, HUI-3	PAT-5D-QOL 0.84 (0.01), HUI-3 0.71 (0.03)	no	no

51	Association of knee pain and different definitions of knee osteoarthritis with health-related quality of life A population-based cohort study	Kiadaliri AA et al.	2016	Sweden	observational (no treatment)	knee OA	1527	69.4 (7.2); 63.8%	NA	cross-sectional	EQ-5D-3L	Mean (95%CI) knee pain without knee OA: 0.71 (0.67-0.75) knee pain with knee OA: 0.67 (0.64-0.69), radiographic knee OA without knee pain: 0.77 (0.72-0.82), radiographic knee OA: 0.75 (0.71-0.79), clinical knee OA: 0.67 (0.63-0.71), clinical and radiographic knee OA: 0.66 (0.61-0.70)	no	no
52	Association of radiographic and symptomatic knee osteoarthritis with health-related quality of life in a population-based cohort study in Japan The ROAD study	Muraki, S et al.	2010	Japan	observational (no treatment)	knee OA	2126	68.9 (10.9); 64%	NA	cross-sectional	EQ-5D-3L	0.90 (0.15)	no	no
53	Quality of well-being in older people with osteoarthritis	Groessl, E. J et al.	2003	US	observational (no treatment)	mixed OA	363	69.2 (5.6); male 35.8%	NA	cross-sectional	QWB	0.643 (0.090)	no	no

54	Clinical tools association in knee osteoarthritis patients. Objective measures vs subjective measures	Martin-Martin, L et al.	2012	Spain	observational (no treatment)	knee OA	120	64.11 (10.36); 80%	NA	cross-sectional	EQ-5D	0.58 (0.23)	yes	no
55	Have cementless and resurfacing components improved the medium-term results of hip replacement for patients under 60 years of age	Jameson, S. S et al.	2015	UK	observational	hip OA	2014	**cemented best**: median age (range) 56 (37–60), 63% **cemented others**: median age (range) 57 (48–60), 72% **Hybirid best**: median age (range) 54 (30–60), 62% **Hybrid others**: median age (range) 56 (28–60), 69% **cementless best**: median age (range) 57 (39–60),52% **cementless others**: median age (range) 54 (25–60),60% **resurfacing best**: median age (range) 52 (32–60), 4% **resurfacing others**: median age (range) 54 (35–60), 72%	THR	preop. Postop.	EQ-5D	**preop**: cemented best 0.40 (0.30) cemented others 0.28 (0.32) Hybrid best 0.41 (0.32) Hybrid others 0.35 (0.33) cementless best 0.32 (0.31) cementless others 0.35 (0.32) resurfacing best 0.47 (0.31) resurfacing others 0.38 (0.34) **postop** [median(range)]: cemented best 0.81 (-0.59–1) cemented others 0.73 (-0.02–1) Hybrid best 0.81 (-0.24–1) Hybrid others 0.82 (-0.35–1) cementless best 0.80 (-0.07–1) cementless others 0.82 (-0.24–1) resurfacing best 1.00 (-0.35–1) resurfacing others 0.81 (-0.24–1)	no	no

56	Inferior Radiographic and Functional Outcomes with Modular Stem in Metal-on-Metal Total Hip Arthroplasty	Laaksonen, I et al.	2018	NA	observational	hip OA	539	mean age (range) 60.1 (23-94), male 54%	THA	post-THA	EQ-5D	median (range) 0.8 (0.0-1.0)	no	yes
57	Health-related quality of life in relation to symptomatic and radiographic definitions of knee osteoarthritis Data from Osteoarthritis Initiative (OAI) 4-year follow-up study	Tormaleh to, S et al.	2018	NA	observational (no treatment)	knee OA	4278	mean age 61.1 (9.2)	NA	baseline	SF-6D	0.801 (0.120)	no	no
58	Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement	Conner-Spady, B. L et al.	2015	Canada	observational (no treatment)	mixed OA	176	64.92 (11.34), 60%	NA	cross-sectional	EQ-5D-5L EQ-5D-3L	EQ-5D-5L: 0.49 (0.25), EQ-5D-3L: 0.46 (0.32)	no	no
59	The burden of osteoarthritis pain from patient's perspective in the United States	Jackson, J et al.	2019	US	observational (no treatment)	mixed OA	623	65.3 (11.3), 60%	opioid	cross-sectional	EQ-5D	0.72 (0.21)	yes	no

60	Health-related quality of life with vertebral fracture, lumbar spondylosis and knee osteoarthritis in Japanese men The ROAD study	Muraki, S et al.	2010	Japan	observational (no treatment)	knee OA	92	Not available	NA	cross-sectional	EQ-5D	0.87 (0.17)	no	no
61	Implant Optimisation for Primary Hip Replacement in Patients over 60 Years with Osteoarthritis	Jameson, S. S et al.	2015	UK	observational	hip OA	9149	**cemented best**: median age (range) 74 (60-93), 61.8% **cemented others**: median age (range) 75.2 (60-94), 65.9% **Hybirid best**: median age (range) 68.1 (60-91), 54.7% **Hybrid others**: median age (range) 71.6 (60-93), 63.7% **cementless best**: median age (range) 72 (60-95), 49.1% **cementless others**: median age (range) 67.8 (60-96), 57.3% **resurfacing best**: median age (range) 64.2 (60-75), 1% **resurfacing others**: median age (range) 62.8 (60-67), 86.7%	THR	preop.	EQ-5D	**female:** cemented best 0.342 (0.313) cemented others 0.319 (0.325) Hybrid best 0.432 (0.301) Hybrid others 0.356 (0.323) cementless best 0.346 (0.317) cementless others 0.366 (0.318) resurfacing best 0.516 resurfacing others 0.586 (0.192) **male:** cemented best 0.425 (0.30) cemented others 0.439 (0.288) Hybrid best 0.439 (0.288) Hybrid others 0.422 (0.302) cementless best 0.418 (0.31) cementless others 0.425 (0.311) resurfacing best 0.551 (0.253) resurfacing others 0.516	no	no

62	The excess burden of osteoarthritis in the province of Ontario, Canada	Tarride, J. E et al.	2012	Canada	observational (no treatment)	mixed OA	1474	66.4 (12.4), male 25.8%	NA	cross-sectional	HUI3	0.68 (0.29)	no	no
63	Is gain in health-related quality of life after a total hip arthroplasty depended on the comorbidity burden	Glassou, E. N et al.	2018	Denmark	observational	hip OA	1582	70 (9); 48%	THA	preop. 3 months follow-up 12 months follow-up	EQ-5D	Mena (95%CI) **preop**: 0.64 (0.63–0.65), **3 months**: 0.85 (0.64–0.86), **12 months**: 0.90 (0.89–0.90)	no	yes
64	Relationship between physical activity and health-related utility among knee osteoarthritis patients	Manheim, L. M et al.	2012	NA	observational (no treatment)	knee OA	142 (low activity 47 & medium activity 49 & high activity 46)	**Low activity**: 72.36 (11.14), 65.96% **medium activity**: 61.48 (12.89), 65.31% **high activity:** 56.54 (9.49), 45.65%	NA	baseline	SF-6D	low activity: 0.74 (0.11), medium activity: 0.76(0.10), high activity: 0.73 (0.13)	no	no
65	Validation of the Chinese (Mandarin) Version of the Oxford Knee Score in Patients with Knee Osteoarthritis	Lin, K.; Bao, L et al.	2017	China	observational	knee OA	114	67 (7); 80%	primary TKA	Pre operation	EQ-5D	0.54 ± 0.19	no	yes
66	Is the use of antidepressants associated with patient-reported outcomes following total hip replacement surgery	Greene, M. E et al.	2016	Sweden	observational	hip OA	9092	69 (9.8); 56%	THR	preop. 1-year postop.	EQ-5D	**preop**: 0.430 (0.308); **1-year postop**: 0.792 (0.228)	no	no

67	Normalization of Widespread Pressure Pain Hypersensitivity After Total Hip Replacement in Patients with Hip Osteoarthritis Is Associated with Clinical and Functional Improvements	Aranda-Villalobos, P et al.	2013	Spain	observational	hip OA	40	**Arthroplasty group:** 65 (12), male 40% **waiting list group:** 65 (13), male 45%	THA vs. waiting list	before surgery 3 months after	EQ-5D	mean (95% CI): **before surgery** arthroplasty group 0.3 (0.2-0.4), waiting list 0.3 (0.2-0.4) **3 months after:** arthroplasty group 0.7 (0.6-0.8), waiting list 0.3 (0.2-0.4)	no	yes
68	Patient-reported outcome in total hip replacement. A comparison of five instruments of health status	Ostendorf, M et al.	2004	NA	observational	hip OA	114	67.6 (10.1); 62.3%	THR	preop. 1-year after	EQ-5D	**preop**: 0.35 (0.31) **1-year after**: 0.76 (0.27)	no	yes
69	Good function after shoulder arthroplasty	Fevang, B. T. S et al.	2012	Norway	observational	shoulder OA	388**	69 (9); 63%	shoulder arthroplasty	preop. Postop.	EQ-5D	**preop**. 0.38 (0.20) **postop**.: 0.65 (0.25)	no	no
70	The impact of different rheumatic diseases on health-related quality of life	Salaffi, F et al.	2018	Italy	observational (no treatment)	mixed OA	399 (knee OA 176 & hip OA 136 &	knee OA: 69.7 (9.1); 60% hip OA: 67.4 (11.6); 58% hand OA: 66.3 (9.5); 70%	NA	cross-sectional	EQ-5D, SF-6D	knee OA: EQ-5D 0.53 (0.16) SF-6D 0.62 (0.09) hip OA: EQ-5D 0.53 (0.15) SF-6D 0.61 (0.08) hand OA: EQ-5D 0.76 (0.10) SF-6D 0.72 (0.06)	no	no

71	Patient-reported outcome is influenced by surgical approach in total hip replacement A study of the Swedish Hip Arthroplasty Register including 42 233 patients	Lindgren, J. V et al.	2014	Sweden	observational	hip OA	42233	**posterior approach:** mean age 69.3 **direct lateral approach:** mean age 68.7	posterior vs. lateral surgical approaches	preop. 1-year post 6-year post	EQ-5D	**preop**: posterior 0.42 (0.31) direct 0.42 (0.31) **1-year**: posterior 0.79 (0.23) direct 0.77 (0.24) **6-year**: posterior 0.76 (0.26) direct 0.73 (0.28)	no	no
72	Evaluation of the Effect of Heviz Mud in Patients with Hand Osteoarthritis A Randomized, Controlled, Single-Blind Follow-Up Study	Gyarmati, N et al.	2017	Hungary	trial	hand OA	47 (treated group 23 & control group 24)	**treated group**: 64.9 (4.4); 96% **control group**: 64.0 (4.7); 96%	Heviz mud	baseline at the end of 3-week treatment 16 week later	EQ-5D	**baseline**: intervention 0.687 (0.150) control 0.665 (0.154) **3 week**: intervention 0.722 (0.150) control 0.709 (0.135) **16 week**: intervention 0.722 (0.198) control 0.716 (0.136)	no	no
73	Generic and condition-specific outcome measures for people with osteoarthritis of the knee	Brazier, J. E et al.	1999	UK	observational (no treatment)	knee OA	230	**patients from TKR waiting list:** mean age (range) 71 (47-87) **patients from clinic:** mean age 64	NA	baseline	EQ-5D	patients from TKR waiting list: 0.447 (0.176), patients from rheumatology clinic: 0.46 (0.168)	no	no
74	Validation of lay descriptions of levels of severity of knee and hip oa in the khoala cohort study	Cross, M et al.	2016	France	observational (no treatment)	mixed OA	184	**no pain or difficulty**: 63.7 (7.7); 51.9% **mild pain**: 61.4 (8.6); 68.6% **moderate pain**: 61.7 (8.6); 63.0% **severe pain**: 62.3 (8.2); 68.0%	NA	cross-sectional	EQ-5D	no pain or difficulty: 0.863 (0.087) mild pain 0.773 (0.171) moderate pain 0.635 (0.239) severe pain 0.356 (0.260)	yes	no

75	Work impairment osteoarthritis and health-related quality of life among employees in Japan	Ken Nakata et al.	2018	Japan	observational (no treatment)	mixed OA	233	54.2 (12.2), 43.8%	NA	cross-sectional	SF-6D	0.69 (0.12)	no	no
76	Prevalence of knee osteoarthritis, risk factors, and quality of life The Fifth Korean National Health and Nutrition Examination Survey	Lee, S et al.	2017	Korea	observational (no treatment)	knee OA	9512 (women 5448 & men 4046)	**women**: 63.3 (0.18) **men**: 61.8 (0.18)	NA	cross-sectional	EQ-5D	**men**: Radiographic knee OA 0.91 (0.89–0.92), Symptomatic radiographic knee OA 0.78 (0.75–0.81) **women**: Radiographic knee OA 0.84 (0.83–0.85), Symptomatic radiographic knee OA 0.74 (0.72–0.76)	no	no
77	A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention	Keeley, T et al.	2015	UK	trial	knee OA	357	63.9(9.83), male 49.3%	usual physiotherapy care	baseline follow-up (at 6 months)	EQ-5D-3L	**baseline** 0.64 (0.23), **follow-up**: 0.70 (0.22)	no	no
78	Health-related quality of life in women with symptomatic hand osteoarthritis A comparison with rheumatoid arthritis patients, healthy controls, and normative data	Slatkowsky-Christensen, B et al.	2007	Norway	observational (no treatment)	hand OA	190	mean age (range) 61.6 (49.9–70.9)	NA	cross-sectional	SF-6D	0.64 (0.01)	no	no

79	A multi-centre, open label, long-term follow-up study to evaluate the benefits of a new viscoelastic hydrogel (Hymovis®) in the treatment of knee osteoarthritis	Benazzo, F et al.	2016	Italy	trial	knee OA	49	male 53.1%	Hymovis	baseline 3months 6 months 13months	EQ-5D	**baseline**: 0.62 (0.18) **3months**: 0.73 (0.07) **6 months**:0.75 (0.12) **13months:** 0.79 (0.11)	no	yes
80	Patient-reported outcomes after revision surgery compared to primary total hip arthroplasty	Postler, A. E et al.	2017	German	observational	hip OA	248 (primary THA 124 & revision THA 124)	**revision THA**: 66.4 (11.3); 70% **primary THA**: 64.1 (11); 70%	revision THA vs. primary THA	preop 6 months postop 3.6-year postop for revision 2.3-year postop for primary	EQ-5D	**preop:** revision 0.5 (0.3) primary 0.5 (0.3) **6-months postop**: revision 0.7 (0.2) primary 0.8 (0.2) **3.6-year post**: revision 0.8 (0.2) **2.3-year post**: primary 0.9 (0.2)	no	yes
81	Patients on the waiting list for total hip replacement a 1-year follow-up study	Bachrach-Lindstrom, M et al.	2008	Sweden	observational	hip OA	229 (men 117 & women 112)	**Men**: 69 (10) **women**: 70 (10)	THR	baseline one wk before surgery 1 yr after surgery	EQ-5D	**baseline**: men 0.47 (0.028), women 0.48 (0.025) **one week before**: men 0.40 (0.028), women 0.37 (0.029) **1 year after:** men 0.88 (0.018), women 0.85 (0.019)	no	yes
82	Short-Term Total Hip Arthroplasty Outcomes in Patients with Psoriatic Arthritis or Psoriatic Skin Disease Compared to Patients with Osteoarthritis	Mandl, L. A et al.	2016	NA	observational	hip OA	915	63.0 (11.1), male 45%	THA	preop 2-year postop	EQ-5D	**preop**: 0.6 (0.2) **2-year postop**: 0.8 (0.2)	no	yes

83	The impact of body mass index on patient reported outcome measures (proms) and complications following primary hip arthroplasty	Jameson, S. S et al.	2014	UK	observational	hip OA	2656	73.3 (7.7), 63.5%	priamry hip arthroplasty	preop 6 months postop	EQ-5D	**preop**: 0.368 (0.313), **6 months postop**: 0.779 (0.225)	no	yes
84	A 6-Week Web-Based osteoarthritis treatment program Observational quasi-experimental study	Nero, H et al.	2017	Sweden	trial	mixed OA	350	62 (10); 68.3%	6-week web-based OA treatment Joint Academy program	baseline follow-up (after 6 weeks)	EQ-5D-3L	**baseline**: 0.65(0.14) **follow-up**: 0.69(0.15)	no	yes
85	A pre-operative group rehabilitation programme provided limited benefit for people with severe hip and knee osteoarthritis	Wallis, J. A et al.	2014	Australia	trial	mixed OA	20	71 (7); 45%	pre-operative exercise and educational programme included self-management strategies	week 1 week 4 week 10	EQ-5D	**week 1**: 0.69 (0.18) **week 4**: 0.69 (0.19) **week 10**: 0.71 (0.15)	no	yes
86	A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis	Raynauld, J. P et al.	2002	Canada	trial	knee OA	255	Appropriate care with hylan GF 20: 62.6 (9.4), 68% Appropriate care without hylan GF 20: 63.5 (10.5), 73%	Appropriate care with hylan GF 20 vs. Appropriate care without hylan GF 20	baseline month 12	HUI3	**AC+H:** baseline 0.5 (0.22) 12 months 0.63 (0.25) **AC:** baseline 0.46 (0.24) 12 months 0.51 (0.28)	no	yes

87	The Influence of Radiological Severity and Symptom Duration of Osteoarthritis on Postoperative Outcome After Total Hip Arthroplasty A Prospective Cohort Study	Al-Amiry, B. S et al.	2018	Sweden	observational	hip OA	222	68 (9.8); 49%	unilateral THA	preop 1 year postop	EQ-5D	**preop**: KL 1-2 0.40 (0.3) KL 3-4 0.50 (0.7); **1 year postop**: KL 1-2 0.83 (0.20) KL 3-4 0.86 (0.17)	no	yes
88	Health-related quality of life and mobility of patients awaiting elective total hip arthroplasty a prospective study	Mahon JL et al.	2002	UK	observational	hip OA	99	68.0 (7.9); 51%	THA	baseline 6 months postop	HUI3	**baseline:** short waits 0.50 (0.18) long waits 0.53(0.23) **6 months post**: short waits 0.71(0.18) long waits 0.71 (0.17)	no	yes
89	A theory of planned behavior-based intervention to improve quality of life in patients with knee hip osteoarthritis a randomized controlled trial	Saffari, M et al.	2018	Iran	trial	mixed OA	120	55.8 (8.9), 75.8%	a theory of planned behaviour-based intervention vs. standard care	baseline 3 months follow-up	EQ-5D-3L	**trial group:** baseline 0.38 (0.33) follow-up 0.66 (0.13) **control group**: baseline 0.37 (0.35) follow-up 0.53 (0.28)	no	yes
90	Waiting for total hip arthroplasty Avoidable loss in quality time and preventable deterioration	Ostendorf, M et al.	2004	NA	observational	hip OA	161	68.4 (9.7); male 34%	THA	preop 3 months postop 1 year postop	EQ-5D	**preop**: 0.33 (0.32) **3 months postop**: 0.71 (0.26) **1 year postop**: 0.75 (0.28)	no	yes

91	Comparative responsiveness of generic health outcome measures at 3 and 12 months following 8 weeks of supervised patient education and exercise therapy	Roos, E. M et al.	2018	Denmark	trial	mixed OA	2904	mean age 64.2; 75%	8-week of supervised patient education and exercise therapy	baseline 3 months 12 months	SF-6D, EQ-5D	**Baseline**: SF-6D 0.727 (0.137); EQ-5D 0.718 (0.103) **3 months**: SF-6D 0.777 (0.137); EQ-5D 0.756 (0.114) **12 months**: SF-6D 0.774 (0.142); EQ-5D 0.761 (0.130)	yes	yes
92	Three-year follow-up study of health related QOL and lifestyle indicators for Japanese patients after total hip arthroplasty	Fujita, K et al.	2016	Japan	observational	hip OA	576	61.6 (9.8); 85%	THA	preop 6 week postop 1 year postop 3 years postop	EQ-5D	**preop**: 0.57 (0.13) **6 week postop**: 0.74 (0.16) **1 year postop**: 0.84 (0.16) **3 years postop**: 0.87 (0.16)	no	yes
93	Quality of life and hip function during the first month after total hip arthroplasty	Uhrbrand, P et al.	2014	Denmark	observational	hip OA	32	56.3 (2.0); 53%	THA	preop 31 days postop	EQ-5D	**preop:** 0.63 (0.02) **31 days postop**: 0.78 (0.03)	no	yes
94	Women in Charnley class C fail to improve in mobility to a higher degree after total hip replacement	Gordon, M et al.	2014	Sweden	observational	hip OA	26249	**women**: 70 (10) **men**: 67 (10)	THR	preop 1 year postop	EQ-5D	**preop**: women 0.72 (0.12) men 0.75 (0.11) **1 year postop**: women 0.87 (0.11) men 0.89 (0.10)	no	no

95	Patient-reported outcomes in the Swedish Hip Arthroplasty Register Results of a nationwide prospective observational study	Rolfson, O et al.	2011	Sweden	observational	hip OA	34960	68.1 (10.4); 58%	THR	preop 1 year postop	EQ-5D	**preop**: 0.41 (0.31), **1 year postop**: 0.78 (0.24)	no	yes
96	An 8-week knee osteoarthritis treatment program of hyaluronic acid injection, deliberate physical rehabilitation, and patient education is cost effective at 2 years follow-up	Miller, L. E et al.	2014	US	trial	knee OA	553	71 (10); male 51%	8-week treatment program of Hainjection, deliberate physical rehabilitation, patient education	baseline	EQ-5D	0.701 (0.051)	no	no
97	An evaluation of a new strengthening and exercise programme that aims to improve the symptoms of knee osteoarthritis by goal setting, using strength to bodyweight ratios	Creasey, J et al.	2017	NA	trial	knee OA	37	60 (7.5)	6-week strengthening and exercise program	week1 week 12	EQ-5D-5L	Median (IQR): **week 1**: 0.6 (0.5-0.7) **week 12**: 0.7 (0.7-0.8)	yes	yes

98	Cross-cultural adaptation and validation of the Spanish version of the Oxford Hip Score in patients with hip osteoarthritis	Martin-Fernandez, J et al.	2017	Spain	observational (no treatment)	hip OA	361	Mean(95%CI) 67.8 (66.7-69.1), 53.2%	observational	cross-sectional	EQ-5D-5L	0.52 (95% CI 0.49-0.55)	no	no
99	Determinants of health status among patients with knee or hip osteoarthritis The role of demographic, clinical and health related quality of life variables	Saffari, M et al.	2014	Iran	observational (no treatment)	knee OA, hip OA	356	63.0 (12.3), male 9.3%	NA	cross-sectional	EQ-5D	all: 0.47 (0.34) knee OA: 0.49 (0.33) hip OA: 0.32 (0.33)	no	no
100	Cost-effectiveness of a supplementary class-based exercise program in the treatment of knee osteoarthritis	Richardson, G et al.	2006	UK	trial	knee OA	214 (home based group 103 &class-based group 111)	**home-based group**: 64.9 (9.7), 60.2% **class-based group**: 64.5 (9.9), 56.8%	home-based exercise program vs. home-based exercise supplemented with an 8-week class-based exercise program	baseline 1 month 6 months 12 months	EQ-5D-3L	**baseline**: home-based 0.50 (0.03), class-based 0.54 (0.03) **1 month**: home based 0.52 (0.03), class-based 0.60 (0.02) **6 months**: home-based 0.54 (0.03), class-based 0.58 (0.02) **12 months**: home-based 0.53 (0.03), class-based 0.58 (0.02)	no	yes

101	Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee	Whitehurst, D. G et al.	2011	UK	trial	knee OA	352	mean age 63, 61%	advice and exercise (AE), advice and exercise plus true acupuncture (AE+TA), advice and exercise plus nonpenetrating acupuncture (AE+NPA)	baseline 6-week 6 months 12 months	EQ-5D	**baseline:** AE 0.617 (0.24), AE+TA 0.568 (0.28) **6-week**: AE 0.648 (0.25), AE+TA 0.682 (0.21) **6-months**: AE 0.635 (0.25), AE+TA 0.636 (0.28) **12-months**: AE 0.620 (0.29), AE+TA 0.635 (0.29)	no	yes
102	Cost-Utility Analysis of High Molecular Weight Hyaluronic Acid for Knee Osteoarthritis in Everyday Clinical Care	Hermans, J et al.	2018	Netherlands	trial	knee OA	156 (intervention group 77 & control group 79)	**intervention group**: 53.6 (8.6), 48% **control group**: 54.8 (6.4), 51%	3 weekly intraarticular injections with Hylan G-F 20 (Sanofi) added to usual care (intervention), or they received usual care only (control)	baseline	EQ-5D-3L	intervention: 0.68 (0.23), control: 0.71 (0.24)	no	yes
103	Cost-utility analysis of interventions to improve effectiveness of exercise therapy for adults with knee osteoarthritis The BEEP trial	Kigozi, J et al.	2018	UK	trial	knee OA	514	mean age 63, 51%	individually tailored exercise (ITE), targeted exercise adherence (TEA) vs. usual physical therapy care (UC)	baseline 3 months 6 months 9 months 18 months	EQ-5D-3L	**baseline**: UC 0.636 (0.230), ITE 0.644 (0.229), TEA 0.629 (0.229) **3 months**: UC 0.686 (0.201) ITE 0.708 (0.188) TEA 0.669 (0.227) **6 months**: UC 0.690 (0.225) ITE 0.692 (0.215) TEA 0.692 (0.217) **9 months**: UC 0.698 (0.217) ITE 0.665 (0.249) TEA 0.702 (0.199) **18 months:** UC 0.700 (0.219) ITE 0.700 (0.206) TEA 0.682 (0.232)	no	yes

104	Dextrose Prolotherapy for Symptomatic Knee Osteoarthritis Feasibility, Acceptability, and Patient-Oriented Outcomes in a Pilot-Level Quality Improvement Project	Rabago, D et al.	2019	NA	trial	knee OA	7	59.6 (9.3); 85.7%	Intra- and extra-articular prolotherapy injections	baseline 8 months follow up	EQ-5D	**baseline**: 0.65 (0.15) **8 months**: 0.82 (0.09)	no	yes
105	Cost-effectiveness of a model consultation to support self-management in patients with osteoarthritis	Oppong, R et al.	2018	UK	trial	mixed OA	525	67.3 (10.4), 59.5%	model OA consultation vs. usual care	baseline 3 months 6 months 12 months	EQ-5D-3L SF-6D	**baseline:** model OA consultation EQ-5D 0.573 (0.298), SF-6D 0.678 (0.139); usual care EQ-5D 0.588 (0.272), SF-6D 0.690 (0.148) **3 months:** model OA consultation EQ-5D 0.615 (0.280), SF-6D 0.688 (0.141); usual care EQ-5D 0.631 (0.264), SF-6D 0.696 (0.141) **6 months**: model OA consultation EQ-5D 0.637 (0.264), SF-6D 0.687 (0.142); usual care EQ-5D 0.638 (0.259), SF-6D 0.707 (0.144) **12 months:** model OA consultation EQ-5D 0.651 (0.262), SF-6D 0.693 (0.139); usual care EQ-5D 0.674 (0.224), SF-6D 0.702 (0.138)	no	yes
106	Do modern total knee replacements offer better value for money A health economic analysis	Hamilton, D. F et al.	2013	NA	trial	knee OA	124 (Kinemax 60 & Triathlon 64)	**Kinemax**: 68.9 (9.8), male 46.7 % **Triathlon**: 69.1 (9.6), male 40.1 %	TKA with Kinemax prosthesis vs. TKA with Triathlon	preop. 6 week postop. 6 months post 1-year post	SF-6D	**preop**: Kinemax 0.631 (0.114) Triathlon 0.623 (0.127) **6 week post**: Kinemax 0.670 (0.143) Triathlon 0.719 (0.116) **6 months post**: Kinemax 0.771 (0.128) Triathlon 0.778 (0.136) **1-year post**: Kinemax 0.773 (0.123) Triathlon 0.766 (0.138)	no	yes

107	Effect of inpatient rehabilitation vs a monitored home-based program on mobility in patients with total knee arthroplasty the HIHO randomized clinical trial	Buhagiar, M. A et al.	2017	Australia	trial	knee OA	212 (165 for RCT, 87 for observational group)	68.4 (9.3), 58%	10 days of inpatient rehabilitation followed by a monitored home-based program after TKA vs. home-based program alone observational: home-based alone	baseline 10 weeks 26 weeks 52 weeks	EQ-5D	**baseline**: Inpatient Rehabilitation 0.39 (0.26), Home Program 0.36 (0.28), Observational 0.37 (0.29) **10 weeks** mean (95%CI): Inpatient Rehabilitation 0.74 (0.69-0.78), Home Program 0.69 (0.65-0.74) **26 weeks**: Inpatient Rehabilitation 0.74 (0.70-0.78), Home Program 0.72 (0.68-0.77), observational 0.72 (0.68-0.77) **52 weeks**: Inpatient Rehabilitation 0.70 (0.66-0.75), Home Program 0.73 (0.69-0.78)	no	yes
108	Cost-Utility and Cost-Effectiveness Analyses of Face-to-Face Versus Telephone-Based Nonpharmacologic Multidisciplinary Treatments for Patients with Generalized Osteoarthritis	Cuperus, N et al.	2016	Netherlands	trial	mixed OA	147	60 (8), 85%	face-to-face treatment program vs. telephone-based treatment program	baseline 6 weeks 13 weeks 26 weeks 36 weeks 52 weeks	EQ-5D-3L TRS SF-6D	**baseline**: FTF EQ-5D 0.57 (0.25), TRS 0.70 (0.20) , SF-6D 0.66 (0.10); T-Based EQ-5D 0.57 (0.26) , TRS 0.69 (0.16), SF-6D 0.68 (0.07) **6 weeks**: FTF EQ-5D 0.63 (0.24), TRS 0.78 (0.12) , SF-6D 0.69 (0.09); T-Based EQ-5D 0.58 (0.26), TRS 0.75 (0.11), SF-6D 0.68 (0.08) **13 weeks**: FTF EQ-5D 0.61 (0.27), TRS0.76 (0.12); T-Based EQ-5D 0.58 (0.32), TRS 0.75 (0.18), **26 weeks**: FTF EQ-5D 0.59 (0.30), TRS 0.76 (0.14) , SF-6D 0.69 (0.09); T-Based EQ-5D 0.59 (0.26), TRS 0.73 (0.13), SF-6D 0.68 (0.09) **39 weeks**: FTF EQ-5D 0.59 (0.30), TRS 0.75 (0.20); T-Based EQ-5D 0.59 (0.31), TRS 0.74 (0.16) **52 weeks**: FTF EQ-5D 0.59 (0.25), TRS 0.74 (0.15) , SF-6D 0.70 (0.11); T-Based EQ-5D 0.57 (0.26), TRS 0.73 (0.13), SF-6D 0.67 (0.10)	no	yes

109	Effectiveness and safety of tapentadol prolonged release with tapentadol immediate release on-demand for the management of severe, chronic osteoarthritis-related knee pain	Steigerwald, I et al.	2012	NA	trial	knee OA	200	67.4 (10.81), 67.5%	tapentadol prolonged release vs. tapentadol immediate release	baseline 6 weeks 8weeks 12 weeks	EQ-5D	**baseline**: 0.42 (0.30) **6 weeks:** 0.66 (0.203) **8 weeks**: 0.67 (0.223) **12 weeks**: 0.69 (0.247)	no	yes
110	Economic evaluation of telephone-based weight loss support for patients with knee osteoarthritis A randomised controlled trial	O'Brien, K. M et al.	2018	Australia	trial	knee OA	119 (intervention group 59 & control group 60)	**intervention group**: 63.0 (11.1), male 34% **control group**: 60.2 (13.9), male 42%	telephone-based weight management vs. usual care	baseline	SF-6D	intervention 0.6 (0.1), control 0.7 (0.1)	no	yes
111	Effect on health-related quality of life of a multimodal physiotherapy program in patients with chronic musculoskeletal disorders	Cuesta-Vargas, A. I et al.	2013	Spain	trial	mixed OA	42	Not available	8-week multimodal physical therapy program	baseline post	EQ-5D	**baseline**: 0.55 (0.32) **post** 0.75 (0.28)	no	yes
112	Impact of co-morbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis	Hosseini, K et al.	2012	France	observational (no treatment)	knee OA, hip OA	878 (hip OA 222 & knee OA 607 & hip and knee OA 49)	**hip OA**: 61.31 (8.76), male 33% **knee OA**: 62.07 (8.47), male 29% **hip and knee OA**: 64.67 (7.99), male 35%	NA	cross-sectional	SF-6D	hip: 0.66 (0.11), knee: 0.66 (0.11), hip and knee: 0.63 (0.11)	no	no

113	Effects of early combined eccentric-concentric versus concentric resistance training following total knee arthroplasty	Suh, M. J et al.	2017	Korea	trial	knee OA	34	71 (6.1), male 12%	ECC-CON following primary TKA vs. CON following primary TKA	before operation 1 month after operation	EQ-5D	**before operation**: ECC-CON 0.6 (0.1), CON 0.6 (0.2) **1 month after**: ECC-CON 0.8 (0.1), CON 0.8 (0.1)	no	yes
114	Adaptation and validation of the Osteoarthritis Knee and Hip Quality of Life (OAKHQOL) questionnaire for use in patients with osteoarthritis in Spain	Gonzalez Saenz de Tejada, M et al.	2011	Spain	observational (no treatment)	knee OA, hip OA	759 (knee OA 409 & hip OA 350)	**knee OA**: 71.26 (7.71), 73.8% **hip OA**: 68.13 (10.61), 50.0%	NA	baseline	EQ-5D	VAS rates: knee 0.449 (0.203) hip 0.408 (0.213) TE rates: knee 0.397 (0.319) hip 0.335 (0.339)	no	no
115	Efficacy and safety of tapentadol prolonged release for moderate to severe chronic osteoarthritis knee pain a pooled analysis of two double blind	Lange, B et al.	2017	12 European countries and US, Canada, New Zealand,	trial	knee OA	2010 (placebo 674 & Tapentadol 663 & Oxycodone 673)	placebo: 60.2 (9.5); 67.8% Tapentadol: 60.3 (9.9); 67.4% Oxycodone: 59.9 (9.9); 62.6%	prolonged release tapentadol vs. controlled release Oxycodone	baseline	EQ-5D	placebo 0.4 (0.3), tapentadol 0.4 (0.3), oxycodone 0.4 (0.3)	no	yes
116	Evaluation of the effect of Lake Heviz thermal mineral water in patients with osteoarthritis of the knee a randomized, controlled, single-blind, follow-up study	Kulisch, A et al.	2014	Hungary	trial	knee OA	77 (thermal water group 38 & control group 39)	**thermal water group**: 65.6 (6.4); male 21% **control group**: 65.5 (7.7); male 23%	Lake Heviz thermal mineral water	baseline	EQ-5D	intervention 0.6281 (0.2406) control 0.6005 (0.2035)	no	no

117	Economic evaluation of aquatic exercise for persons with osteoarthritis	Patrick, D. L et al.	2001	US	trial	mixed OA	249 (treatment group 125 & control group	**Treatment group**: mean age 65.7, 85.3% **control group**: mean age 66.1, 87.1%	20-wk aquatic class vs. usual activities pattern	baseline post class	QWB	**baseline:** treatment 0.597 (0.068), control 0.599 (0.065) **post class**: treatment 0.606 (0.069), control 0.599 (0.079)	no	yes
118	Five-year results of a randomised controlled trial comparing mobile and fixed bearings in total knee replacement	Breeman, S et al.	2013	UK	trial	knee OA	539 (mobile bearing group 276 & fixed bearing group 263)	**mobile bearing group**: 69(8), male 38.7% **fixed bearing group**: 69(9), male 41.1%	mobile bearing vs. fixed bearing in TKR	baseline 3 months postop 1-year post 2-year post 3-year post 4-year post 5-year post	EQ-5D	**baseline**: mobile bearing 0.32 (0.32) fixed bearing 0.34 (0.31) **3 months postop**: mobile bearing 0.66 (0.28) fixed bearing 0.66 (0.24) **1-year post**: mobile bearing 0.69 (0.31) fixed bearing 0.67 (0.30) **2-year post**: mobile bearing 0.67 (0.32) fixed bearing 0.64 (0.30) **3-year post**: mobile bearing 0.66 (0.32) fixed bearing 0.62 (0.34) **4-year post**: mobile bearing 0.60 (0.36) fixed bearing 0.59 (0.35) **5-year post**: mobile bearing 0.59 (0.36) fixed bearing 0.61 (0.34)	no	yes
119	Knee joint distraction compared with high tibial osteotomy randomized controlled trial	Van Der Woude, AJ. A et al.	2017	Netherlands	trial	knee OA	69 (HTO group 45 & KJD group 22)	**HTO group**: 49.4 (1.0), male 60% **KJD group**: 51.2 (1.1), male 73%	HTO vs. KJD	baseline 3 months 6 months 12 months	EQ-5D	**baseline**: HTO 0.64 (0.2), KJD 0.63 (0.2) **3 months**: HTO 0.68 (0.2), KJD 0.52 (0.3) **6 months**: HTO 0.68 (0.3), KJD 0.69 (0.2) **12 months**: HTO 0.79 (0.3), KJD 0.77 (0.1)	no	yes

120	Economic evaluation of arthritis self-management in primary care	Patel, A et al.	2009	UK	trial	mixed OA	812 (intervention group 406 & control group 406)	**intervention group**: 68.4(8.2), male 37% **control group**: 68.7 (8.6), male 37%	six sessions of an arthritis self-management programme plus an education booklet (intervention group) vs. education booklet alone (standard care control group)	baseline 4 months 12 months	EQ-5D	**baseline:** intervention 0.570 (0.25) control 0.535 (0.28) **4 months**: intervention 0.552 (0.28) control 0.556 (0.27) **12 months:** intervention 0.578 (0.25) control 0.559 (0.27)	no	yes
121	One-year follow-up of mud-bath therapy in patients with bilateral knee osteoarthritis a randomized, single-blind controlled trial	Fioravanti A et al.	2015	Italy	trial	knee OA	103 (MBT group 53 & control group 50)	**MBT group**: 68.49 (9.01), male 43% **control group**: 69.66 (11.1), male 12%	Mud-bath vs. usual care	baseline	EQ-5D	MBT 0.46 (0.31) control 0.37 (0.36)	no	no
122	The effect of sulphurous water in patients with osteoarthritis of hand. Double-blind, randomized, controlled follow-up study	Kovacs, C et al.	2012	Hungary	trial	hand OA	45 (intervention group 24 & control group 21)	**intervention group**: mean age (rang) 58 (47-71) **control group**: mean age (rang) 61 (50–73)	Spa water vs. tap water	baseline 3 weeks 3 months 6 months	EQ-5D	**baseline**: spa water 0.481 (0.206) tap water 0.470 (0.208) **3 weeks**: spa water 0.570 (0.226) tap water 0.475 (0.201) **3 months**: spa water 0.510 (0.181) tap water 0.429 (0.19) **6 months**: spa water 0.495 (0.168) tap water 0.418 (0.192)	no	no
123	Effect of an education programme for patients with osteoarthritis in primary care - A randomized controlled trial	Hansson, E. E et al.	2010	Sweden	trial	mixed OA	114 (intervention group 61 & control group 53)	**intervention:** mean age 62 (9.43) **control:** mean age 63 (9.51)	education program vs. usual care	baseline	EQ-5D	intervention 0.58 (0.25), control 0.56 (0.30)	no	yes

124	Group education and exercise is feasible in knee and hip osteoarthritis	Skou, S. T et al.	2012	Denmark	trial	mixed OA	39	mean (range) 59.3 (56-65), 86%	education and neuromuscular exercise	baseline 3 months post	EQ-5D	**baseline** 0.779 (0.086) **3 months post** 0.825 (0.104)	no	no
125	Pain Catastrophizing Is Independently Associated with Quality of Life in Patients with Severe Hip Osteoarthritis	Hayashi, K et al.	2018	Japan	observational (no treatment)	hip OA	70	mean (range) 68 (60-75), 89%	NA	cross-sectional	EQ-5D	0.59 (0.48-0.65)	no	no
126	Factors associated with patients' willingness to consider joint surgery after completion of a digital osteoarthritis treatment program A prospective cohort study	Cronstrom, A et al.	2018	Sweden	trial	mixed OA	458	62 (5.6), 67.8%	6-wks digital non-surgical OA treatment program comprising education, exercise and asynchronous chat	baseline	EQ-5D	0.64 (0.2)	no	yes
127	Better early functional outcome after short stem total hip arthroplasty	van Oldenrijk, J et al.	2017	Netherlands	trial	hip OA	150 (short stem 75 & conventional stem 75)	**short stem**: 60.3 (6.8); male 28% **conventional stem**: 60.5 (7.1); male 29%	different THA: the Collum Femoris Preserving short stem vs. the Zweymuller Alloclassic conventional stem.	baseline 6weeks after 3 months after 6 months after 1 year after 2 years	EQ-5D	**short stem** (95%CI): baseline 0.62 (0.57-0.68), 6 weeks 0.72 (0.68-0.76), 3 months 0.82 (0.79-0.85), 6 months 0.86 (0.83-0.89), 1 year 0.88 (0.85-0.91), 2 years 0.87 (0.84-0.91) **conventional stem** (95%CI): baseline 0.59 (0.54-0.65), 6 weeks 0.76 (0.73-0.79), 3 months 0.81 (0.78-0.85), 6 months 0.85 (0.82-0.89), 1 year 0.85 (0.80-0.89), 2 years 0.86 (0.82-0.91)	no	yes
128	The long-term outcome of simple trapeziectomy	Yeoman, T. F. M et al.	2019	NA	observational	hand OA	205	mean (range) 66 (46-87), men 10%	simple trapeziectomy	preop postop (mean 8.2 years after)	EQ-5D	**preop**: 0.50 (0.24) **post**: 0.56 (0.31)	no	no

129	Pharmacist-initiated intervention trial in osteoarthritis A multidisciplinary intervention for knee osteoarthritis	Marra, C. A et al.	2012	Canada	trial	knee OA	139 9usual care 66 & intervention care 73)	**usual care**: 60.8 (7.2), male 44% **intervention care**: 62.7 (9.2), male 42%	multidisciplinary intervention vs. usual care	baseline	HUI3	**usual care**: 0.679 (0.253), **intervention** 0.750 (0.170)	no	yes
130	Effectiveness of exercise therapy added to general practitionercare in patients with hip osteoarthritis a pragmatic randomizedcontrolled trial	Teirlinck, C et al.	2016	Netherlands	trial	hip OA	203 (GP+ET group 101 & GP group 102)	**GP+ET**: 64 (8.5), 62% **GP**: 67 (9.6), 55%	exercise therapy added to general practitioner care vs. general practitioner care	baseline 6 weeks 3 months 6 months 9 months 12 months	EQ-5D	**baseline**: GP+ET 0.778 (0.122) GP 0.748 (0.161) **6 weeks**: GP+ET 0.788 (0.126) GP 0.756 (0.177) **3 months**: GP+ET 0.780 (0.162) GP 0.777 (0.147) **6 months**: GP+ET 0.771 (0.187) GP 0.759 (0.174) **9 months**: GP+ET 0.781 (0.176) GP 0.763 (0.197) **12 months**: GP+ET 0.784 (0.198) GP 0.784 (0.151)	no	yes
131	Physical Therapist-Delivered Pain Coping Skills Training and Exercise for Knee Osteoarthritis Randomized Controlled Trial	Bennell, K. L et al.	2016	Australia	trial	knee OA	222 (exercise 75 & PCST 74 & PCST/exercise 73)	**exercise**: 62.7 (7.9), 59% **PCST**: 63.0 (7.9), 61% **PCST/exercise**: 64.6 (8.3), 60%	PCST/exercise vs. exercise or PCST alone	week 0 week 12 week 32 week 52	AQoL-6D	**week 0**: exercise 0.71 (0.14) PCST 0.71 (0.16) PCST/exercise 0.74 (0.12) **week 12**: exercise 0.78 (0.17) PCST 0.78 (0.15) PCST/exercise 0.80 (0.15) **week 32**: exercise 0.76 (0.15) PCST 0.79 (0.16) PCST/exercise 0.84 (0.12) **week 52**: exercise 0.78 (0.16) PCST 0.81 (0.12) PCST/exercise 0.84 (0.13)	no	yes
132	Pilot study of massage in veterans with knee osteoarthritis	Juberg, M et al.	2015	US	trial	knee OA	25	56.96 (11.98), men 68%	8 weeks massage	pre post	EQ-5D-5L	**pre**: 0.60 (0.16) **post**: 0.68 (0.15)	no	no
133	Platelet-rich plasma (PRP) therapy for knee arthritis a feasibility study in primary care	Glynn, L. G et al.	2018	Ireland	trial	knee OA	12	72.6 (10.4), male 58%	Platelet-rich plasma	baseline 4 months post	EQ-5D-3L	**baseline**: 0.45 (0.19) **4 months post**: 0.77 (0.25)	no	yes

134	Validity and reliability of the Swedish version of the Patient Specific Functional Scale in patients treated surgically for carpometacarpal joint osteoarthritis	Rosengren, J et al.	2013	Sweden	observational (no treatment)	hand OA	58	median (range) 62 (48-82), 83%	NA	cross-sectional	EQ-5D	median (range) 0.7 (0-1)	no	no
135	Prospective, randomized, double-blinded, double-dummy and multicenter phase IV clinical study	Ha, C. W et al.	2016	Korea	trial	knee OA	124 (experimental group 63 & control group 61)	**experimental group**: 63.8 (7.2), male 12.7% **control group**: 65.4 (7.0), male 11.5%	PG201 vs. SKI306X	baseline 8 weeks 12 weeks	EQ-5D	**baseline**: experimental 0.73 (0.11), control 0.72 (0.14) **8 weeks**: experimental 0.76 (0.12), control 0.76 (0.10) **12 weeks**: experimental 0.77 (0.12), control 0.79 (0.11)	no	yes
136	The effect of education and supervised exercise on physical activity, pain, quality of life and self-efficacy - an intervention study with a reference group	Jonsson, T et al.	2018	NA	trial	mixed OA	264 (intervention group 195 & control group 69)	**intervention group**: 60 (10), 33% **control group**: 66 (7), 84%	education and supervised exercise vs. standard care	baseline 3 months 12 months	EQ-5D	median (IQR): **baseline** intervention 0.73 (0.62–0.80) control 0.66 (0.16–0.73) **3 months** intervention 0.73 (0.69–0.80) control 0.52 (0.09–0.73) **12 months** intervention 0.73 (0.69–0.80)	no	yes
137	Spironolactone for People Age 70 Years and Older with Osteoarthritic Knee Pain A Proof-of-Concept Trial	McMurdo, M. E. T et al.	2016	Scotland	trial	knee OA	86 (Spironolactone group 43 & placebo group 43)	**Spironolactone group**: 77.4 (4.8), 60% **placebo group:** 76.1 (5.2), 63%	Spironolactone vs. placebo	baseline	EQ-5D-3L	Spironolactone 0.68 (0.19) placebo 0.60 (0.28)	no	yes

138	STICKS study - Short-sTretch Inelastic Compression bandage in Knee Swelling following total knee arthroplasty - a feasibility study	Brock, T. M et al.	2017	UK	trial	knee OA	49 (control group 25 & compression group 24)	**control group**: 69.5 (6.8), 64% **compression group**: 67.3 (8.2), 67%	short-stretch, inelastic compression bandage after TKA vs. standard wool and crepe bandage after TKA	preop 6 months postop	EQ-5D	**preop**: control 0.554 (0.270) compression 0.570 (0.240) **6 months post**: control 0.651 (0.331) compression 0.812 (0.183)	no	yes
139	Sub-vastus approach is more effective than a medial parapatellar approach in primary total knee arthroplasty A randomized controlled trial	Bridgma nn, S. A et al.	2009	NA	trial	knee OA	231 (sub-vastus group 116 & medial parapatellar 115)	**sub-vastus group**: 70.1 (8.0), male 52% **medial parapatellar**: 70.9 (8.1), male 51%	sub-vastus approach to TKA vs. medial parapatellar approach	baseline 1 week 6 week 12 week 52 week	EQ-5D	**baseline**: sub-vastus 0.27 (0.27) medial parapatellar 0.29 (0.27) **1 week**: sub-vastus 0.42 (0.19) medial parapatellar 0.35 (0.19) **6 week**: sub-vastus 0.61 (0.26) medial parapatellar 0.58 (0.24) **12 week**: sub-vastus 0.77 (0.24) medial parapatellar 0.73 (0.29) **52 week**: sub-vastus 0.87 (0.21) medial parapatellar 0.80 (0.26)	no	yes
140	Effects of sulfur bath on hip osteoarthritis a randomized, controlled, single-blind, follow-up trial a pilot study	Kovacs, C et al.	2016	Hungary	trial	hip OA	41 (balneotherapy group 21 & control group 20)	**balneotherapy group**: 59.14 (7.55), **control group**: 60.66 (7.6)	combination of balneotherapy and home exercise therapy vs. home exercise alone	baseline 3 week 12 week	EQ-5D	**baseline**: Balneotherapy 0.483 (0.218), control 0.483 (0.219) **3 week**: Balneotherapy 0.645 (0.206), control 0.595 (0.215) **12 week**: Balneotherapy 0.637 (0.196), control 0.514 (0.216)	no	no

141	Accelerated perioperative care and rehabilitation intervention for hip and knee replacement is effective A randomized clinical trial involving 87	Larsen, K et al.	2008	Denmark	trial	knee OA, hip OA	87 (intervention group 45 & control group 42)	**intervention group**: 64 (10.8), 56% **control group**: 66 (9.2), 45%	accelerated perioperative care and rehabilitation vs. current intervention	baseline 3 months post	EQ-5D	**THA patients accelerated group**: baseline 0.45 (0.30), 3 months: 0.88 (0.17) **THA patients control group**: baseline 0.49 (0.22), 3 months: 0.76 (0.23) **TKA patients accelerated group**: baseline 0.44 (0.24), 3 months 0.86 (0.11) **TKA patients control group**: baseline 0.60 (0.22), 3 months: 0.86 (0.09) **UKA patients accelerated group**: baseline 0.67 (0.05), 3 months: 0.85 (0.21) **UKA patients control group**: baseline 0.66 (0.22), 3 months: 0.80 (0.06)	no	yes
142	Patient-reported outcome after total hip arthroplasty comparison between lateral and posterior approach	Rosenlund, S et al.	2017	Denmark	trial	hip OA	80 (lateral approach 38 & posterior approach 39)	**lateral approach**: 60 (7), male 68% **posterior approach**: 62 (6), male 67%	lateral approach THA vs. posterior approach THA	preop.	EQ-5D-3L	**lateral approach** 0.6 (0.2) **posterior approach** 0.6 (0.2)	no	yes
143	Better management of patients with osteoarthritis - evidence based education and exercise delivered nationwide in Sweden	Jonsson, T. S et al.	2019	Sweden	trial	knee OA, hip OA	46935	Not available	BOA program offered hip and knee OA information and individually adapted exercise program	baseline 3 months 12 months	EQ-5D	**knee OA** baseline 0.636 (0.220), 3 months 0.702 (0.197); 12 months 0.692 (0.198) **hip OA** baseline 0.611 (0.232), 3 months 0.654 (0.225); 12 months 0.650 (0.218)	yes	yes

144	The effect of Neydharting mud-pack therapy on knee osteoarthritis A randomized, controlled, double-blind follow-up pilot study	Tefner, I. K et al.	2013	Hungary	trial	knee OA	53 (intervention group 27 & control group 26)	**intervention group**: 63.42 (8.86) **control group**: 63.55 (9.53)	hot mud pack vs. control	baseline 2 week 6 week 12 week	EQ-5D	**baseline:** hot-mud 0.49 (0.22), control 0.56 (0.17) **2 week**: hot-mud 0.63 (0.25), control 0.63 (0.196) **6 week**: hot-mud 0.70 (0.267), control 0.63 (0.20) **12 week**: hot-mud 0.72 (0.247), control 0.66 (0.161)	no	no
145	Topical (intra-articular) tranexamic acid reduces blood loss and transfusion rates following total knee replacement A randomized controlled trial (TRANX-K)	Alshryda, S et al.	2013	UK	trial	knee OA	157 (placebo 78 & Traneexamic acid group 79)	**placebo**: 67.1 (10.2), male 56% **Traneexamic acid**: 65.5 (9.6), male 38%	Tranexamic acid for TKR vs. placebo	preop. 3 months postop	EQ-5D-3L	**preop**: Placebo 0.431 (0.33), Tranexamic Acid 0.377 (0.31) **3 months postop**: Placebo 0.780 (0.24), Tranexamic Acid 0.705 (0.31)	no	yes
146	Total Knee Arthroplasty Using Bicruciate-Stabilized or Posterior-Stabilized Knee Implants Provided Comparable Outcomes at 2 Years	Scarvell, J. M et al.	2017	Australia	trial	knee OA	240 (Posterior Cruciate-Stabilized Implant Group 124 & Bicruciate-Stabilized	Posterior Cruciate-Stabilized Implant: 66.80 (7.49), 49% Bicruciate-Stabilized Implant: 66.66 (7.54), 53%	Posterior Cruciate-Stabilized Implant vs. Bicruciate-Stabilized Implant for TKA	1 year postop 2 year postop	EQ-5D	**1 year postop**: PCS 0.91 (SE 0.01) BCS 0.89 (SE 0.01) **2 year postop**: PCS 0.89 (SE 0.01) BCS 0.88 (SE 0.01)	no	yes
147	Stable migration pattern of an ultra-short anatomical uncemented hip stem	Mahmoud, A. N et al.	2017	Sweden	trial	hip OA	25	mean (range) 51.4 (16.7-68.2), male 52%	Proxima stem during THA	preop 1-year post	EQ-5D	**preop** 0.47 (0.31) **1-year post** 0.90 (0.17)	no	yes

148	Western medical acupuncture in a group setting for knee osteoarthritis Results of a pilot randomised controlled trial	White, A et al.	2016	UK	trial	knee OA	60 (group acupuncture 20 & individual acupuncture 20 & standard 20)	**group acupuncture**: 64.7 (7.7), male 50% **individual acupuncture**: 65.1 (9.9), male 60% **standard:** 64.9 (10.8), male 50%	standardised advice and exercise booklet, booklet plus group acupuncture, booklet plus individual acupuncture	baseline 14 week	EQ-5D	**baseline**: group acupuncture 0.545 (0.255), individual acupuncture 0.480 (0.250), standard 0.555 (0.274) **14 week**: group acupuncture 0.639 (0.308), individual acupuncture 0.660 (0.227), standard 0.560 (0.271)	no	yes
149	Customized foot insoles have a positive effect on pain, function, and quality of life in patients with medial knee osteoarthritis	Skou, S. T et al.	2013	Denmark	trial	knee OA	42	mean age 63, 48%	Customized foot insoles	baseline after intervention	EQ-5D-3L	**baseline**: 0.59 (0.22) **after intervention**: 0.79 (0.18)	no	no
150	Total hip arthroplasty through the mini-incision (Micro-hip) approach versus the standard transgluteal (Bauer) approach A prospective, randomised study	Dienstknecht, T et al.	2014	NA	trial	hip OA	143 (Micro-hip 55 & Bauer 88)	**Micro-hip**: 61.9 (12.1), 60% **Bauer:** 61.3 (11.6), 53%	THA through the mini-incision approach vs. standard transgluteal approach	preop 6 week postop 3 months postop	EQ-5D	**preop:** micro 0.473 (0.235) Bauer 0.466 (0.253) **6 week postop**: micro 0.847 (0.167) Bauer 0.810 (0.169) **3 months postop**: micro 0.850 (0.216) Bauer 0.845 (0.230)	no	yes
151	The effects of inpatient versus outpatient spa therapy on pain, anxiety, and quality of life in elderly patients with generalized osteoarthritis a pilot study	Ozkuk, K et al.	2018	Turkey	trial	mixed OA	150 (Etodolac group 50 & outpatient spa 50 & inpatient spa 50)	**Etodolac group**: 69.2 (3.9), 74% **outpatient spa**: 69.0 (3.5), 64% **inpatient spa**: 69.4 (3.3), 66%	Etodolac, outpatient spa, inpatient spa	baseline 2 weeks 6 weeks	EQ-5D	**baseline**: Etodolac 0.37 (0.03), outpatient spa 0.39 (0.03), inpatient spa 0.36 (0.03) **2 weeks:** Etodolac 0.54 (0.03), outpatient spa 0.71 (0.03), inpatient spa 0.85 (0.03) **6 weeks:** Etodolac 0.51(0.03), outpatient spa 0.67 (0.03), inpatient spa 0.87 (0.03)	no	no

Note: 15D=15-dimensional questionnaire; AQoL-6D=Assessment of Quality of Life-6 dimension; BCS: Bicruciate-Stabilized Implant; EQ-5D=EuroQol-5 dimension; FTF: face-to-face treatment program; GP: general practitioner care; GP+ET: exercise therapy added to general practitioner care; HSUVs=health state utility values; HUI=health utility index; HTO= high tibial osteotomy; IQR: interquartile range; KJD=knee joint distraction; NSAIDs=Nonsteroidal anti-inflammatory drugs; NA=not available; OA=osteoarthritis; PCS: Posterior Cruciate-Stabilized Implant; PCST: Pain Coping Skills training; preop=pre-operation; postop=post-operation; QWB=quality of well-being; SD=standard deviation; SF-6D=Short-Form Six-Dimension; SG=standard gamble; T-Based: telephone-based treatment program; TRS=the rating scale; TJR=total joint replacement; THR/THA=total hip replacement/arthroplasty; TKR/TKA=total knee replacement/arthroplasty. Mixed OA including a variety of OA patients without specifying their affected OA joint site.

Supplement 4.2.2 Details of HSUVs included in the meta-analyses for knee OA core interventions.

					Baseline			1 month			6-week			12-week			14-week			6-month			32-week			9-month			12-month			18-month		
No.	Title	Authors/ Year published	Interventions	Elicitation Method	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
1	An evaluation of a new strengthening and exercise programme that aims to improve the symptoms of knee osteoarthritis by goal setting, using strength to bodyweight ratios	Creasey, J et. al, 2017	strengthening and exercise program	EQ-5D-5L	32	0.60	0.16							29	0.736	0.078																	
2	Better management of patients with osteoarthritis - evidence based education and exercise delivered nationwide in Sweden	Jonsson, T. S et. al, 2019	information and individually adapted exercise program	EQ-5D	29,865	0.64	0.22							29,865	0.702	0.197													20,199	0.692	0.198			
3	Cost-effectiveness of a supplementary class-based exercise program in the treatment of knee osteoarthritis	Richardson, G et. al, 2006	home-based exercise program	EQ-5D-3L	103	0.50	0.30	103	0.52	0.30										103	0.54	0.30							103	0.53	0.30			
			home-based exercise supplemented with class-based exercise program	EQ-5D-3L	111	0.54	0.32	111	0.6	0.21										111	0.58	0.21							111	0.58	0.21			

4	Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee	Whitehurst, D. G et. al, 2011	advice and exercise (AE)	EQ-5D	92	0.62	0.24				92	0.648	0.25				92	0.635	0.25				92	0.62	0.29
			advice and exercise plus true acupuncture (AE+TA)	EQ-5D	92	0.57	0.28				92	0.682	0.21				92	0.636	0.28				92	0.635	0.29
5	Cost-utility analysis of interventions to improve effectiveness of exercise therapy for adults with knee osteoarthritis The BEEP trial	Kigozi, J et. al, 2018	usual physical therapy care (UC)	EQ-5D-3L	175	0.64	0.23				175	0.686	0.201				175	0.69	0.225				175	0.698	0.217				175	0.7	0.219
			individually tailored exercise (ITE)	EQ-5D-3L	176	0.64	0.23				176	0.708	0.188				176	0.692	0.215				176	0.665	0.249				176	0.7	0.206
			targeted exercise adherence (TEA)	EQ-5D-3L	163	0.63	0.23				163	0.669	0.227				163	0.692	0.217				163	0.702	0.199				163	0.682	0.232
6	Economic evaluation of telephone-based weight loss support for patients with knee osteoarthritis A randomised controlled trial	O'Brien, K. M et. al, 2018	telephone-based weight management	SF-6D	59	0.60	0.10																							
			usual care	SF-6D	60	0.70	0.10																							

7	Pharmacist-initiated intervention trial in osteoarthritis A multidisciplinary intervention for knee osteoarthritis	Marra, C. A et. al, 2012	multidisciplinary intervention	HUI3	66	0.68	0.25												
			usual care	HUI3	73	0.75	0.17												
8	Physical Therapist-Delivered Pain Coping Skills Training and Exercise for Knee Osteoarthritis Randomized Controlled Trial	Bennell, K. L et. al, 2016	Exercise alone	AQoL-6D	75	0.71	0.14				75	0.78	0.17				75	0.76	0.15				75	0.78	0.16
			Physical Therapist-Delivered Pain Coping Skills Training (PCST) alone	AQoL-6D	74	0.71	0.16				74	0.78	0.15				74	0.79	0.16				74	0.81	0.12
			PCST and Exercise	AQoL-6D	73	0.74	0.12				73	0.8	0.17				73	0.84	0.14				73	0.84	0.15
9	Western medical acupuncture in a group setting for knee osteoarthritis Results of a pilot randomised controlled trial	White, A et. al, 2016	booklet plus group acupuncture	EQ-5D	20	0.55	0.26				20	0.639	0.308												
			booklet plus individual acupuncture	EQ-5D	20	0.48	0.25				20	0.66	0.227												

			standardised advice and exercise booklet	EQ-5D															
					20	0.56	0.27									20	0.56	0.271	

Note: *The highlighted data were not included in the meta-analysis as Post-treatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups). OA=osteoarthritis; HSUVs=health state utility values; AQoL-6D=Assessment of Quality of Life-6 dimension; EQ-5D=EuroQol-5 dimension; SF-6D=Short-Form Six-Dimension; HUI=Health utility Index; SD=standard deviation. N means the number of participants.*

Supplement 4.2.3 Details of HSUVs included in the meta-analyses for knee OA medication treatments.

					baseline			3-week			6-week			8-week			12-week			6-month		
No.	Title	Authors/ year published	Intervention	Elicitation method	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD
1	Cost-utility analysis and economic burden of knee osteoarthritis treatment the analysis from the real clinical practice	Turajane, T et al, 2012	traditional NSAIDs	EQ-5D	939	-0.079	0.19													939	0.595	0.12
			Celecoxib	EQ-5D	380	-0.11	0.18													380	0.602	0.12
2	Effectiveness and safety of tapentadol prolonged release with tapentadol immediate release on-demand for the management of severe, chronic osteoarthritis-related knee pain	Steigerwald, I et al, 2012	tepentadol	EQ-5D	195	0.42	0.3				159	0.66	0.203	153	0.67	0.223	125	0.69	0.247			
3	Efficacy and safety of tapentadol	Lange, B et al., 2017	placebo	EQ-5D	674	0.4	0.3															

	prolonged release for moderate to severe chronic osteoarthritis knee pain a pooled analysis of two double blind		tapentadol	EQ-5D	663	0.4	0.3							
			oxycodone	EQ-5D	673	0.4	0.3							
4	Prospective, randomized, double-blinded, double-dummy and multicenter phase IV clinical study	Ha, C. W et al., 2016	PG201	EQ-5D	63	0.73	0.11				63	0.76	0.12	63	0.77	0.12
			SKI306X	EQ-5D	61	0.72	0.14				61	0.76	0.1	61	0.79	0.11
5	Spironolactone for People Age 70 Years and Older with Osteoarthritic Knee Pain A Proof-of-Concept Trial	McMurdo, M. E. T et al., 2016	Spironolactone	EQ-5D	43	0.68	0.19									
			placebo	EQ-5D	43	0.6	0.28									
6	The effects of adherence to non-Steroidal anti-inflammatory drugs and factors influencing drug adherence in patients with knee osteoarthritis	Park, K. K et al., 2016	NSAIDs	EQ-5D	1334	0.71	0.2	1334	0.76	0.1						

Note: *OA=osteoarthritis; HSUVs=health state utility values; EQ-5D=EuroQol-5 dimension; SD=standard deviation; NSAIDs=Nonsteroidal anti-inflammatory drugs. N means the number of participants.*

Supplement 4.2.4 Details of HSUVs included in the meta-analyses for knee OA injection treatments.

	baseline	3-month	4-month	6-month	8-month	12-month	13-month

No	Title	Authors/ Year published	Interventions	Elicitation method	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD
1	A multi-centre, open label, long-term follow-up study to evaluate the benefits of a new viscoelastic hydrogel (Hymovis®) in the treatment of knee osteoarthritis	Benazzo, F et al., 2016	viscoelastic hydrogel	EQ-5D	49	0.62	0.18	49	0.73	0.07				49	0.75	0.12							49	0.79	0.11
2	A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis	Raynauld, J. P et al., 2002	Appropriate care with hylan GF 20	HUI3	123	0.5	0.22													122	0.63	0.25			
			Appropriate care without hylan GF 21	HUI3	126	0.46	0.24													107	0.51	0.28			
3	Cost-Utility Analysis of High Molecular Weight Hyaluronic Acid for Knee Osteoarthritis in Everyday Clinical Care	Hermans, J et al., 2018	hylan GF 20 with usual care	EQ-5D	77	0.68	0.23																		
			usual care	EQ-5D	79	0.71	0.24																		
4	Dextrose Prolotherapy for Symptomatic Knee Osteoarthritis Feasibility, Acceptability, and	Rabago, D et al., 2019	Intra- and extra-articular prolotherapy injections	EQ-5D	7	0.65	0.15										7	0.82	0.09						

	Patient-Oriented Outcomes in a Pilot-Level Quality Improvement Project																		
5	Platelet-rich plasma (PRP) therapy for knee arthritis a feasibility study in primary care	Glynn, L. G et al., 2018	Platelet-rich plasma	EQ-5D	12	0.45	0.19				12	0.77	0.25						

Note: *The highlighted data were not included in the meta-analysis as Post-treatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups). OA=osteoarthritis; HSUVs=health state utility values; EQ-5D=EuroQol-5 dimension; HUI=Health utility Index; SD=standard deviation. N means the number of participants.*

Supplement 4.2.5 Details of HSUVs included in the meta-analyses for knee OA primary surgery treatments

					baseline	1-week	1-month	6-week	10-week	12-week	3-month	6-month	1-year	18-month	2-year	33-year	4-years	5-year	49-month	7-year
No.	Title	Authors/Year published	Intervention	Elicitation Method	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD	n/mean/SD
1	An exploratory study of response shift in health-	Zhang, X. H et al., 2012	TKA	SF-6D	71/0.62/0.008							62/0.68/0.07		71/0.77/0.18						

	related quality of life and utility assessment among patients with osteoarthritis		TKA	EQ-5D	68/0.52/0.42	68/0.86/0.22
2	Comparative outcomes and cost-utility following surgical treatment of focal lumbar spinal stenosis compared with osteoarthritis of the hip or knee	Rampersaud, Y. R et al.,2014	TKA	SF-6D	99/0.55/0.08	99/0.78/0.15

3	Comparison of lifetime incremental cost utility ratios of surgery relative to failed medical management for the treatment of hip, knee and spine osteoarthritis modelled	Tso, P et al.,2012	TKA	SF-6D			99/0.64/0.15
4	Comparing the validity and responsiveness of the EQ-5D-5L to the Oxford hip and knee scores and SF-12 in osteoarthritis patients 1 year following total joint replacement	Conner-Spady, B. L et al.,2018	TKA	EQ-5D	268/0.39/0.27		

5	Cost effectiveness and quality of life in knee arthroplasty	Lavernia, C. J et al.,1997	TKA	QWB	100/0.56/0.05		45/0.56/0.08	60/0.57/0.09	52/0.59/0.08	18/0.59/0.07
6	Cost-Effectiveness of five commonly used prosthesis brands for total knee replacement in the UK A study using the NJR Dataset	Pennington, M et al.,2016	TKA	EQ-5D	13635/0.38/0.31					
			TKA	EQ-5D	5005/0.41/0.31					
			TKA	EQ-5D	3364/0.39/0.31					
			TKA	EQ-5D	4187/0.41/0.31					

			TKA	EQ-5D	3585/0.40/0.31				
7	Determining Cost-Effectiveness of Total Hip and Knee Arthroplasty Using the Short Form-6D Utility Measure	Elmallah, R. K et al.,2017	TKA	SF-6D	844/0.62/0.10			844/0.77/0.11	
8	Do modern total knee replacements offer better value for money A health economic analysis	Hamilton, D. F et al.,2013	TKA	SF-6D	60/0.63/0.11	60/0.67/0.14	60/0.77/0.13	60/0.77/0.12	
			TKA	SF-6D	64/0.62/0.13	64/0.72/0.12	64/0.78/0.14	64/0.77/0.14	

9	Does patella resurfacing really matter Pain and function in 972 patients after primary total knee arthroplasty An observational study from the Norwegian Arthroplasty Register	Lygre, S. H. L et al.,2010	TKA	EQ-5D	972/0.46/0.22				
10	Early Clinically Relevant Improvement in Quality of Life and Clinical Outcomes 1 Year Postsurgery in Patients with Knee and Hip Joint Arthroplasties	Neuprez, A et al.,2018	TKA	EQ-5D	280/0.46/0.23		280/0.66/0.2	280/0.69/0.19	280/0.67/0.22

11	Effectiveness of hip or knee replacement surgery in terms of quality adjusted life years and costs	Räsänen, P et al.,2007	TKA	15D	103/0.81/0.09			103/0.83/0.11	103/0.84/0.11				
12	Effects of early combined eccentric-concentric versus concentric resistance training following total knee arthroplasty	Suh, M. J et al.,2017	TKA	EQ-5D	16/0.60/0.10	16/0.80/0.10							
			TKA	EQ-5D	18/0.60/0.20	18/0.80/0.10							
13	Five-year results of a randomised controlled trial comparing	Breeman, S et al.,2013	TKA	EQ-5D	276/0.32/0.32		276/0.66/0.28		276/0.69/0.31	276/0.67/0.32	276/0.66/0.32	276/0.60/0.36	276/0.59/0.36

	mobile and fixed bearings in total knee replacement		TKA	EQ-5D	262/0.34/0.31		262/0.66/0.24	262/0.67/0.30	262/0.64/0.30	262/0.62/0.34	262/0.59/0.35	262/0.61/0.34
14	Outcome of total hip arthroplasty, but not of total knee arthroplasty, is related to the preoperative radiographic severity of osteoarthritis	Tilbury, C et al.,2016	TKA	EQ-5D	74/0.60/0.30							
			TKA	EQ-5D	197/0.60/0.30							
15	Patient-reported outcomes after total and unicompartmental knee arthroplasty A study of 14 076 matched patients from the national joint registry for EngLand and Wales	Liddle, A. D et al.,2015	UKA	EQ-5D	3519/0.48/0.29							

16	Predicting the Long-Term Gains in Health-Related Quality of Life After Total Knee Arthroplasty	Schilling, C. G et al.,2017	TKA	SF-6D	488/0.57/0.10	488/0.71/0.15		488/0.69/0.15
17	Quality of Life and Cost-Effectiveness 1 Year After Total Hip Arthroplasty	Lavernia, C. J, Alcerro, J.C et al.,2011	TKA	EQ-5D	5398/0.68/0.17		5398/0.84/0.15	
			UKA	EQ-5D	240/0.71/0.14		240/0.87/0.13	
			TKA	SF-6D	5398/0.63/0.11		5398/0.76/0.13	5398/0.79/0.13

			UKA	SF-6D	240/0.65/0.11							240/0.80/0.13	240/0.80/0.14
18	Quality of life benefits of knee arthroplasty for osteoarthritis	Neuprez, A; Francois, G et al.,2014	TKA	EQ-5D	279/0.46/0.23				279/0.66/0.20	279/0.68/0.19			
19	Sub-vastus approach is more effective than a medial parapatellar approach in primary total knee arthroplasty A randomized controlled trial	Bridgman, S. A et al.,2009	TKA	EQ-5D	116/0.27/0.27	116/0.42/0.19	116/0.61/0.26	116/0.77/0.24			116/0.87/0.21		
			TKA	EQ-5D	115/0.29/0.27	115/0.35/0.19	115/0.58/0.24	115/0.73/0.29			115/0.80/0.26		
20	The effects of age on patient-reported outcome measures in	Williams, D. P et al.,2013	TKA	EQ-5D	108/0.36/0.32					108/0.69/0.30		108/0.76/0.24	

total knee replacements		TKA	EQ-5D	489/0.37/0.33		489/0.73/0.24	489/0.73/0.29
		TKA	EQ-5D	9370.43/0.32		937/0.74/0.24	937/0.76/0.27
		TKA	EQ-5D	792/0.45/0.31		792/0.74/0.22	792/0.74/0.24
		TKA	EQ-5D	130/0.42/0.31		130/0.70/0.26	130/0.66/0.27
21	Topical (intra-articular) tranexamic acid reduces blood loss	Alshryda, S et al.,2013	TKA	EQ-5D	78/0.43/0.33	78/0.78/0.24	

	and transfusion rates following total knee replacement A randomized controlled trial (TRANX-K)		TKA	EQ-5D	79/0.38/0.31		79/0.71/0.31		
22	Total or partial knee replacement Cost-utility analysis in patients with knee osteoarthritis based on a 2-year observational study	Xie, F et al.,2010	TKA	SF-6D	431/0.65/0.09		431/0.68/0.10	431/0.67/0.10	
			UKA	SF-6D	102/0.66/0.12		102/0.67/0.12	102/0.68/0.11	
23	Does 3-Dimensional In Vivo Component Rotation Affect	Liow, M. H. L et al.,2016	TKA	EQ-5D	58/0.69/0.23				58/0.83/0.19

	Clinical Outcomes in Unicompartmental Knee Arthroplasty		TKA	EQ-5D	58/0.78/0.16				58/0.87/0.13
24	Rheumatoid arthritis does not increase risk of short-term adverse events after total knee arthroplasty A retrospective case-control study	LoVerde, Z. J et al.,2015	TKA	EQ-5D	318/0.65/0.18				
25	Total Knee Arthroplasty Using Bicruciate-Stabilized or Posterior-Stabilized Knee Implants Provided Comparable Outcomes at 2 Years	Scarvell, J. M et al.,2017	TKA	EQ-5D			124/0.91/0.11	124/0.89/0.11	
			TKA	EQ-5D			116/0.89/0.11	116/0.88/0.11	

26	Effect of inpatient rehabilitation vs a monitored home-based program on mobility in patients with total knee arthroplasty the HIHO randomized clinical trial	Buhagiar, M. A et al.,2017	TKA	EQ-5D	81/0.39/0.26	79/0.74/0.19	80/0.74/0.17	80/0.70/0.19
			TKA	EQ-5D	84/0.36/0.28	78/0.69/0.20	80/0.72/0.21	80/0.73/0.21
			TKA	EQ-5D	87/0.37/0.29		87/0.72/0.21	
27	Evaluation of 1031 primary titanium nitride coated mobile bearing total knee arthroplasties in an orthopedic clinic	Breugem, S. J. M et al.,2017	TKA	EQ-5D				663/0.88/0.16
			revision TKA	EQ-5D				8/0.61/0.38

28	Knee joint distraction compared with high tibial osteotomy A randomized controlled trial	Van Der Woude, J. A et al.,2017	HTO	EQ-5D	45/0.64/0.20			45/0.68/0.20	45/0.68/0.30	45/0.79/0.30
			KJD	EQ-5D	22/0.63/0.20			22/0.52/0.30	22/0.69/0.20	22/0.77/0.10
29	STICKS study - Short-sTretch Inelastic Compression bandage in Knee Swelling following total knee arthroplasty - a feasibility study	Brock, T. M et al.,2017	TKA	EQ-5D	25/0.55/0.27				25/0.65/0.33	
			TKA	EQ-5D	24/0.57/0.24				24/0.81/0.18	
30	Accelerated perioperative care and rehabilitation intervention for hip and	Larsen, K et al.,2008	TKA	EQ-5D	15/0.44/0.24			15/0.86/0.11		

	knee replacement is effective A randomized clinical trial involving 87		TKA	EQ-5D	12/0.60/0.22	12/0.86/0.09
			UKA	EQ-5D	2/0.67/0.05	2/0.85/0.21
			UKA	EQ-5D	2/0.66/0.22	2/0.80/0.06
31	Validation of the Chinese (Mandarin) Version of the Oxford Knee Score in Patients with Knee Osteoarthritis	Lin,K et al.,2017	TKA	EQ-5D	114/0.54/0.19	

Note: OA=osteoarthritis; HSUVs=health state utility values; 15 D=15-dimension questionnaire; EQ-5D=EuroQol-5 dimension; SF-6D=Short-Form Six-Dimension; HUI=Health utility Index; QWB=quality of well-being; TKA= total knee arthroplasty; UKA= Unicompartmental Knee arthroplasty; KJD=Knee joint distraction; HTO= high tibial osteotomy. n/mean/SD=number of participants/mean HSUVs/standard deviation of HSUVs.

Supplement 4.2.6 Details of HSUVs included in the meta-analyses for hip OA core interventions.

					baseline			3-week			6-week			12-week			6-month			9-month			12-month		
No.	Title	authors	intervention	Elicitation method	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD
1	Better management of patients with osteoarthritis - evidence based education and exercise delivered nationwide in Sweden	Jonsson, T. S. et al., 2019	information and individually adapted exercise program	EQ-5D	13570	0.611	0.232							13570	0.654	0.225							8320	0.65	0.218
2	Effectiveness of exercise therapy added to general practitioner care in patients with hip osteoarthritis a pragmatic randomized controlled trial	Teirlinck, C et al., 2016	exercise therapy + general practitioner care	EQ-5D	101	0.778	0.122				101	0.788	0.126	101	0.78	0.162	101	0.771	0.187	101	0.781	0.176	101	0.784	0.198
			general practitioner care	EQ-5D	102	0.748	0.161				102	0.756	0.177	102	0.777	0.147	102	0.759	0.174	102	0.763	0.197	102	0.784	0.151

Note: *The highlighted data were not included in the meta-analysis as Post-treatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups). OA=osteoarthritis; HSUVs=health state utility values; EQ-5D=EuroQol-5 dimension; SD=standard deviation. N means the number of participants.*

Supplement 4.2.7 Details of HSUVs included in the meta-analyses for hip OA primary surgery treatments.

				preop			1-month			6-week			3-month			6-month			1-year			2-year			3-year			5-year			6-year			13-year		
No.	Title	Authors/ Year published	Intervention	Elicitation	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD
1	A clinical and radiographic 13 year follow up study of 138 Charnley hip arthroplasties in patients 50-70 years old Comparison of university hospital	Hulleberg, G et al, 2008	THA	EQ-5D																													89	0.75	0.24
2	Association between changes in global femoral offset after total hip arthroplasty	Mahmood, S. S et al, 2016	THA	EQ-5D	71	0.44	0.26													71	0.82	0.19														

	and function, quality of life, and abductor muscle strength		THA	EQ-5D	73	0.43	0.22											73	0.86	0.17			
			THA	EQ-5D	78	0.51	0.66											78	0.86	0.19			
3	Better early functional outcome after short stem total hip arthroplasty	van Oldenrijk, J et al, 2017	THA	EQ-5D	75	0.62	0.24	75	0.72	0.177	75	0.82	0.133	75	0.86	0.13	75	0.88	0.13	75	0.87	0.16
			THA	EQ-5D	75	0.59	0.24	75	0.76	0.133	75	0.81	0.155	75	0.85	0.16	75	0.85	0.20	75	0.86	0.20
4	Body Mass Index Class Is Independently Associated with Health-Related Quality of Life After Primary Total Hip Arthroplasty an Institutional Registry-Based Study	McLawhorn, A. S et al, 2017	THA	EQ-5D-3L	2733	0.64	0.21														2733	0.89	0.16

5	Changes in the WOMAC, EuroQol and Japanese lifestyle measurements among patients undergoing total hip arthroplasty	Fujita, K et al. 2009	THA	EQ-5D	451	0.56	0.13	451	0.74	0.16				451	0.79	0.17						
6	Comparative outcomes and cost-utility following surgical treatment of focal lumbar spinal stenosis compared with osteoarthritis of the hip or knee	Rampersaud, Y. R et al. 2014	THA	SF-6D	99	0.52	0.22										99	0.72	0.32	99	0.75	0.16
7	Comparing Short Form 6D, Standard Gamble, and Health Utilities Index Mark 2 and Mark 3 utility scores Results from total hip arthroplasty patients	Feeny, D et al. 2012	THA	SG	63	0.61	0.33				63	0.76	0.25									
			THA	HUI2	63	0.59	0.10				63	0.69	0.11									
			THA	HUI3	63	0.55	0.20				63	0.76	0.15									
			THA	SF-6D	63	0.49	0.21				63	0.72	0.18									

8	Comparing the validity and responsiveness of the EQ-5D-5L to the Oxford hip and knee scores and SF-12 in osteoarthritis patients 1 year following total joint replacement	Conner-Spady, B. L et al. 2018	THA	EQ-5D-5L	269	0.35	0.25									
9	Determining Cost-Effectiveness of Total Hip and Knee Arthroplasty Using the Short Form-6D Utility Measure	Elmallah, R. K et al. 2017	THA	SF-6D	224	0.61	0.13							224	0.80	0.13
10	Early Clinically Relevant Improvement in Quality of Life and Clinical Outcomes 1 Year Postsurgery in Patients with Knee and Hip Joint Arthroplasties	Neuprez, A et al. 2018	THA	EQ-5D	346	0.44	0.23	346	0.71	0.24	346	0.74	0.25	346	0.71	0.27
11	Effectiveness of hip or knee replacement surgery in terms of quality adjusted life years and costs	Räsänen, P et al. 2007	THA	15D	96	0.81	0.08				96	0.87	0.09	96	0.86	0.12
			revision THA	15D	24	0.81	0.08				24	0.84	0.09	24	0.82	0.10

12	Inferior Radiographic and Functional Outcomes with Modular Stem in Metal-on-Metal Total Hip Arthroplasty	Laaksonen, I et al. 2018	THA	EQ-5D										539	0.79	0.16
13	Is gain in health-related quality of life after a total hip arthroplasty depended on the comorbidity burden	Glassou, E. N et al. 2018	THA	EQ-5D	1582	0.64	0.20	1582	0.85	2.232	1582	0.90	0.10			
14	Normalization of Widespread Pressure Pain Hypersensitivity After Total Hip Replacement in Patients with Hip Osteoarthritis Is Associated with Clinical and Functional Improvements	Aranda-Villalobos, P et al. 2013	THA	EQ-5D	20	0.30	0.21	20	0.7	0.214						
15	Patient-reported outcome after total hip arthroplasty comparison between lateral and posterior approach	Rosenlund, S et al. 2017	THA	EQ-5D-3L	38	0.60	0.20									
			THA	EQ-5D-3L	39	0.60	0.20									

16	Patient-reported outcome in total hip replacement. A comparison of five instruments of health status	Ostendorf, M et al. 2004	THA	EQ-5D	114	0.35	0.31							114	0.76	0.27				
17	Patient-reported outcomes after revision surgery compared to primary total hip arthroplasty	Postler, A. E et al. 2017	revision THA	EQ-5D	124	0.50	0.30				124	0.7	0.2				124	0.8	0.2	
			THA	EQ-5D	124	0.50	0.30				124	0.8	0.2		124	0.9	0.2			
18	Outcome of total hip arthroplasty, but not of total knee arthroplasty, is related to the preoperative radiographic severity of osteoarthritis	Tilbury, C et al. 2016	THA	EQ-5D	77	0.60	0.20													
			THA	EQ-5D	225	0.60	0.30													
19	Patients on the waiting list for total hip replacement a 1-year follow-up study	Bachrach-Lindstrom, M et al. 2008	THA	EQ-5D	117	0.40	0.03							117	0.88	0.02				
			THA	EQ-5D	112	0.37	0.03							112	0.85	0.02				

20	Quality of Life and Cost-Effectiveness 1 Year After Total Hip Arthroplasty	Lavernia, C. J et al. 2011	THA	EQ-5D	5463	0.63	0.19						5463	0.88	0.14				
			hip resurfacing	EQ-5D	843	0.68	0.16						843	0.93	0.11				
			THA	SF-6D	5463	0.60	0.11						5463	0.79	0.13		5463	0.81	0.13
			hip resurfacing	SF-6D	843	0.63	0.10						843	0.82	0.12		843	0.84	0.12
21	Short-Term Total Hip Arthroplasty Outcomes in Patients with Psoriatic Arthritis or Psoriatic Skin Disease Compared to Patients with Osteoarthritis	Mandl, L. A et al. 2016	THA	EQ-5D	915	0.60	0.20						915	0.8	0.2				
22	Stable migration pattern of an ultra-short anatomical uncemented hip stem	Mahmoud, A. N et al. 2017	THA	EQ-5D	25	0.47	0.31						25	0.9	0.17				

23	The impact of body mass index on patient reported outcome measures (proms) and complications following primary hip arthroplasty	Jameson, S. S et al. 2014	THA	EQ-5D	2656	0.37	0.31							2656	0.78	0.23			
24	The Influence of Radiological Severity and Symptom Duration of Osteoarthritis on Postoperative Outcome After Total Hip Arthroplasty A Prospective Cohort Study	Al-Amiry, B. S et al. 2018	THA	EQ-5D	73	0.40	0.30										73	0.83	0.2
			THA	EQ-5D	149	0.50	0.70										149	0.86	0.17
25	Total hip arthroplasty through the mini-incision (Micro-hip) approach versus the standard transgluteal (Bauer) approach A prospective, randomised study	Dienstknecht, T et al. 2014	THA	EQ-5D	55	0.47	0.24			55	0.85	0.17	55	0.85	0.22				
			THA	EQ-5D	88	0.47	0.25			88	0.81	0.17	88	0.85	0.23				
26	Health-related quality of life and mobility of patients awaiting elective total hip arthroplasty a prospective study	Mahon JL et al. 2002	THA	HUI3	63	0.50	0.18							63	0.71	0.18			
			THA	HUI3	36	0.53	0.23							36	0.71	0.17			

27	Waiting for total hip arthroplasty Avoidable loss in quality time and preventable deterioration	Ostendorf, M; Buskens, E et al. 2004	THA	EQ-5D	161 0.33 0.32			161 0.71 0.26	161 0.75 0.28		
28	Three-year follow-up study of health related QOL and lifestyle indicators for Japanese patients after total hip arthroplasty	Fujita, K et al. 2016	THA	EQ-5D	576 0.57 0.13		576 0.74 0.16		576 0.84 0.16		576 0.87 0.16
29	Quality of life and hip function during the first month after total hip arthroplasty	Uhrbrand, P et al. 2014	THA	EQ-5D	32 0.63 0.02	32 0.78 0.03					
30	Patient-reported outcomes in the Swedish Hip Arthroplasty Register Results of a nationwide prospective observational study	Rolfson, O et al. 2011	THA	EQ-5D	34960 0.41 0.31				34960 0.78 0.24		
31	Comparison of lifetime incremental cost utility ratios of surgery relative to failed medical management for the treatment of hip, knee and spine osteoarthritis modelled	Tso, P et al. 2012	THA	SF-6D	99 0.52 0.22					99 0.72 0.32	

32	Accelerated perioperative care and rehabilitation intervention for hip and knee replacement is effective A randomized clinical trial involving 87	Larsen, K et al. 2008	THA	EQ-5D	28	0.45	0.30							28	0.88	0.17
			THA	EQ-5D	28	0.49	0.22							28	0.76	0.23

Note: *OA=osteoarthritis; HSUVs=health state utility values; EQ-5D=EuroQol-5 dimension; SF-6D=Short-Form Six-Dimension; HUI=Health utility Index; 15 D= 15-dimensional questionnaire; SG=standard gamble; SD=standard deviation; THA=total hip arthroplasty. N means the number of participants.*

Supplement 4.2.8 Details of HSUVs included in the meta-analyses for mixed OA core interventions

					baseline			4-week			6-week			8-week			10-week			13-week			4-month			20-week			6-month			36-week			39-week			12-month		
No.	Title	Authors/published year	Intervention	Elicitation	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD
1	A 6-Week Web-Based osteoarthritis treatment program Observational quasi-experimental study	Nero, H et al., 2017	web-based OA treatment Joint Academy program	EQ-5D	235	0.65	0.14				235	0.69	0.15																												

2	A pre-operative group rehabilitation programme provided limited benefit for people with severe hip and knee osteoarthritis	Wallis, J. A et al., 2014	pre-operative exercise and educational programme	EQ-5D	20	0.69	0.18	20	0.69	0.19	20	0.71	0.15						
3	A theory of planned behavior-based intervention to improve quality of life in patients with knee hip osteoarthritis a randomized controlled trial	Saffari, M et al., 2018	behaviour-based intervention	EQ-5D	60	0.38	0.33							53	0.66	0.13			
			standard care	EQ-5D	60	0.37	0.35							54	0.53	0.28			
4	An innovative care model coordinated by a physical therapist and nurse practitioner for osteoarthritis of the hip and knee in specialist care a prospective study	Voorn, V. M et al., 2013	care model coordinated by a physical therapist and nurse practitioner	EQ-5D	87	0.48	0.30				63	0.59	0.29						
5	Comparative responsiveness of generic health outcome measures	Roos, E. M et al., 2018	education and exercise therapy	SF-6D	2904	0.73	0.14							2904	0.78	0.14	2904	0.77	0.14

	at 3 and 12 months following 8 weeks of supervised patient education and exercise therapy		education and exercise therapy	EQ-5D	2904	0.72	0.10						2904	0.76	0.11							2904	0.76	0.13
			model consultation	EQ-5D	288	0.57	0.30						288	0.62	0.28				288	0.64	0.26		288	0.65	0.26
6	Cost-effectiveness of a model consultation to support self-management in patients with osteoarthritis	Oppong, R et al., 2018	model consultation	SF-6D	288	0.68	0.14						288	0.69	0.14				288	0.69	0.14		288	0.69	0.14
			usual care	EQ-5D	237	0.59	0.27						237	0.63	0.26				237	0.64	0.26		237	0.67	0.22
			usual care	SF-6D	237	0.69	0.15						237	0.70	0.14				237	0.71	0.14		237	0.70	0.14
7	Cost-Utility and Cost-Effectiveness Analyses of Face-to-Face Versus Telephone-Based Nonpharmacologic Multidisciplinary Treatments for	Cuperus, N et al., 2016	Face-to-Face nonpharmacologic multidisciplinary treatments	EQ-5D	75	0.57	0.25		75	0.63	0.24		75	0.61	0.27		75	0.59	0.30		75	0.59	0.3	75	0.59	0.25
			Face-to-Face nonpharmacologic multidisciplinary treatments	TRS	75	0.70	0.20		75	0.78	0.12		75	0.76	0.12		75	0.76	0.14		75	0.75	0.2	75	0.74	0.15

	Patients with Generalized Osteoarthritis		Face-to-Face nonpharmacologic multidisciplinary treatments	SF-6D	75 0.66 0.10	75 0.69 0.09				75 0.69 0.09		75 0.70 0.11
			telephone-based nonpharmacologic multidisciplinary treatments	EQ-5D	72 0.57 0.26	72 0.58 0.26		72 0.58 0.32		72 0.59 0.26	72 0.59 0.31	72 0.57 0.26
			telephone-based nonpharmacologic multidisciplinary treatments	TRS	72 0.69 0.16	72 0.75 0.11		72 0.75 0.18		72 0.73 0.13	72 0.74 0.16	72 0.73 0.13
			telephone-based nonpharmacologic multidisciplinary treatments	SF-6D	72 0.68 0.07	72 0.68 0.08				72 0.68 0.09		72 0.67 0.10
8	Effect on health-related quality of life of a multimodal physiotherapy program in patients with chronic musculoskeletal disorders	Cuesta-Vargas, A. I et al., 2013	physical therapy	EQ-5D	42 0.55 0.32		42 0.75 0.28					
9	Economic evaluation of aquatic exercise	Patrick, D. L et al., 2001	aquatic class	QWB	125 0.60 0.07				101 0.61 0.07			

	for persons with osteoarthritis		usual activity pattern	QWB	124	0.60	0.07										121	0.60	0.08	
10	Economic evaluation of arthritis self-management in primary care	Patel, A et al., 2009	Self-management programme plus an education booklet	EQ-5D	381	0.57	0.25										299	0.55	0.28	285 0.58 0.25
			education booklet alone	EQ-5D	375	0.54	0.28										331	0.56	0.27	312 0.56 0.27
11	Effect of an education programme for patients with osteoarthritis in primary care - A randomized controlled trial	Hansson, E. E et al., 2010	education program	EQ-5D	61	0.58	0.25													
			usual care	EQ-5D	53	0.56	0.30													
12	Factors associated with patients' willingness to consider joint surgery after completion of a digital osteoarthritis treatment program A prospective cohort study	Cronstrom, A et al., 2018	non-surgical OA treatment program comprising education, exercise and asynchronous chat	EQ-5D	458	0.64	0.20													

13	The effect of education and supervised exercise on physical activity, pain, quality of life and self-efficacy - an intervention study with a reference group	Jonsson, T et al., 2018	education and supervised exercise	EQ-5D	195	0.72	0.13				195	0.74	0.08				195	0.74	0.08
			standard care	EQ-5D	69	0.51	0.43				69	0.44	0.49						

Note: *The highlighted data were not included in the meta-analysis as Post-treatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups). OA=osteoarthritis; HSUVs=health state utility values; EQ-5D=EuroQol-5 dimension; SF-6D=Short-Form Six-Dimension; TRS=the rating scale; QWB=Quality of well-being; SD=standard deviation. N means the number of participants.*

Supplement 4.3. The list of included interventions under each category of treatment for knee, hip and mixed OA from studies included in meta-analyses

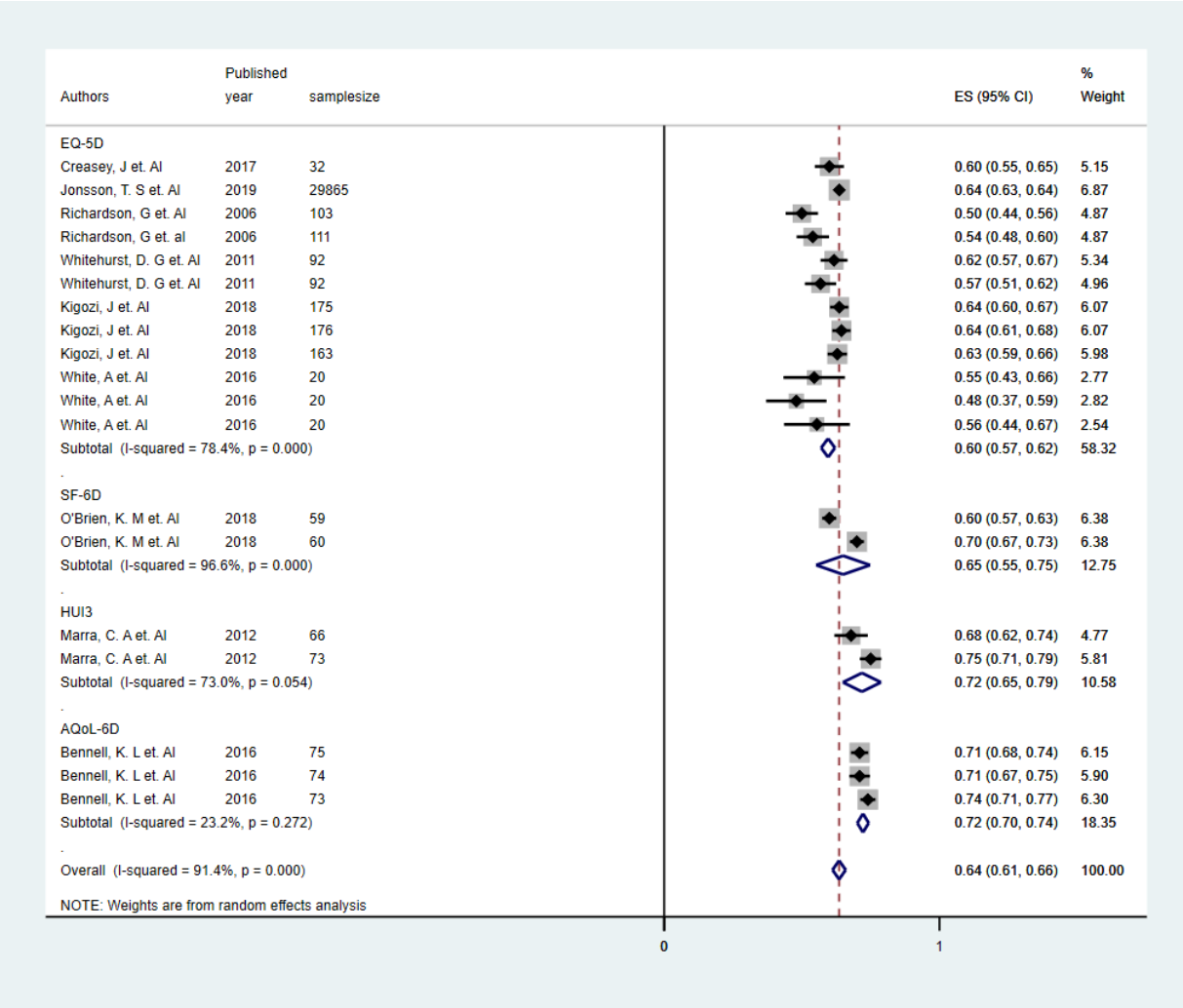
Knee OA	Frequency
Lifestyle treatment (n=9)	
Strengthening and exercise program	1
Better management of patients with OA program offering hip and knee OA information and an individually adapted exercise program	1
Home-based exercise program vs. home-based exercise supplemented with an 8-week class-based exercise program	1
Advice and exercise (AE), advice and exercise plus true acupuncture (AE+TA), advice and exercise plus nonpenetrating acupuncture (AE+NPA)	1
Individually tailored exercise (ITE), targeted exercise adherence (TEA) vs. usual physical therapy care (UC)	1
Telephone-based weight management vs. usual care	1
Multidisciplinary intervention vs. usual care	1
Pain Coping Skills Training (PCST) and exercise vs. exercise or PCST alone	1
Standardised advice and exercise booklet	1
Medication (n=6)	
Traditional NSAIDs vs. celecoxib	1
Tapentadol prolonged release vs. tapentadol immediate release	1
Prolonged release tapentadol vs. controlled release Oxycodone	1
PG201 vs. SKI306X (herb medicines)	1

Spironolactone vs. placebo	1
NSAIDs	1
Injection (n=5)	
Viscoelastic hydrogel	1
Appropriate care (AC)+Hylan G-F 20 vs. AC	1
Intraarticular injections with Hylan G-F 20 (Sanofi) added to usual care (intervention), or usual care only (control)	1
Intra- and extra-articular prolotherapy injections	1
Platelet-rich plasma	1
Primary surgery (n=31)	
Total knee arthroplasty (TKA)	26
Unicompartmental Knee Arthroplasty (UKA)	2
TKA vs. UKA	3
Knee joint distraction (KJD) vs. high tibial osteotomy (HTO)	1
Hip OA	
Lifestyle treatment (n=2)	
BOA program offering hip and knee OA information and an individually adapted exercise program	1
Exercise therapy added to general practitioner care vs. general practitioner care	1
Primary surgery (n=32)	
Total hip arthroplasty (THA)	31
THA vs. hip resurfacing	1
Mixed OA	
Lifestyle treatment (n=13)	
Web-based OA treatment Joint Academy program providing information, exercises, an online physiotherapist, and education regarding factors of relevance to OA, including lifestyle	1
Pre-operative exercise and educational programme including self-management strategies	1
A theory of planned behaviour-based intervention vs. standard care	1
An innovative care model coordinated by a physical therapist and nurse practitioner	1
Supervised patient education and exercise therapy	1
Model OA consultation vs. usual care	1

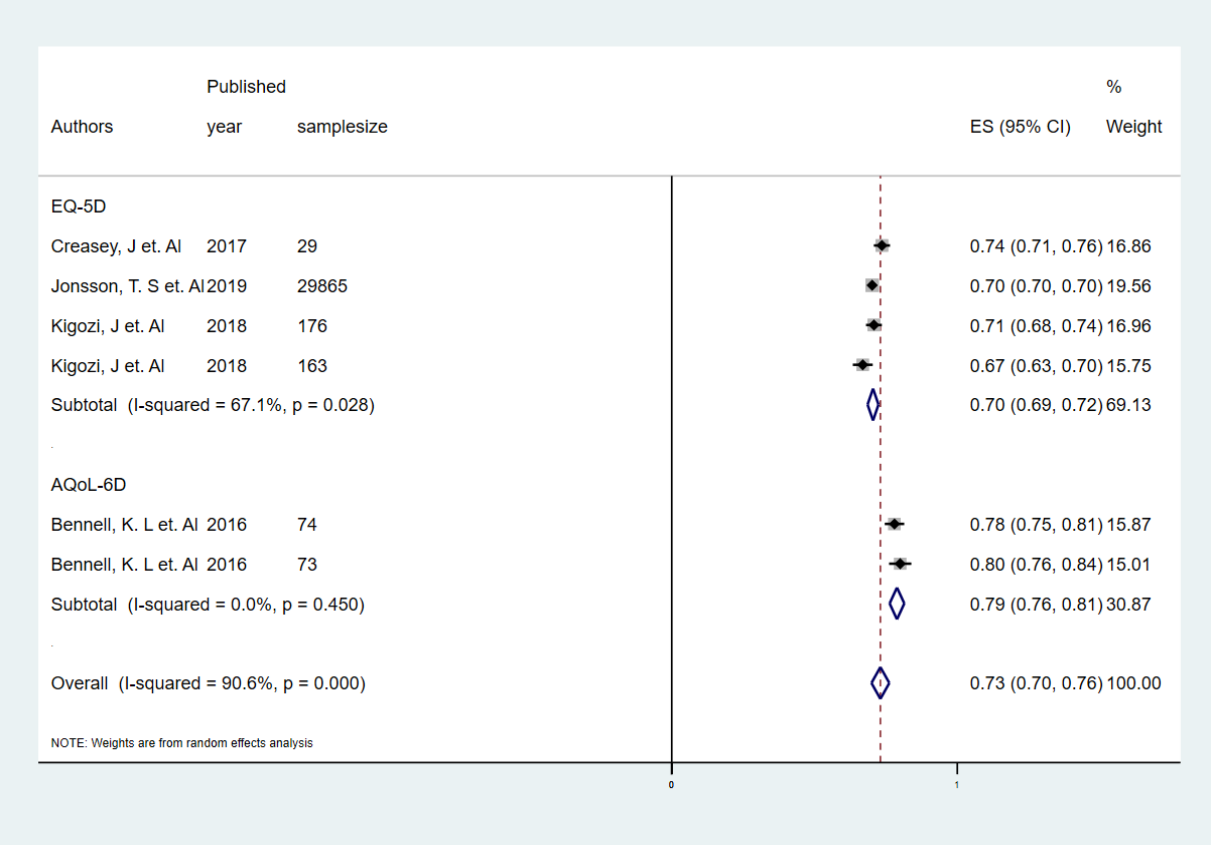
Face-to-face treatment program vs. telephone-based treatment program	1
Multimodal physical therapy program	1
Aquatic exercise vs. usual activities pattern	1
Six sessions of an arthritis self-management programme plus an education booklet (intervention group) or the education booklet alone (standard care control group)	1
Education program vs. usual care	1
Digital non-surgical OA treatment program comprising education, exercise and asynchronous chat	1
Education and supervised exercise vs. standard care	1

Note: n means the number of studies in each category.

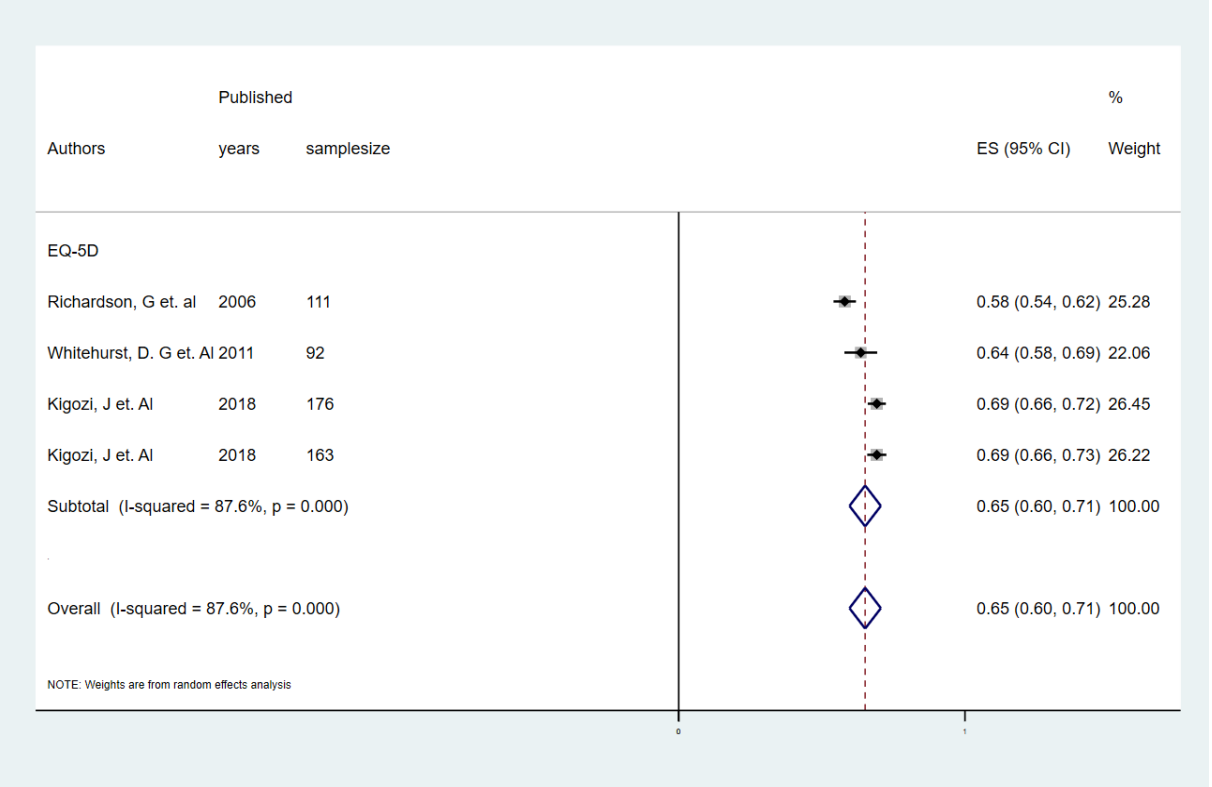
Supplement Figure 4.1: The forest plots of the meta-analyses for knee OA core intervention.



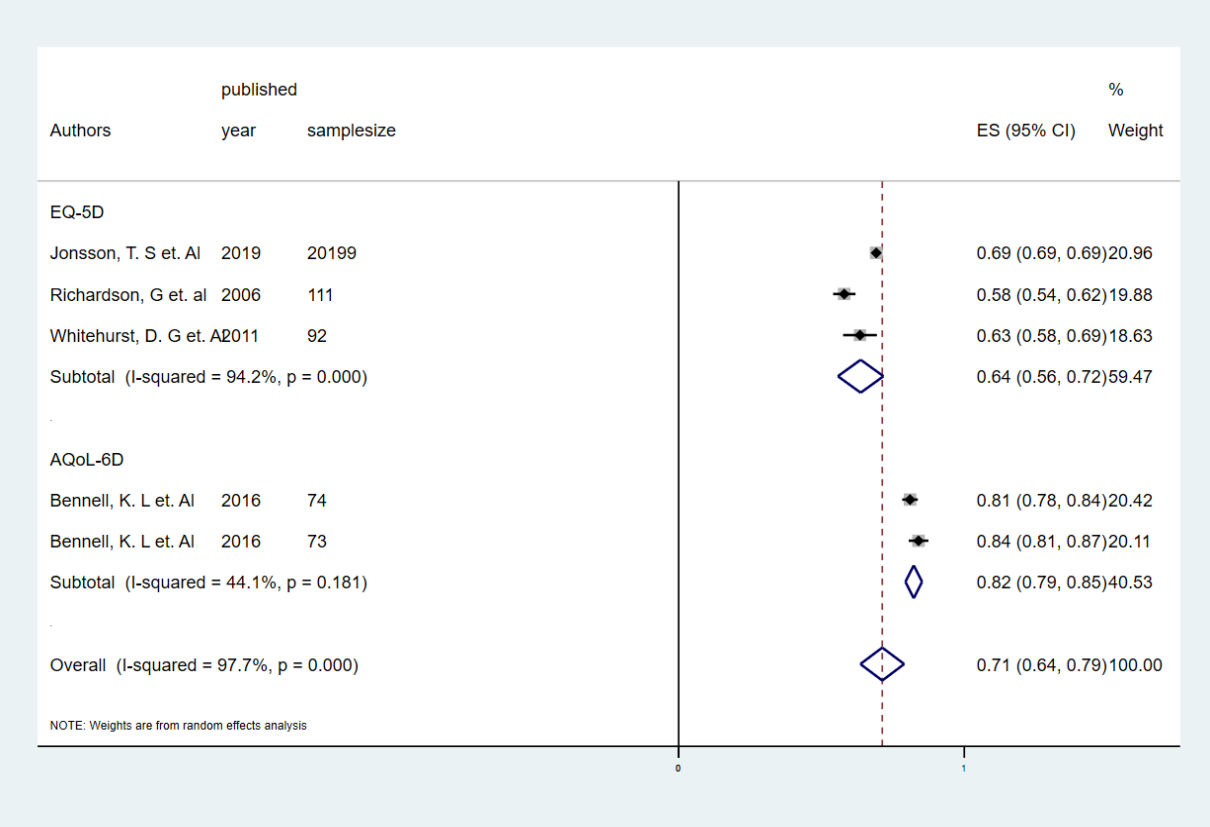
Supplement figure 4.1A. The forest plot for knee OA core intervention at Baseline



Supplement figure 4.1B. The forest plot for knee OA core intervention at 3-month post-treatment.

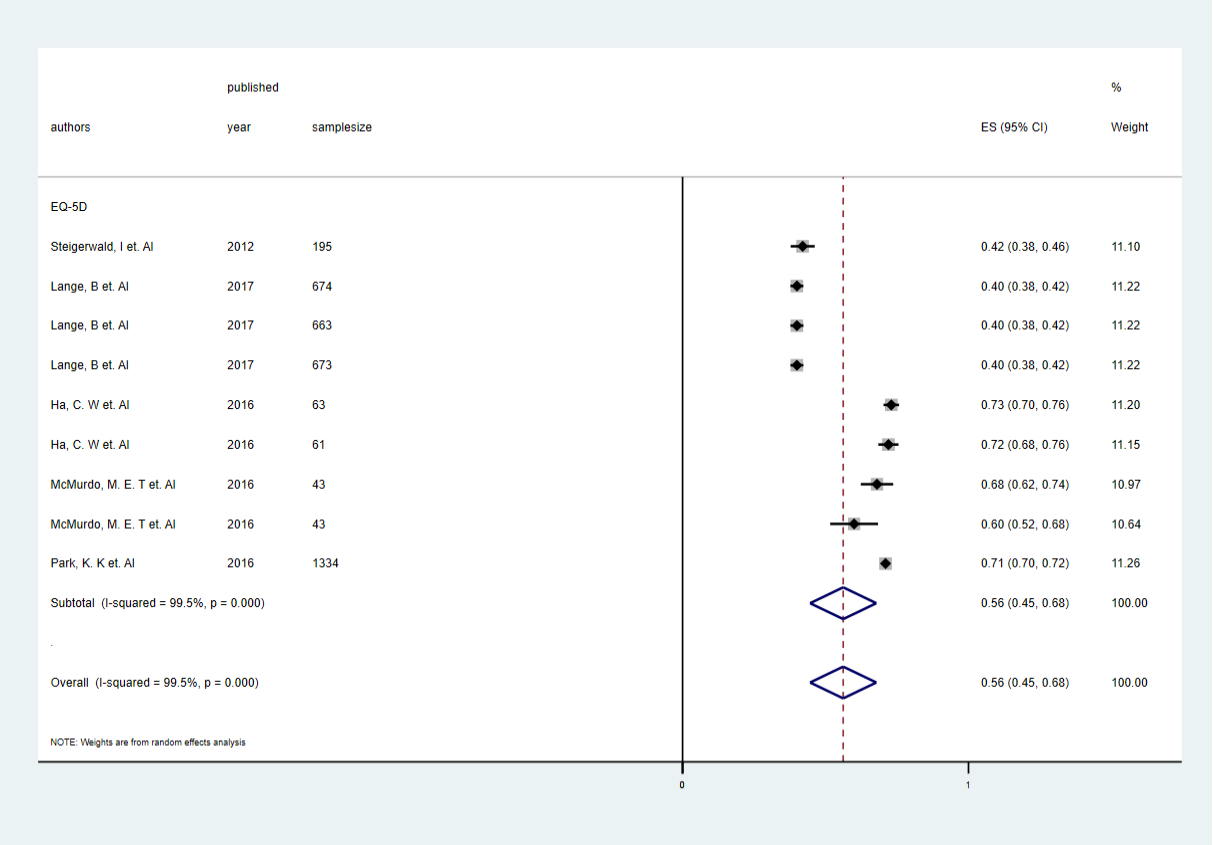


Supplement figure 4.1C. The forest plot for knee OA core intervention at 6-month post-treatment.

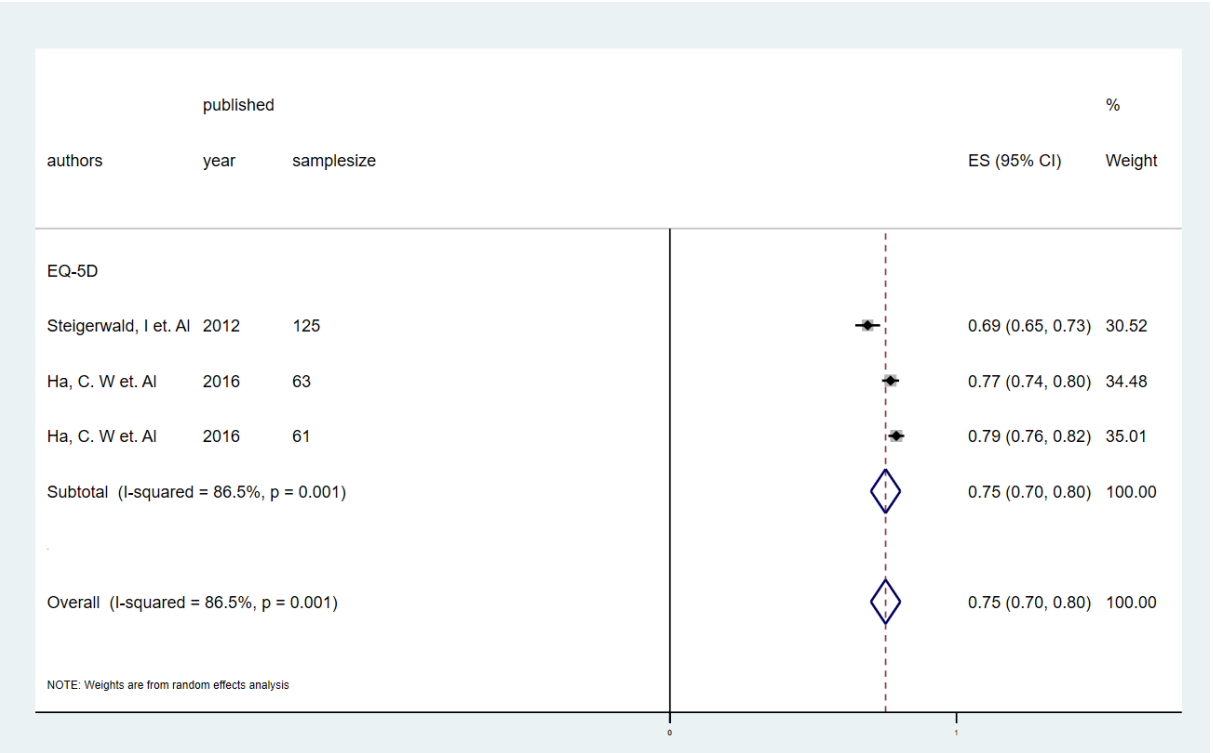


Supplement figure 4.1D. The forest plot for knee OA core intervention at 1-year post-treatment.

Supplement Figure 4.2: The forest plots of the meta-analyses for knee OA medication treatment.

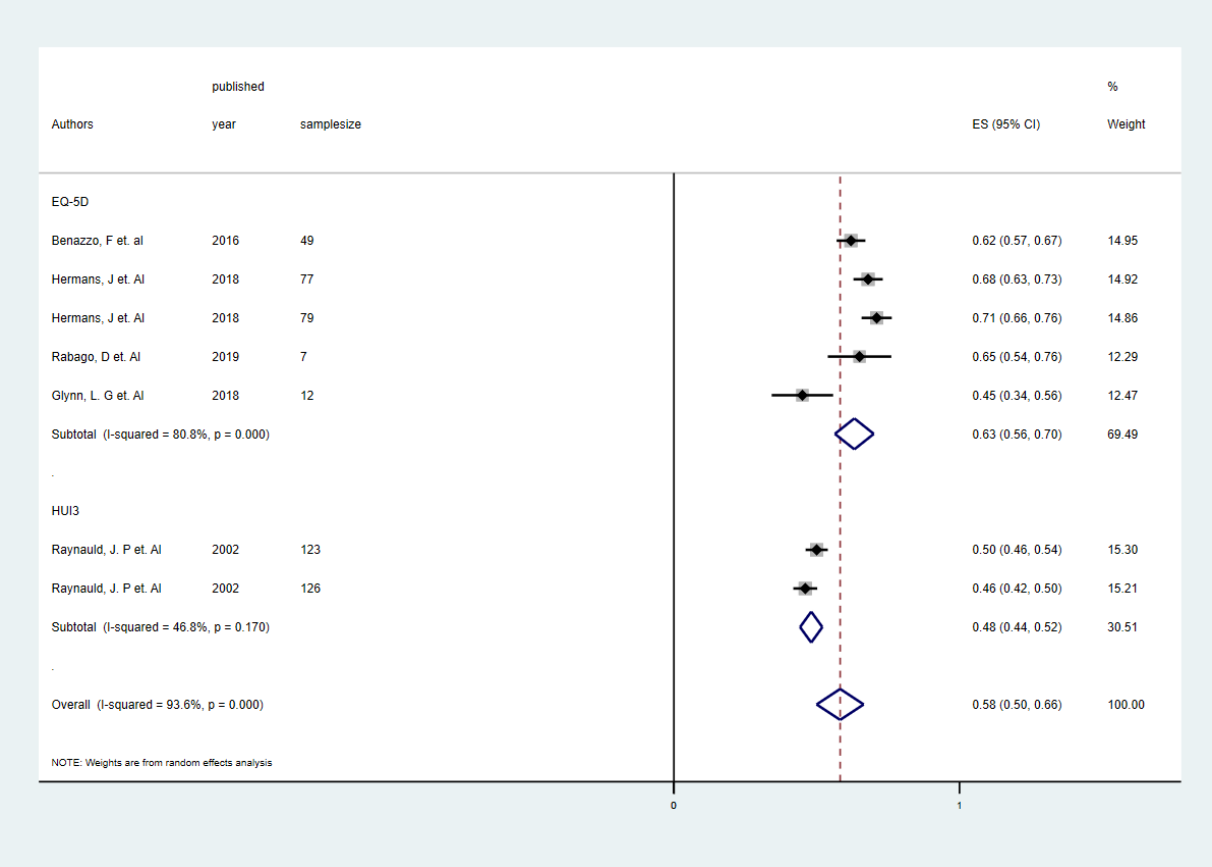


Supplement figure 4.2A. The forest plot for knee OA medication treatment at Baseline.



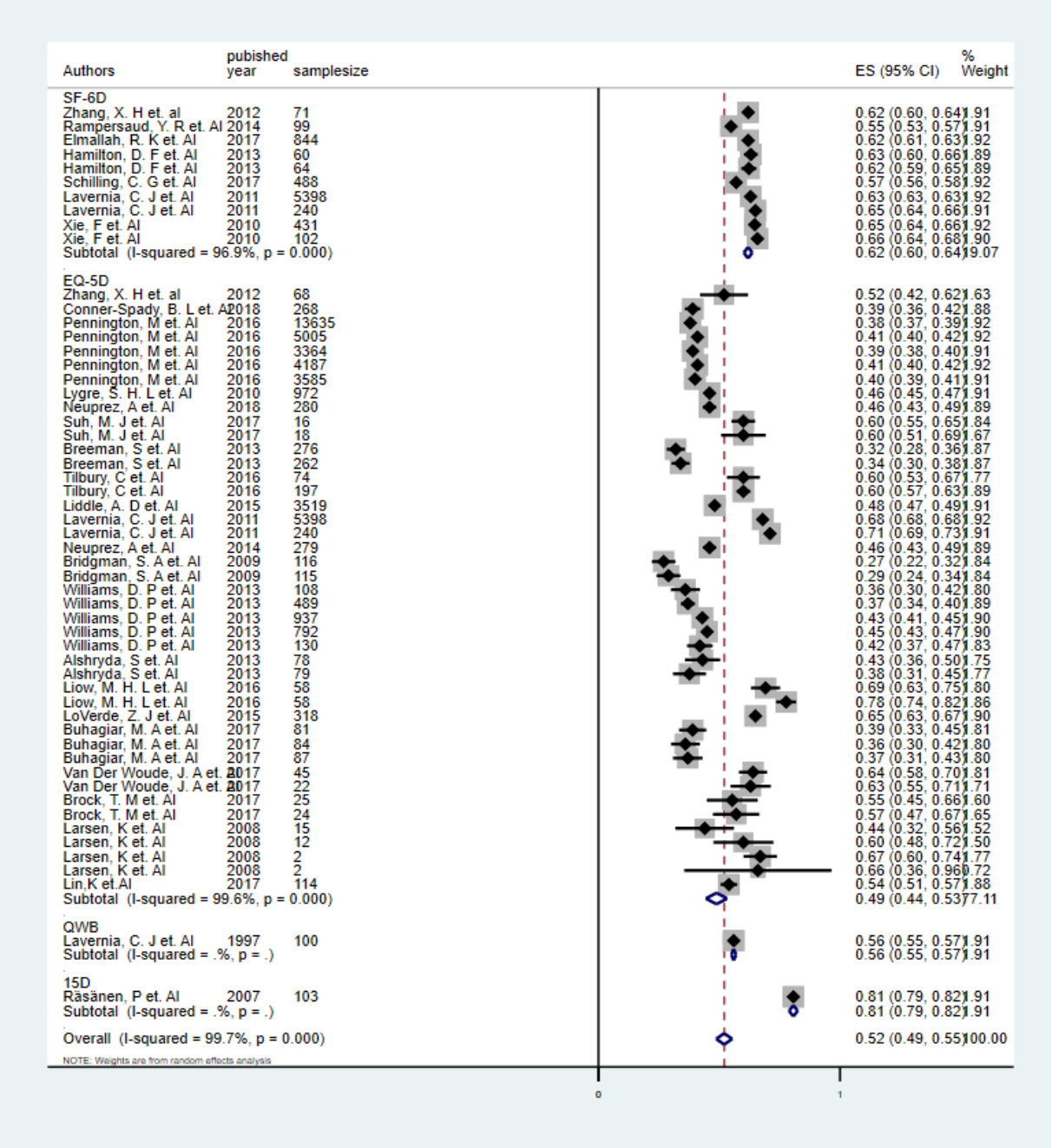
Supplement figure 4.2B. The forest plot for knee OA medication treatment at 3-month post-treatment.

Supplement Figure 4.3: The forest plots of the meta-analyses for knee OA injection treatment.

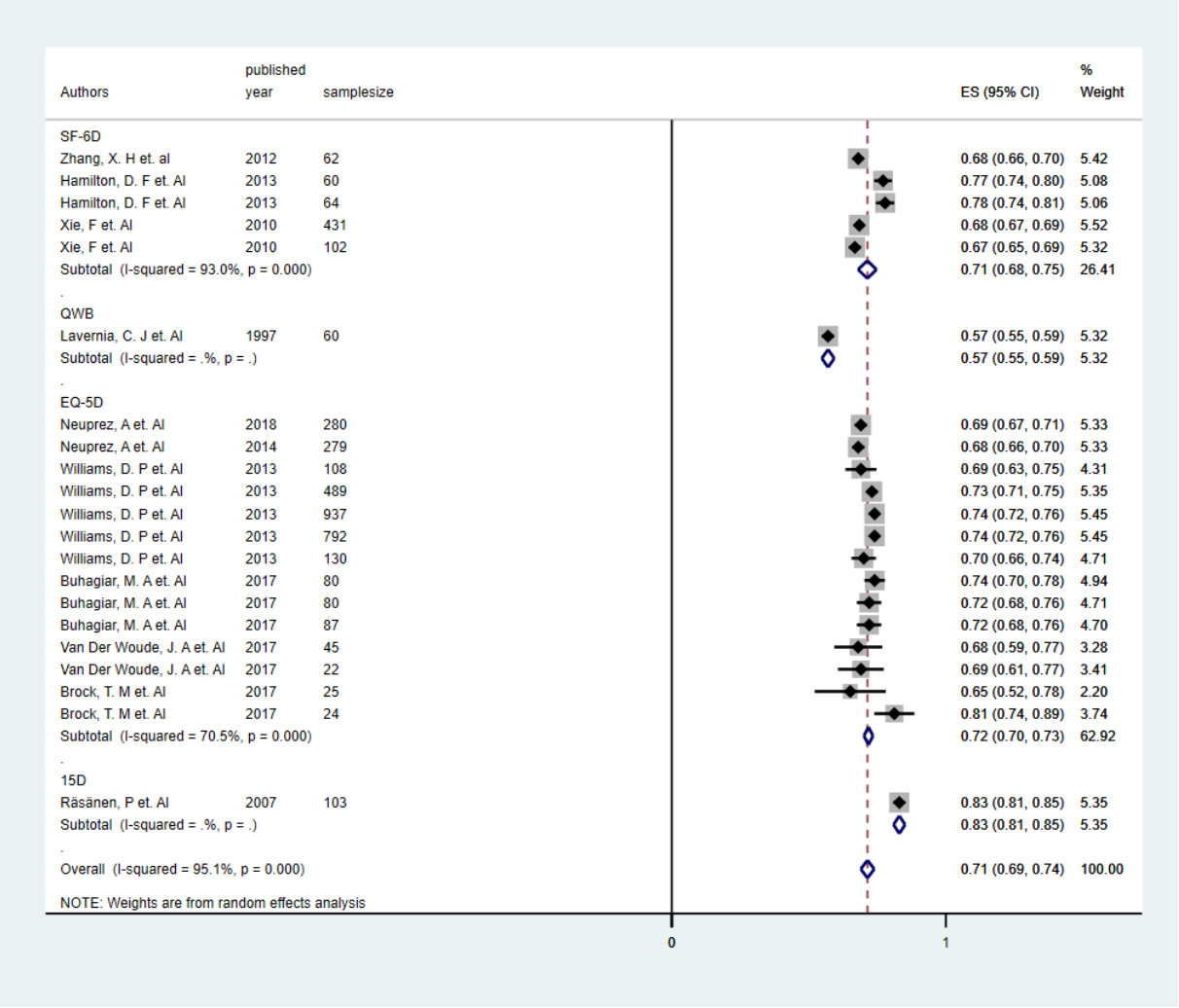


Supplement figure 3. The forest plot for knee OA injection treatment at baseline.

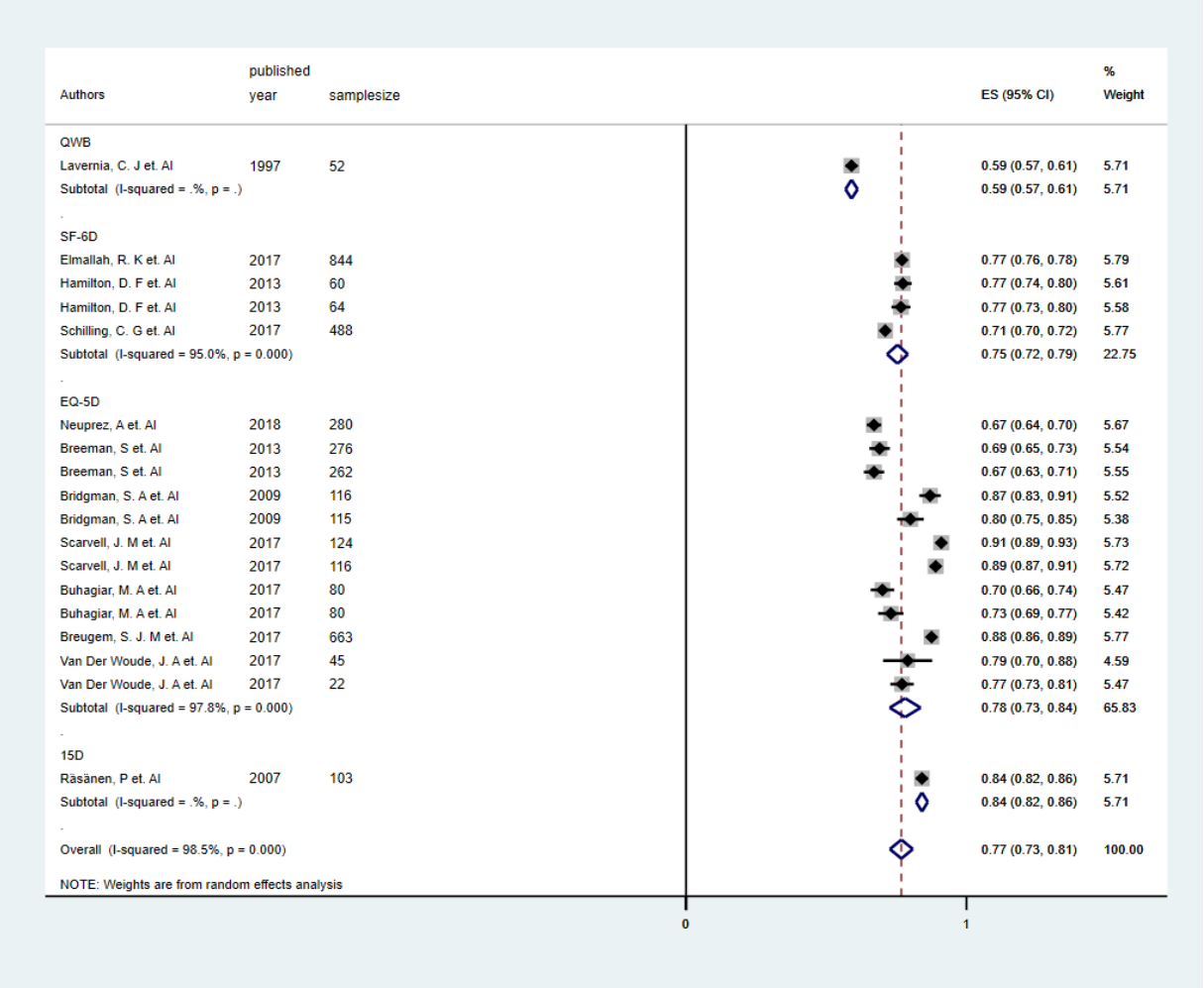
Supplement Figure 4.4: The forest plots of the meta-analyses for knee OA primary surgery treatment.



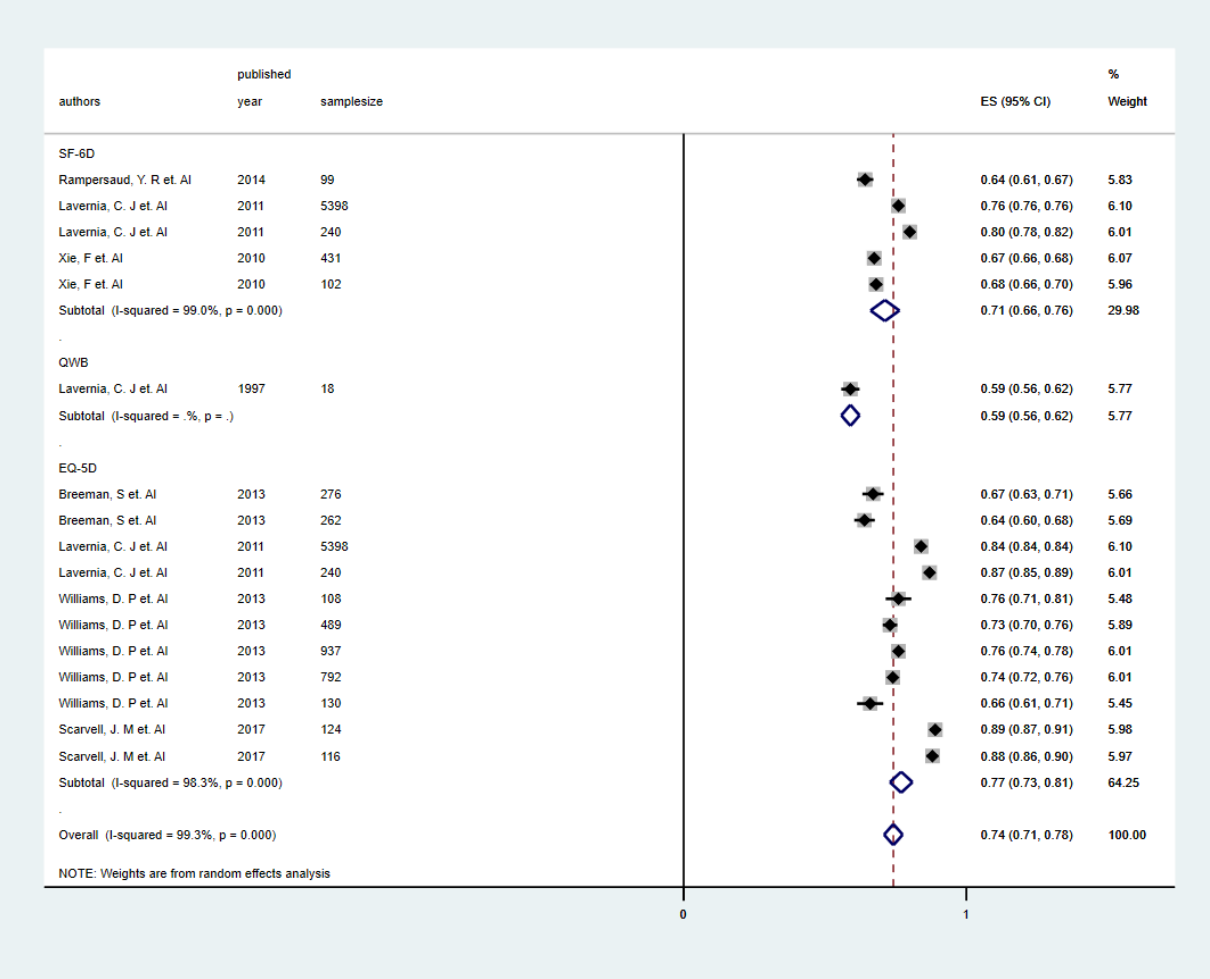
Supplement figure 4.4A. The forest plot for knee OA primary surgery treatment at baseline.



Supplement figure 4.4B. The forest plot for knee OA primary surgery treatment at 6-month post-treatment.

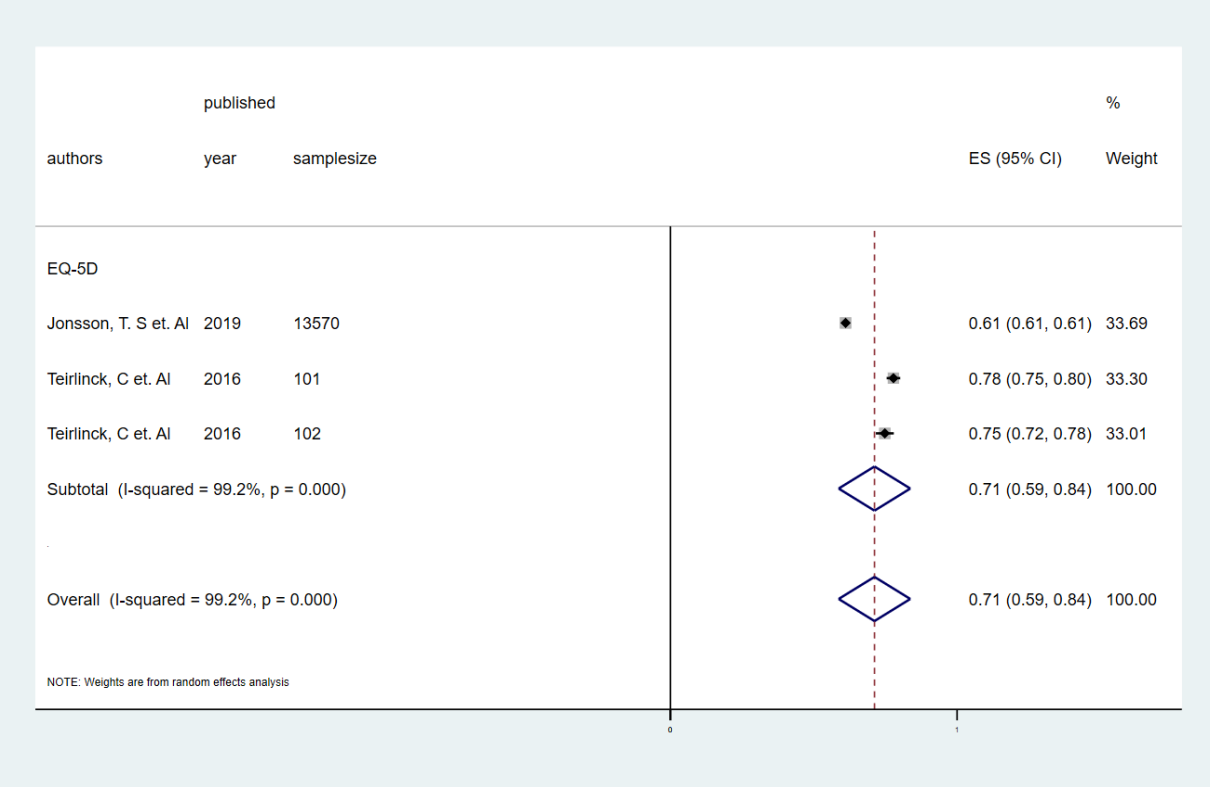


Supplement figure 4.4C. The forest plot for knee OA primary surgery treatment at 1-year post-treatment.

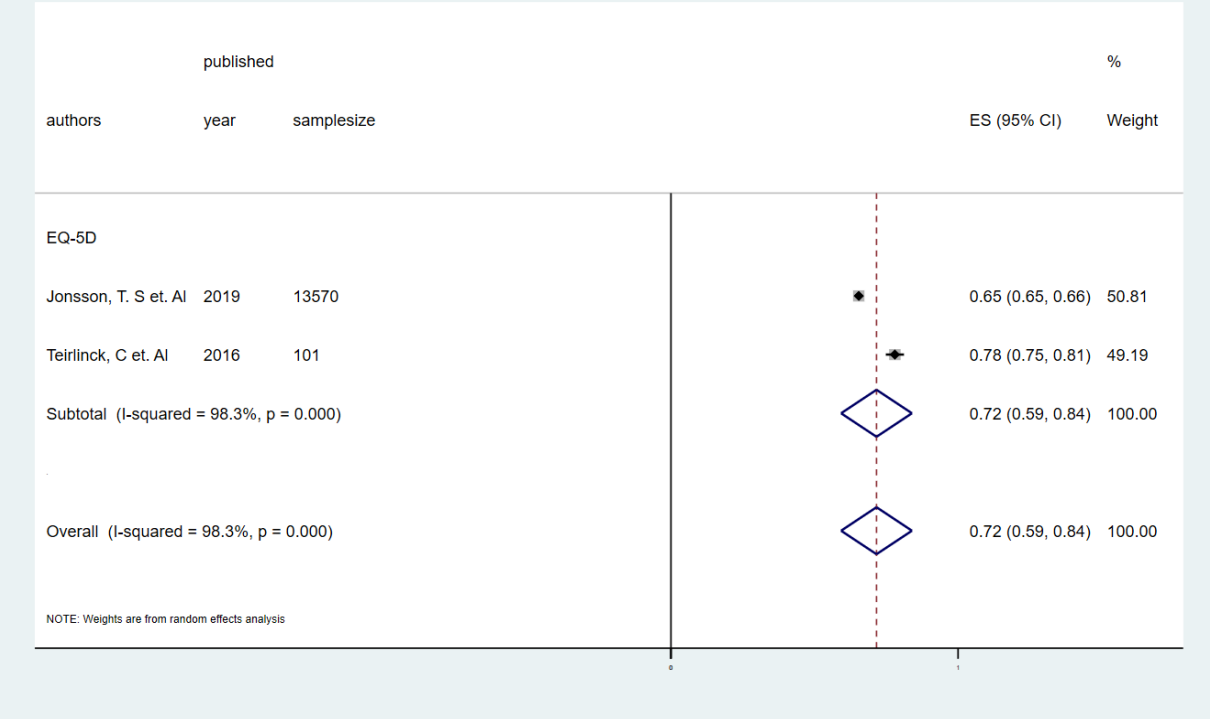


Supplement figure 4.4D. The forest plot for knee OA primary surgery treatment at 2-year post-treatment.

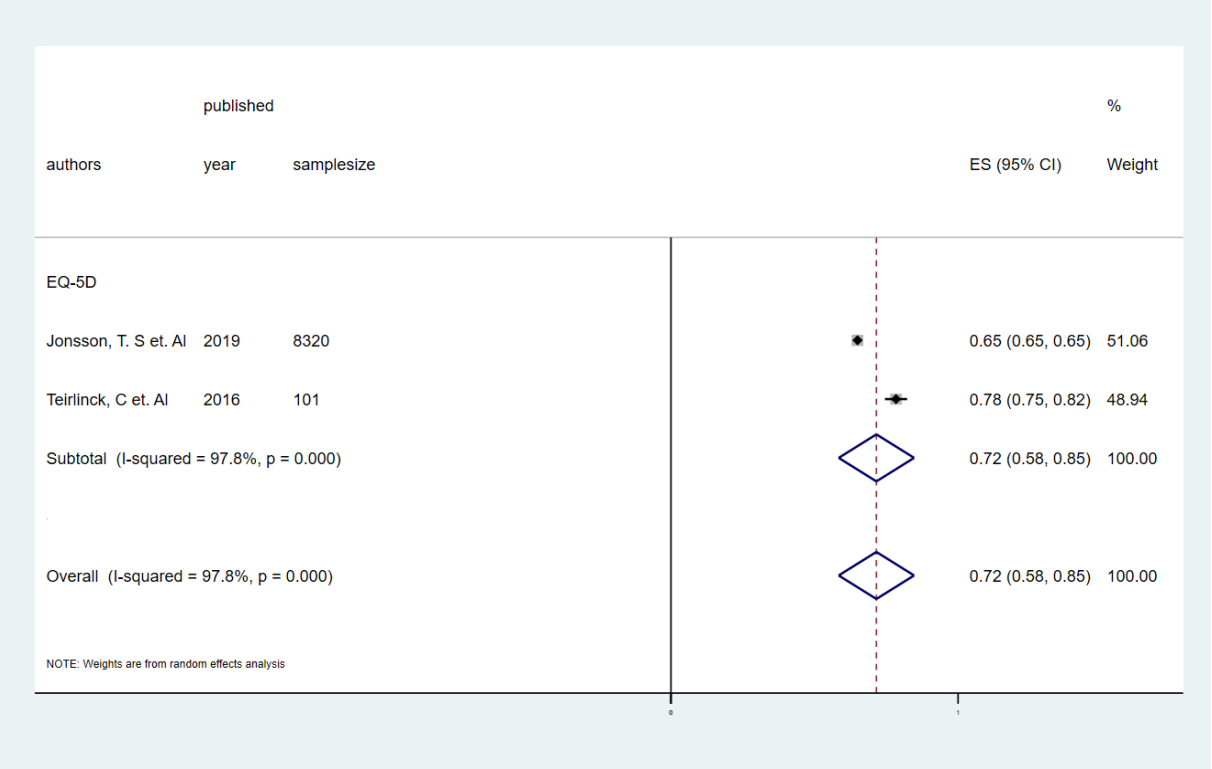
Supplement Figure 4.5: The forest plots of the meta-analyses for hip OA core intervention.



Supplement figure 4.5A. The forest plot for hip OA lifestyle treatment at baseline.

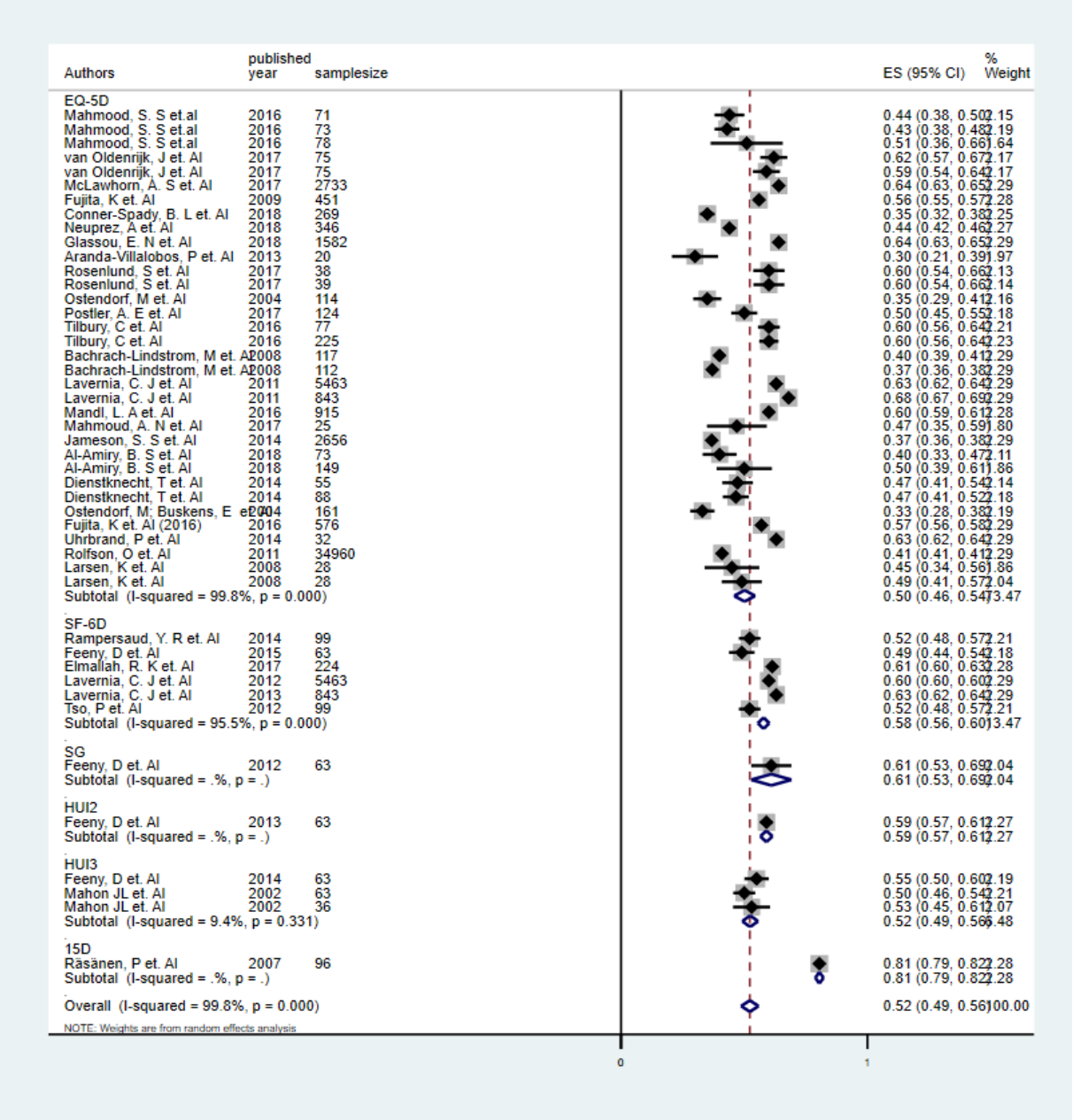


Supplement figure 4.5B. The forest plot for hip OA lifestyle treatment at 3-month post-treatment.

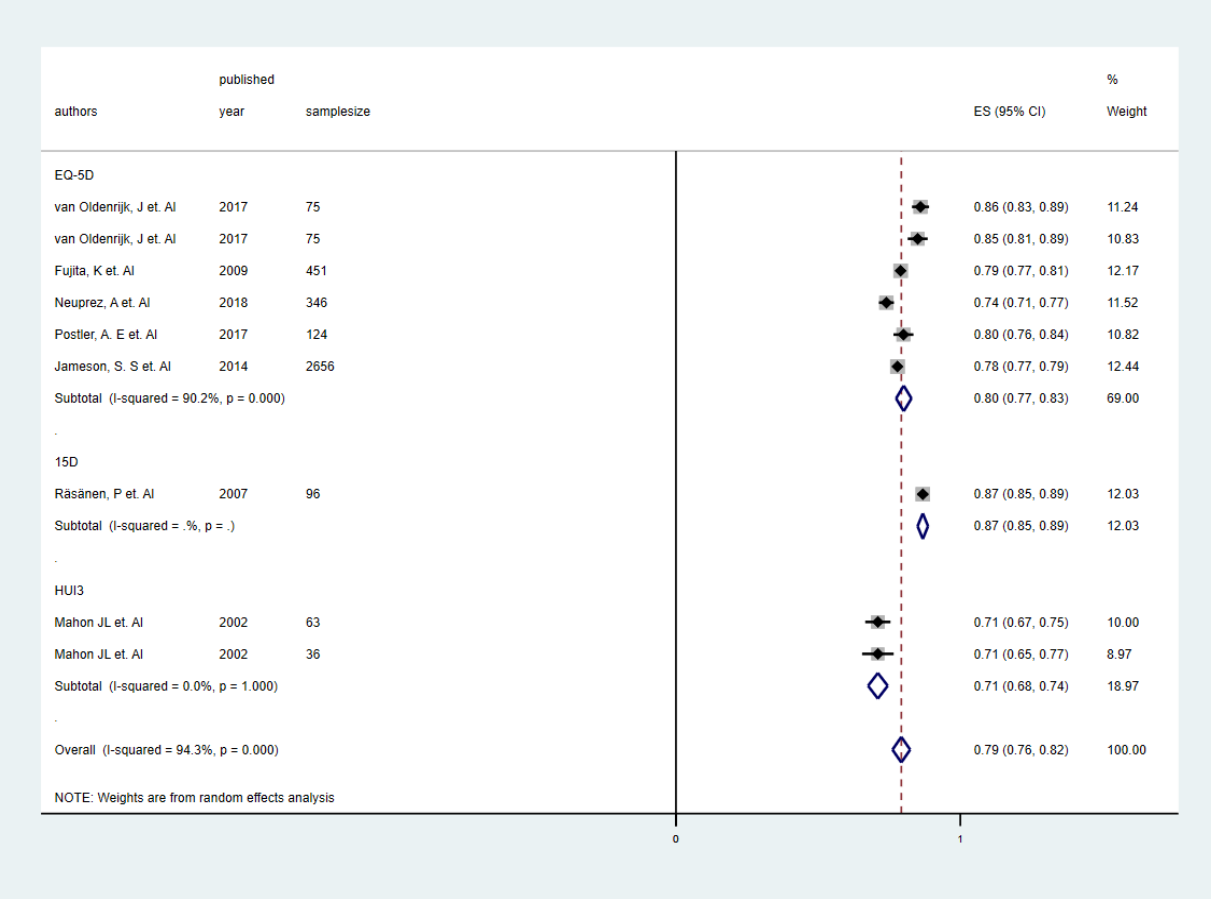


Supplement figure 4.5C. The forest plot for hip OA lifestyle treatment at 1-year post-treatment.

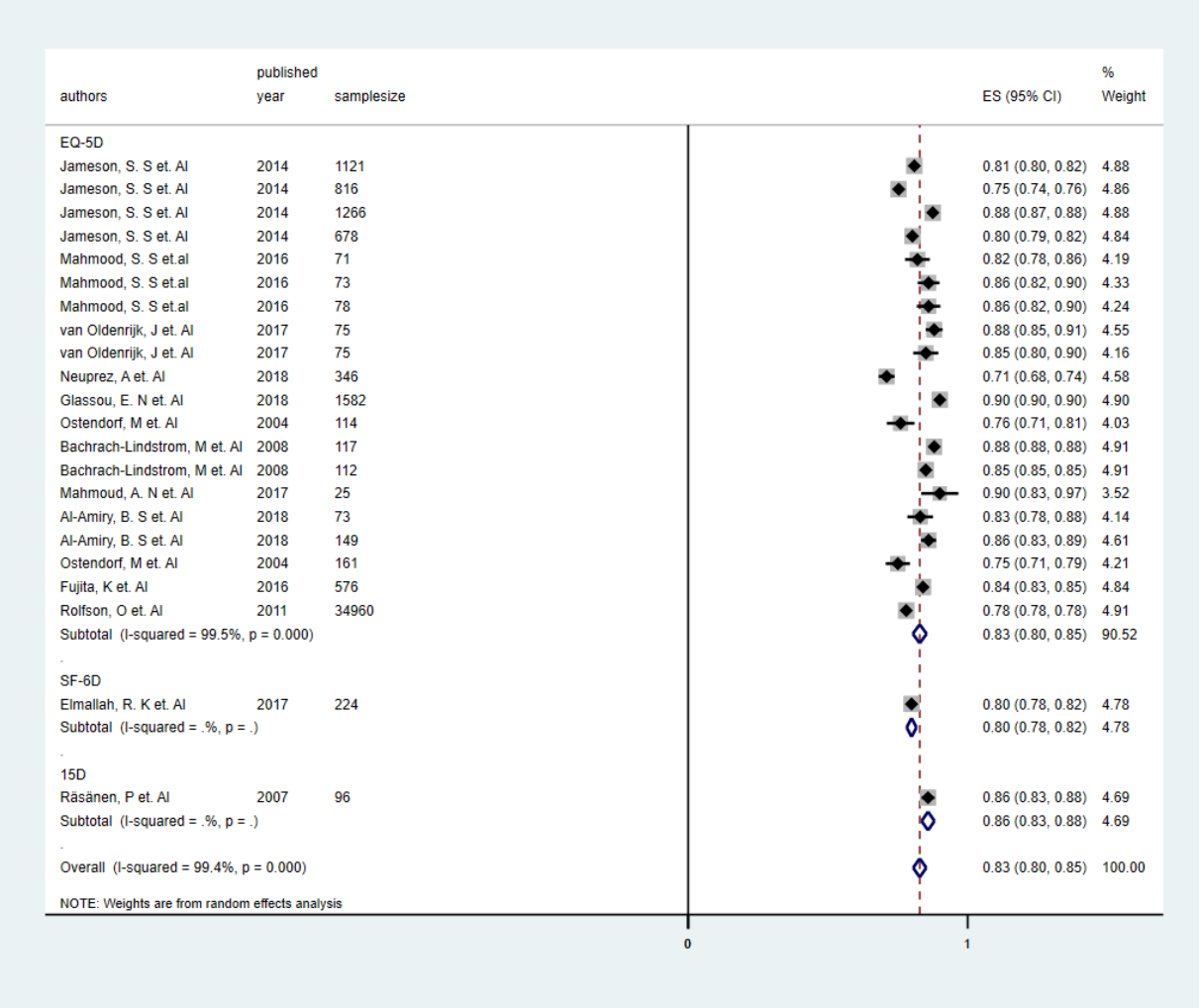
Supplement Figure 4.6: The forest plots of the meta-analyses for hip OA primary surgery treatment.



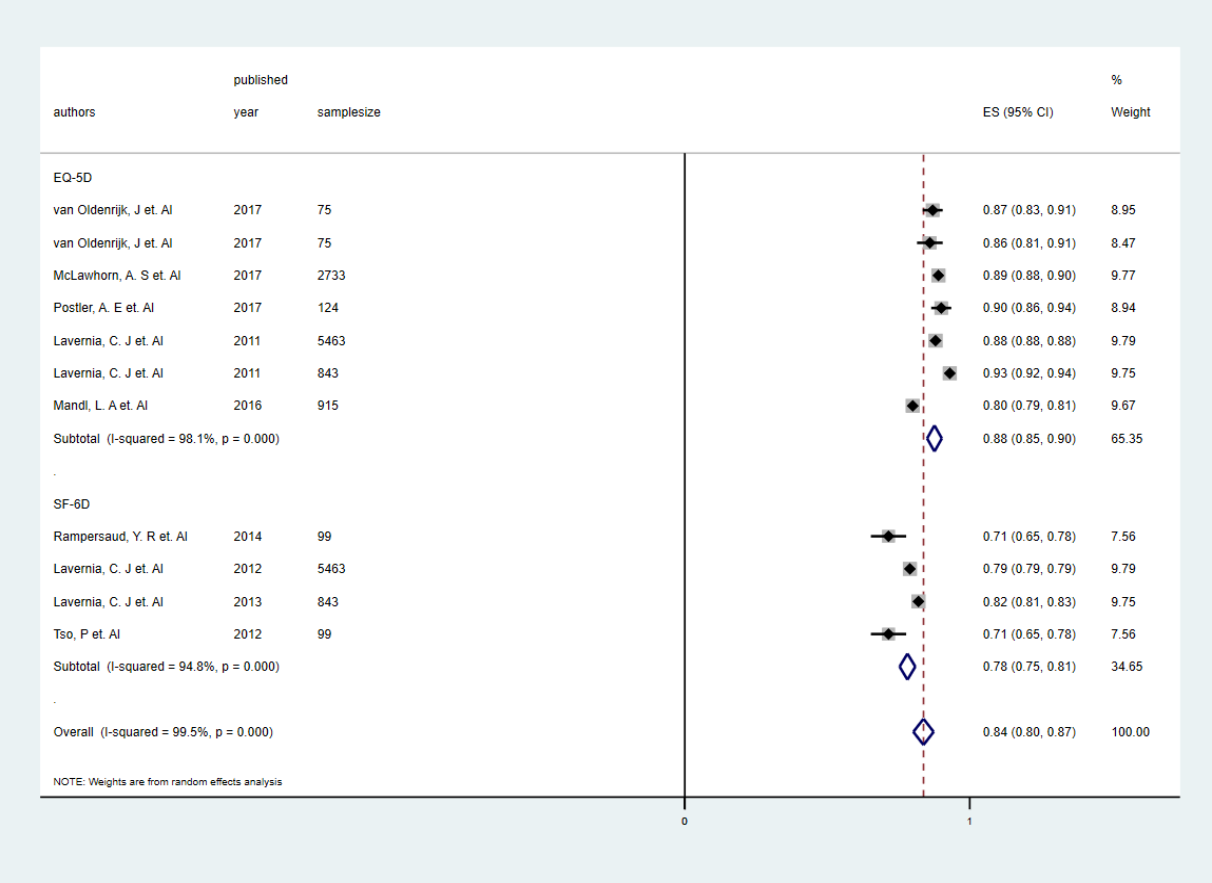
Supplement figure 4.6A. The forest plot for hip OA primary surgery treatment at baseline.



Supplement figure 4.6B. The forest plot for hip OA primary surgery treatment at 6-month post-treatment.

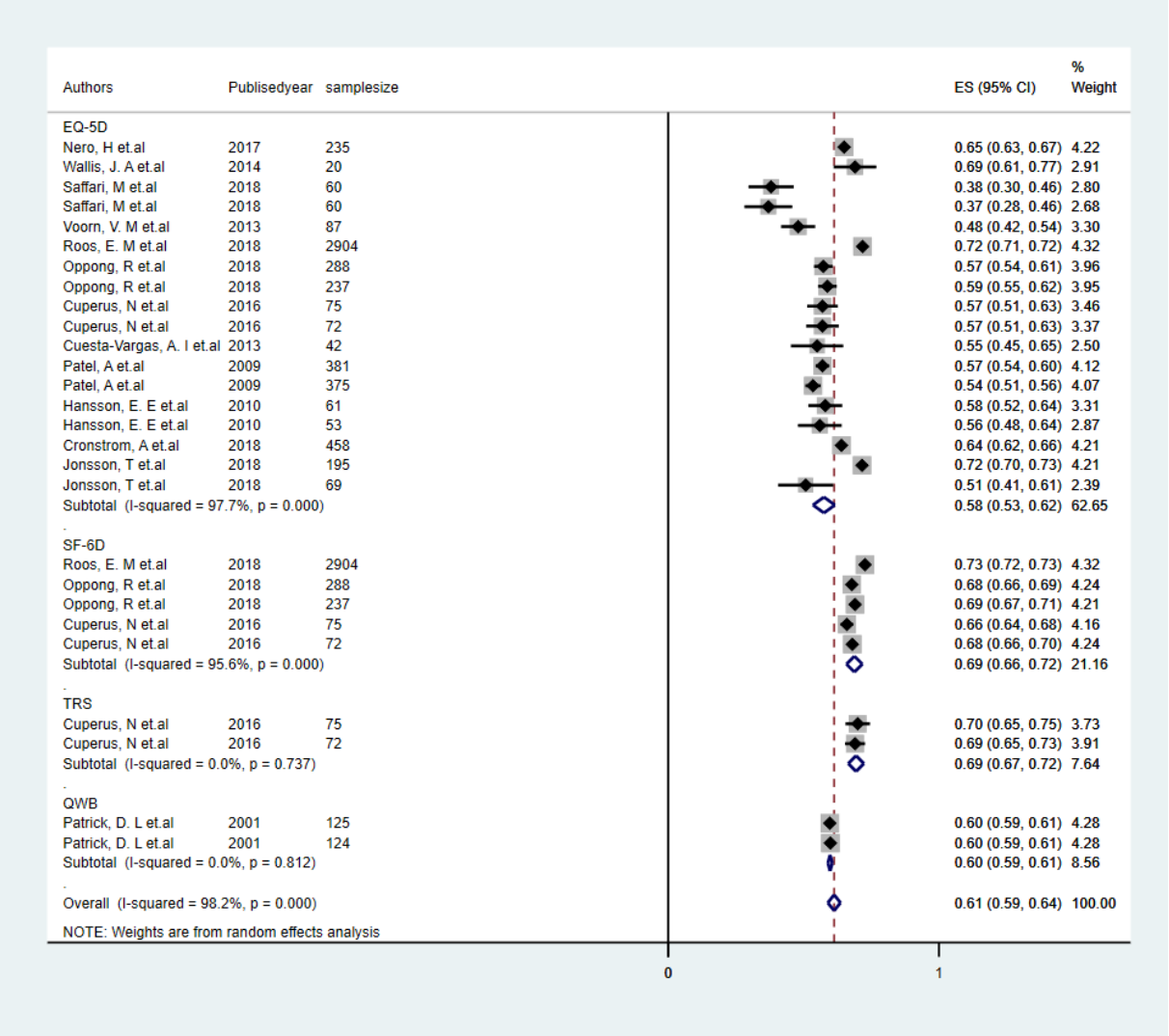


Supplement figure 4.6C. The forest plot for hip OA primary surgery treatment at 1-year post-treatment.

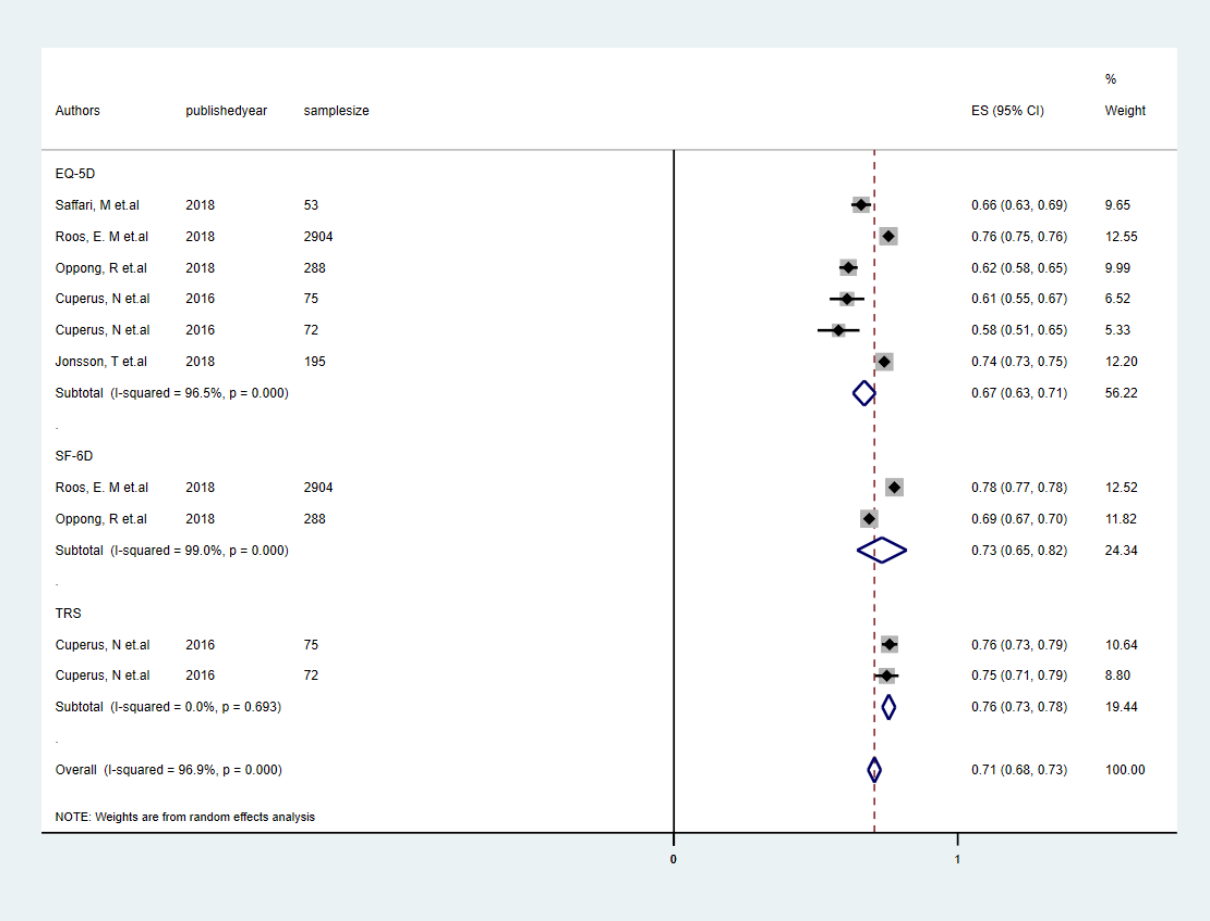


Supplement figure 4.6D. The forest plot for hip OA primary surgery treatment at 2-year post-treatment.

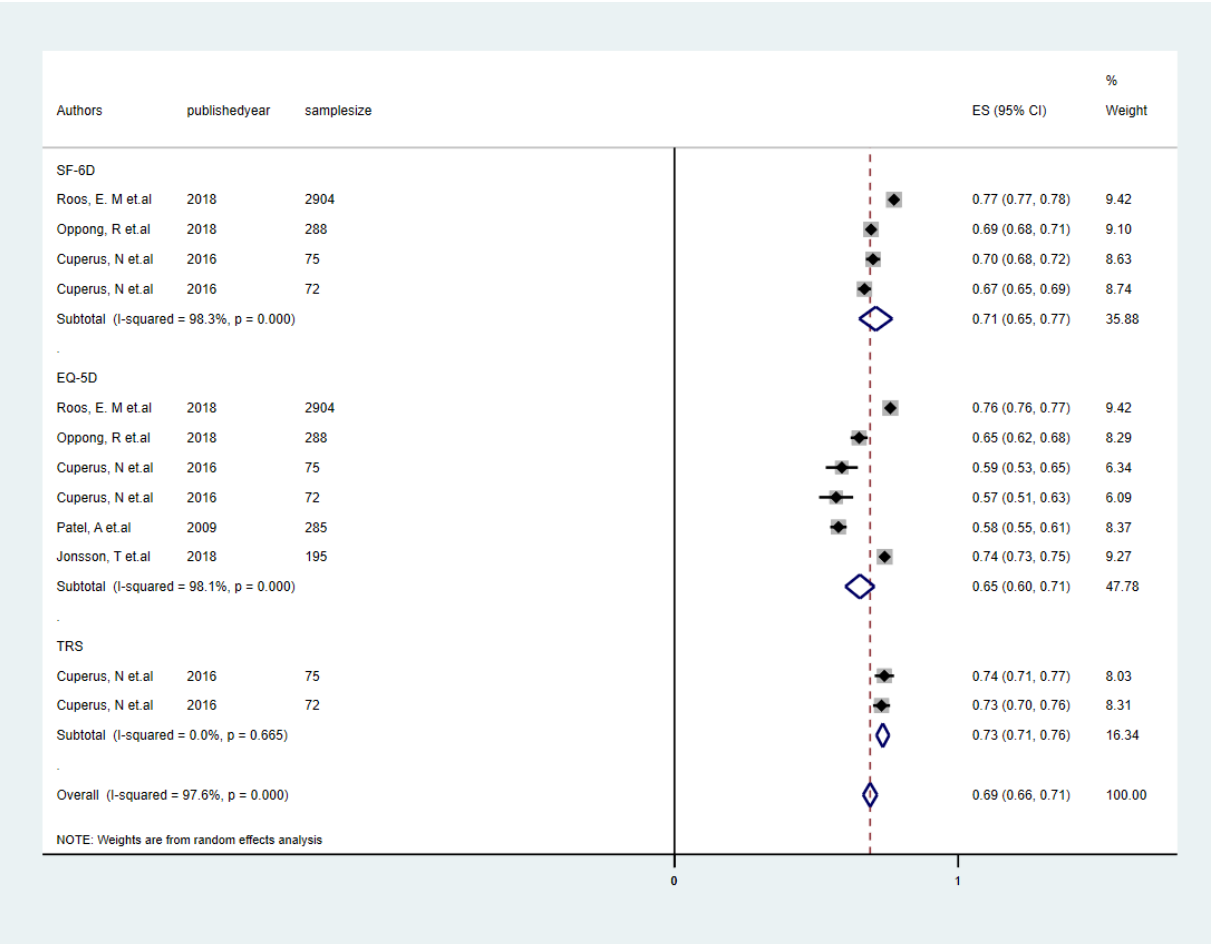
Supplement Figure 4.7: The forest plots of the meta-analyses for mixed OA core intervention.



Supplement figure 4.7A. The forest plot for mixed OA lifestyle treatment at baseline.



Supplement figure 4.7B. The forest plot for mixed OA lifestyle treatment at 3-month post-treatment.



Supplement figure 4.7C. The forest plot for mixed OA lifestyle treatment at 1-year post-treatment.

Chapter 5: Cross-sectional and temporal differences in health-related quality of life of people with and without OA: a ten-year prospective study

5.1 Preface

Chapter 5 presents Study 3 which provided a comprehensive assessment of the effects of OA on HRQoL, measured in terms of HSUVs and dimensional scores, using Australian data and Assessment of Quality of Life (AQoL)-4D instrument. Chapter 5's key aim is to addresses the paucity of data on HSUVs impacts of OA in Australia by investigating the cross-sectional and longitudinal difference in HSUVs and health-dimension scores of Australians with OA compared with those without. Chapter 5 also identifies the physical and psychosocial health drivers of longitudinal changes in HSUV scores of Australians with OA compared with those without. A further important aim of Chapter 5 is to generate HSUVs input for use in the future health economic models of OA for Australian and similar populations.

The text in Chapter 5 has been published in *Rheumatology* (Zhao, Ting, Hasnat Ahmad, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Graeme Jones, and Andrew J. Palmer. "Cross-sectional and temporal differences in health-related quality of life of people with and without osteoarthritis: a 10-year prospective study." Rheumatology (2021) (Supplement 5A).

5.2 Abstract

Objective: To describe the impact of osteoarthritis (OA) for health-related quality of life (HRQoL) in the forms of health state utility values (HSUVs) and health-dimension scores, and to investigate the longitudinal changes in HRQoL of people with compared to without OA using an Australian population-based longitudinal cohort.

Methods: Participants of the Tasmanian Older Adult Cohort (interviewed at baseline [n=1,093], 2.5 years [n=871], 5 years [n=760] and 10 years [n=562]), with data on OA diagnosis and HRQoL were included. The mean (standard deviation) age of the TASOAC participants at baseline was 62.5 (7.51) years and 51% of the sample were females. HRQoL was assessed using the Assessment of Quality of Life-4-Dimensions and analysed using multivariable linear mixed regressions.

Results: Compared to participants without OA, HSUVs for those with OA were 0.07 (95% confidence interval: -0.09, -0.05) units lower on average over ten years. HSUVs for participants with knee and/or hip OA were similar to those with other types of OA at 2.5 years follow-up and then diverged, with HSUVs of the former being up to 0.09 units lower than the latter. Those with OA had lower scores for psychological wellness, independent living and social relationships compared to those without OA. Independent living and social relationships were mainly impacted by knee and/or hip OA with the effect on the former increasing over time.

Conclusion: The mean HSUV of OA participants was 0.07 units lower than that of those without OA over ten years, driven largely by lower scores on psychological wellness, independent living and social relationships.

5.3 Introduction

Osteoarthritis (OA) is one of the most common chronic joint diseases, affecting knees, hips and the small joints of the hands. OA is characterised by joint pain, stiffness, swelling, loss of function and disability; which in turn, negatively impacts individuals' health-related quality of life (HRQoL) (1). HRQoL focuses on the impact of physical and psychosocial health status on quality of life. It can be reflected as an overall index defined by health state utility value (HSUV) or scores for different dimensions of health. HSUVs measure the strength of preference for a given health state and are represented as a number between 0 (death) and 1 (optimal health). Health states worse than death may exist, with negative HSUVs assigned (2).

The indirect measurement of HSUVs using multi-attribute utility instruments (MAUIs) such as the EuroQoL five-dimensions (EQ-5D), Assessment of Quality of Life-4 Dimensions (AQoL-4D) and Short-Form-6D (SF-6D) has become increasingly popular (3). MAUIs are a set of questions and response categories which seek to describe a person's health and an accompanying formula, or a set of weights elicited from a sample of the general population for converting responses into HSUVs (4, 5). In addition to HSUVs, some MAUIs (e.g.: AqoL-4D) can also provide summary scores for specific health dimensions, for example physical dimension and psychosocial dimensions.

Studies assessing the impact of OA on HRQoL have mostly been undertaken in Europe and the United States and are cross-sectional in nature (1, 6, 7). There is a paucity of data on HRQoL impacts of OA in Australia. For instance, a 2015 Australian study estimated the

impacts of knee and/or hip OA on HRQoL using data from a multi-centre, cross-sectional survey (8). While this added to our understanding of knee and/or hip OA impacts on HRQoL of younger people, it used cross-sectional data from a small number (n=21) of community participants who had knee and hip OA. Furthermore, no study has investigated longitudinal changes in the mean HSUVs and individual health-dimension scores of people with OA compared to those without OA. Our study aims to extend the scope of previous research by investigating the cross-sectional and longitudinal differences in HRQoL in the forms of HSUVs and health-dimension scores of people with OA compared to those without OA using an Australian population-based cohort.

5.4 Methods

5.4.1 Study Design

The Tasmanian Older Adult Cohort (TASOAC) is a prospective, population-based study aimed at identifying the environmental, genetic, and biochemical factors associated with the development and progression of OA at multiple site (knee, hip, hand, and spine). Further details about the TASOAC study have been described elsewhere (9). Participants aged 50–80 years in 2002 were selected from the electoral roll in Southern Tasmania, using sex-stratified simple random sampling without replacement. Participants were excluded if they resided in an aged care facility or were unable to have a knee magnetic resonance imaging scan. A total of 1,099 adults (response rate=57%) consented to participate in the study. The research was approved by the Southern Tasmanian Health and Medical Human Research Ethics Committee and written informed consent to participate in the study was obtained from all participants. Participants completed questionnaires and attended necessary clinical tests (not reported in this study). Data were collected using interview-administered questionnaires at 4 periods: baseline (from February 2002 to September 2004), 2.5 years later (range 1.4−4.8 years), 5 years later (range 3.6−6.9 years), and 10 years later (range 9−13 years).

5.4.2 Measurement of Health-related Quality of life (HRQoL)

HRQoL was assessed at baseline, 2.5, 5 and 10 years using the AQoL-4D questionnaire, which consists of 12 items covering four health dimensions, each with three items and four response levels (0,1,2 and 3). The four dimensions and their corresponding items are: (1) independent living (self-care, activities of daily living, and mobility), (2) physical senses (sight, hearing and communication), (3) social relationships (social isolation, relationship and family role)

and (4) psychological wellbeing (sleep, anxiety and pain). The items were combined according to the AQoL-4D algorithm to calculate dimension scores and an overall index of HSUVs (10).

5.4.3 Diagnosis of Osteoarthritis

At baseline, participants were asked “Have you had been told by a doctor that you have osteoarthritis at any of these sites”. Seven sites were listed: neck, back, hands, shoulders, hips, knees, and feet. Participants were given the choice between answering "yes" or "no", and all participants answering “yes” were considered diagnosed with OA.

5.4.4 Categorisation of OA patients

We first categorised our sample into two groups based on their baseline OA diagnosis: 1) without OA (no joint site diagnosed with OA); 2) with OA (one or more of the seven joint sites diagnosed with OA). We further classified OA participants into two groups based on their OA types: 2a) participants with knee and/or hip OA (at least one of the sites of knee and hip diagnosed with OA); and 2b) other types of OA (at least one of the five joint sites other than knee and hip diagnosed with OA).

5.4.5 Other Characteristics

Data on sex and date of birth were collected at baseline. Weight was measured to the nearest 0.1 kg (with no shoes/socks/bulky clothing/headwear) using a single pair of calibrated electronic scales (Seca Delta Model 707). Height was measured to the nearest 0.1 cm (barefooted) using a stadiometer. Weight and height were collected at all timepoints. Body mass index (BMI) was calculated as [weight (kg)/ (height (m)^2], and participants were grouped into 3 categories of BMI (normal weight [BMI<25.0kg/m^2], overweight [25.0kg/m^2≤BMI≤29.9kg/m^2] and obese [BMI≥30.0kg/m^2]) following the WHO criteria (11). Education data was collected at baseline, and participants were grouped into three categories (i.e.: Low education [having no formal qualification], middle education [holders of school or intermediate certificate, higher school or leaving certificate, trade/apprenticeship certificate/diploma] and high education [holders of university degree/higher degree]) based on their highest education levels. Employment status was collected at all timepoints, and participants were categorised into 5 groups [1, employed/self-employed (part-time/full-time); 2, retired; 3, unemployed; 4, disability pensioners; 5, others (including home duties, students, sole parent pension)]. Presence or absence of ten comorbidities were assessed at all timepoints by asking “Have you ever been diagnosed by a doctor as having any of the following”: diabetes,

heart attack, hypertension, thrombosis, asthma, bronchitis/emphysema, osteoporosis, hyperthyroidism, hypothyroidism, and rheumatoid arthritis. The number of comorbidities was then calculated as the total number of diagnosed chronic conditions at each timepoint.

5.4.6 Statistical analysis

We summarised the demographic and other features of our sample using descriptive statistics. The difference in HSUVs between groups and over time was evaluated based on both the statistical and clinical significance. Statistical significance was set as a P-value ≤0.05 (two-tailed) and the clinical significance followed the minimum clinically important difference (MID) in the AQoL-4D HSUVs of 0.06 for the Australian population (12). Linear mixed model (LMM) with random intercepts by patient identification number (ID) (the variable designating the clusters) using maximum likelihood estimation was conducted to estimate the impacts of OA on HRQoL (i.e.: 1, HSUVs; 2, independent living; 3, physical senses; 4, social relationships; and 5, psychological wellness) over ten years. We estimated two sets of separate LMM regressions (each comprising 5 regression equations): first set of regressions evaluated the HRQoL differences between those with and without OA; and the second one compared these differences among participants with knee and/or hip OA, those with other types of OA and those without OA (Supplement 5.1). We adjusted the models for age (50-59=0, 60-69=1, 70-80=2), sex (male=0, female=1), BMI (normal=0, overweight=1, obese=2), education (low=0, middle=1, high=2), employment status (employed=0, retired=1, unemployed=2, disability pensioner=3, others=4) and number of comorbidities (0, 1, 2, ≥3). Regression estimates and 95% confidence interval (CI) of the fixed effects are reported in the tables. All statistical analyses were performed using STATA (version16.0, StataCorp, College Station, Texas, USA).

5.5 Results

5.5.1 OA diagnosis

A total of 1,093 participants with complete information on OA diagnosis and HRQoL assessments at baseline were included in the analysis with the mean (standard deviation) age of 62.5 (7.51) years old and 51% of female. Of these, 871 (80%), 760 (70%), and 562 (51%) completed the 2.5, 5- and 10-years follow-up assessments. Participants lost to follow-up at 10 years were more likely to report presence of knee and/or hip OA, be aged between 70 and 80 years, and have lower levels of education at baseline than those completing 10-year follow-up

(Supplement 5.2). From Table 5.1, 36% (n=398) of participants were diagnosed with some form of OA at baseline (203 knee and/or hip OA, 195 with types of OA other than hip and/or knee). The proportion of those who reported OA at baseline was similar at all four time points.

Table 5.1. The distribution of baseline OA diagnosis at each time point

Time point	Participants with OA			Participants without OA	All participants
	Knee and/or hip OA	Other types of OA	Any OA	No OA	Any OA + No OA
	n (%)	n (%)	n (%)	n (%)	n (%)
Baseline	203 (19)	195 (18)	398 (36)	695 (64)	1,093 (100)
2.5 years	149 (17)	155 (18)	304 (35)	567 (65)	871 (100)
5 years	127 (17)	142 (19)	269 (35)	491 (65)	760 (100)
10 years	82 (15)	109 (19)	191 (34)	371 (66)	562 (100)

Notes: Knee and/or hip OA included at least one of the sites of knee and hip diagnosed with OA; other types of OA included at least one of the five joint sites other than knee and hip diagnosed with OA. OA= Osteoarthritis.

5.5.2 Characteristics of the study population

Table 5.2 shows the baseline characteristics of participants by OA diagnosis. Compared to those without OA, the OA group had higher proportions of females (61% vs. 46%), participants aged ≥60 (65% vs. 52%), and with ≥2 comorbidities (42% vs. 28%); and a lower proportion of employed participants (30% vs. 45%). Distributions of BMI and education were similar between the two groups. Compared to those with other types of OA, the knee and/or hip OA group had a higher proportion of obesity (36% vs. 25%), retired participants (45% vs. 38%), and participants with ≥3 comorbidities (24% vs. 18%).

Table 5.2 The baseline characteristic of participants by OA diagnosis

Variables	Without OA	With OA	Knee and/or hip OA	Other types of OA
Sex (n, %)				
Male	377 (54)	157 (39)	82 (40)	75 (38)
Female	318 (46)	241 (61)	121 (60)	120 (62)
Age (n, %)				
50-59	330 (48)	139 (35)	72 (36)	67 (34)
60-69	244 (35)	154 (39)	80 (39)	74 (38)
70-80	121 (17)	105 (26)	51 (25)	54 (28)
BMI (n, %)				

Normal (BMI<25. 0kg/m^2)	196 (28)	120 (30)	53 (26)	67 (34)
Overweight (25.0kg/m^2≤BMI≤29.9kg/m^2)	320 (46)	156 (39)	76 (37)	80 (41)
Obesity (BMI≥30.0kg/m^2)	179 (26)	122 (31)	74 (36)	48 (25)
Education [a] (n, %)				
Low	104 (15)	75 (19)	39 (19)	36 (18)
Middle	509 (73)	280 (70)	141 (69)	139 (71)
High	81 (12)	43 (11)	23 (11)	20 (10)
Employment (n, %)				
Employed	310 (45)	121 (30)	58 (29)	63 (32)
Retired	251 (36)	166 (42)	92 (45)	74 (38)
Unemployment	15 (2)	4 (1)	1 (0)	3 (2)
Disability pension	30 (4)	42 (11)	23 (11)	19 (10)
Others [b]	89 (13)	65 (16)	29 (14)	36 (18)
Number of comorbidity (n, %)				
0	277 (40)	98 (25)	50 (25)	48 (25)
1	228 (33)	133 (33)	59 (29)	74 (38)
2	115 (17)	83 (21)	45 (22)	38 (20)
≥3	75 (11)	84 (21)	49 (24)	35 (18)

Note: OA=osteoarthritis; BMI=body mass index; Other types of OA included OA at sites of neck, back, hands, shoulders and feet.

a: Education categorization was based on participants' highest education levels, low education included those having no formal qualification, middle education included holders of school or intermediate certificate, higher school or leaving certificate, trade/apprenticeship certificate/diploma, and high education included holders of university degree/higher degree.

b: others employment included home duties, students and sole parent pension

5.5.3 Linear mixed model regression of HSUVs

Table 5.3 shows the results of LMMs, comparing HSUVs and health-dimension scores between those with and without OA. Compared to participants without OA, those with OA had clinically important 0.07 (95% CI: -0.09, -0.05) units lower HSUVs at all time points. The mean baseline psychological wellness score of participants with OA was 0.04 (95% CI: -0.05, -0.03) units lower than those without OA and this difference remained stable over the ten years follow up period. The mean baseline independent living score of participants with OA was 0.02 (95% CI: -0.03, -0.01) units lower than those without OA and this difference increased to 0.04 units at 10-years follow-up. The mean baseline social relationship score of participants with OA was 0.02 (95% CI: -0.03, 0.00) units lower than those without OA. Supplement 5.3

provides the estimated marginal means of HRQoL scores over time for participants with and without OA.

Table 5.3. Estimates of linear mixed models for HRQoL comparing those with and without OA

	HSUVs	Independent living	Social relationships	Physical senses	Psychological wellness
OA diagnosis					
Without OA	Reference	Reference	Reference	Reference	Reference
With OA	**-0.07 (-0.09, -0.05)**	**-0.02 (-0.03, -0.01)**	**-0.02 (-0.03, 0.00)**	-0.01 (-0.02, 0.00)	**-0.04 (-0.05, -0.03)**
Timepoints					
Baseline	Reference	Reference	Reference	Reference	Reference
2.5years	-0.01 (-0.03, 0.00)	0.00 (-0.01, 0.00)	**-0.02 (-0.02, -0.01)**	0.00 (-0.01, 0.00)	0.00 (-0.01, 0.01)
5 years	0.00 (-0.01, 0.02)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.00 (0.00, 0.01)
10years	-0.02 (-0.04, 0.00)	**-0.02 (-0.03, -0.01)**	0.00 (-0.01, 0.02)	**-0.01 (-0.02, 0.00)**	0.00 (-0.01, 0.01)
OA diagnosis × time points					
OA × 2.5years	0.01 (-0.01, 0.03)	0.00 (-0.02, 0.01)	0.01 (0.00, 0.03)	0.00 (-0.01, 0.01)	0.01 (-0.01, 0.02)
OA × 5years	-0.01 (-0.03, 0.01)	-0.01 (-0.03, 0.00)	-0.01 (-0.02, 0.01)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)
OA × 10years	-0.01 (-0.04, 0.02)	**-0.02 (-0.04, 0.00)**	-0.01 (-0.03, 0.00)	0.01 (-0.01, 0.02)	0.01 (-0.01, 0.02)
Intercept	0.77 (0.74, 0.80)	1.00 (0.99, 1.02)	0.94 (0.92, 0.96)	0.90 (0.89, 0.92)	0.89 (0.87, 0.91)

Note: 95% CI=95% confidence interval, OA=osteoarthritis, HSUVs=health state utility values. HRQoL = health-related quality of life. Other types of OA include OA at sites of the neck, back, hands, shoulders, and feet. All the models were adjusted for age, sex, education, body mass index, employment and number of comorbidities. OA diagnosis × time points indicate the interaction effects between OA diagnosis and the time of follow-up. Bold indicates statistical significance (P-value < 5%).

Table 5.4 Estimates of linear mixed models for HRQoL scores comparing participants with different types of OA

	HSUVs	Independent living	Social relationships	Physical senses	Psychological wellness
OA diagnosis					
Without OA	Reference	Reference	Reference	Reference	Reference
Knee and/or hip OA	**-0.09 (-0.12, -0.06)**	**-0.04 (-0.05, -0.02)**	**-0.03 (-0.05, -0.02)**	-0.01 (-0.02, 0.01)	**-0.05 (-0.07, -0.04)**
Other types of OA	**-0.05 (-0.07, -0.02)**	-0.01 (-0.02, 0.01)	0.00 (-0.02, 0.02)	-0.01 (-0.02, 0.01)	**-0.03 (-0.05, -0.02)**
Timepoints					
Baseline	Reference	Reference	Reference	Reference	Reference

2.5years	-0.01 (-0.03, 0.00)	0.00 (-0.01, 0.00)	**-0.02 (-0.03, -0.01)**	0.00 (-0.01, 0.00)	0.00 (-0.01, 0.01)
5 years	0.00 (-0.01, 0.02)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.00 (-0.01, 0.01)	0.00 (0.00, 0.01)
10years	**-0.02 (-0.04, -0.01)**	**-0.02 (-0.03, -0.01)**	0.00 (-0.01, 0.02)	**-0.01 (-0.02, 0.00)**	0.00 (-0.01, 0.01)
OA diagnosis × time points					
Knee and/or hip OA × 2.5years	0.01 (-0.02, 0.04)	-0.01 (-0.03, 0.01)	0.01 (-0.01, 0.03)	0.00 (-0.01, 0.01)	0.01 (-0.01, 0.03)
Knee and/or hip OA × 5years	**-0.04 (-0.07, -0.01)**	**-0.03 (-0.05, -0.01)**	0.00 (-0.02, 0.02)	-0.01 (-0.03, 0.00)	0.00 (-0.02, 0.01)
Knee and/or hip OA × 10years	-0.03 (-0.06, 0.01)	**-0.03 (-0.05, -0.01)**	0.00 (-0.03, 0.02)	-0.01 (-0.02, 0.01)	0.00 (-0.02, 0.02)
Other types OA × 2.5years	0.01 (-0.01, 0.04)	0.00 (-0.02, 0.02)	0.01 (-0.01, 0.03)	0.01 (-0.01, 0.02)	0.00 (-0.01, 0.02)
Other types OA × 5years	0.01 (-0.02, 0.04)	0.00 (-0.02, 0.02)	-0.01 (-0.03, 0.01)	0.02 (0.00, 0.03)	0.00 (-0.01, 0.02)
Other types OA × 10years	0.00 (-0.03, 0.03)	-0.01 (-0.03, 0.01)	-0.02 (-0.05, 0.00)	0.02 (0.00, 0.03)	0.01 (-0.01, 0.03)
Intercept	0.77 (0.74, 0.80)	1.00 (0.98, 1.02)	0.94 (0.92, 0.96)	0.90 (0.89, 0.92)	0.89 (0.87, 0.91)

Note: 95% CI=95% confidence interval, OA=osteoarthritis, HSUVs=health state utility values. HRQoL = health-related quality of life. Other types of OA include OA at the neck, back, hands, shoulders and feet. All the models were adjusted by age, sex, education, body mass index, employment and number of comorbidities. OA diagnosis × time points indicate the interaction effects between OA diagnosis and the time of follow-up. Bold indicates statistical significance (P-value < 5%).

Table 5.4 shows the results of LMMs, comparing HSUVs and health-dimension scores of those with knee and/or hip OA, and of those with other types of OA to those without OA. Baseline HSUVs of people with knee and/or hip OA and with other types of OA were lower than those without OA, with a mean clinically important difference of 0.09 (95% CI: -0.12, -0.06) and statistically significant 0.05 (95% CI: -0.07, -0.02) units, respectively. For those with knee and/or hip OA, baseline scores on the health dimensions of independent living (-0.04 [95% CI: -0.05, -0.02] units), social relationships (-0.03 [95% CI: -0.05, -0.02] units) and psychological wellness (-0.05 [95% CI: -0.07, -0.04] units) were lower; and for those with other types of OA, the only difference was in psychological wellness (-0.03 [95% CI: -0.05, -0.02] units) score. Whilst the mean HSU difference between those with other types of OA and those without OA did not change overtime, participants with knee and/or hip OA recorded a further reduction of 0.04 units at 5-years follow-up, with an additional 0.03 units reduction in their independent living score over the same period. The physical senses dimension score was similar across OA groups. The HSUVs for the two OA sub-groups stayed similar up to 2.5 years follow-up, then widened such that HSUVs of knee and/or hip OA were up to 0.09 units lower than other types of OA after 5-years follow-up. Supplement 5.4 provides the estimated marginal means of HRQoL scores over time for two OA groups of and those without OA.

5.6 Discussion

This is the first long-term prospective study to investigate the cross-sectional and temporal differences in HRQoL in the forms of HSUVs and health-dimension scores of people with and without OA using a population-based cohort. Participants with OA had clinically important 0.07 (95% CI: -0.09, -0.05) units lower HSUVs over ten years compared to participants without OA. The main dimensions impacted by OA were psychological wellness, independent living, and social relationships. Whilst psychological wellness was equally impacted for both knee and/or hip OA and other types of OA, independent living and social relationships scores were substantially lower for people with knee and/or hip OA and the effect on the former increased over time. Appropriate and timely support to maintain independent living, social relationships and psychological wellness of OA patients may therefore have the potential to minimize the negative HRQoL impacts of OA.

The mean HSU of OA participants was 0.07 units lower than for people without OA over ten years follow-up, suggesting that the difference is likely to be clinically important as it exceeds the minimum clinically important difference (0.06) of the Australian population. Our results

are similar to the findings from the AQoL-4D based nationally representative cross-sectional Australian National Survey of Mental Health and Wellbeing, which found 0.07 unit difference in HSUVs between those with and without arthritis (13). However, no suitable published data were available to compare with our estimates of longitudinal changes in HSUVs, so further research to confirm our findings in other populations would be helpful.

Our study is the first to compare the difference in HSUVs of knee and/or hip OA and other (relatively understudied) type of OA, and to track this difference over a period of 10 years. Our findings indicate that HSUVs in people with knee and/or hip OA were 0.04 units lower compared to those with other types of OA up to 2.5 years follow-up, and the HSU gap widened (up to 0.09 units) thereafter. This was driven predominantly by the increased impacts of knee and/or hip OA on independent living over time. We therefore recommend the use of interventions that are tailored to specific OA types and stages of disease process to maintain patients' HSUVs.

A further important finding is that the mean HSU of participants with other types of OA was 0.05 (95% CI: -0.07, -0.02) units lower than for people without OA. Previous research has mostly focused on the HRQoL consequences of knee and/or hip OA with few studies exploring the HRQoL consequences of OA at other joint sites including hand, feet, shoulder, neck and back (1, 8, 14). In a cross-sectional study, patients with hand OA had 0.13 units lower mean HSU (measured with the SF-6D) than healthy controls (15). In another cross-sectional study the physical function score of those with feet OA was significantly lower compared to those without OA, however, the HSU was not assessed in this study (16). Given our results and these other published data, future researchers are recommended to give more attention to the HRQoL impact of these joint sites.

The health dimension most negatively impacted by OA was psychological wellness, and hence this is the key overall driver of lower HSUVs in patients with OA. Upon investigating the impact of OA on health dimensions by OA sub-groups, both knee and/or hip OA and other types of OA had noticeable impact on psychological wellness. Our results are consistent with previous findings showing worse scores in pain, sleeping quality, and mental health (the key determinants of psychological wellness) in people with OA (7). We further found that the impact on psychological wellness was sustained over time in both OA sub-groups. As psychosocial wellness is impacted irrespective of the type of OA and the time of assessment, we expect that the introduction and promotion of interventions/preventions to effectively

manage the psychological wellness (e.g.: pain, sleep, and anxiety/depression) of people with all types of OA throughout the disease process is important to minimise the negative HSU impact of OA.

Knee and/or hip OA had a noticeable (0.03 units) impact on social relationships at baseline. This aligns with previous findings suggesting that people with knee or hip OA may have difficulties keeping or developing their social network due to pain and disability which obstruct access to social support (17, 18). As social supports play a critical role in improving HRQoL (19-21), we recommend improved access to social support for people with OA, especially those with knee and/or hip OA.

There was a noticeable impact of knee and/or hip OA on independent living (0.04 units) at baseline and this impact increased over time. This is expected as the knee and hip are the main load-bearing joints with huge potential to impact mobility, and consistent with the chronic progressive nature of knee and/or hip OA (22). Our results are supported by previous literature suggesting the increased likelihood of developing difficulty over time in people with knee and/or hip OA's lower extremity tasks, activities of daily living (23), and self-reported functioning (24). Increased and targeted support to maintain independent living of people with knee and/or hip OA, particularly those who have been living with the disease for longer, should therefore be on the agenda for healthcare providers and policy makers.

As OA impacts each health dimension with different intensity and timing, so interventions to improve HRQoL may need to be tailored to specific OA types and health dimensions, and the intensity of the interventions should be tailored over time. Such interventions might include those to manage psychological wellness targeting at pain, sleep and anxiety/depression, maintain social relationships and independent living. Support to maintain psychological wellness of people with OA should be provided irrespective of the type and disease duration. However, support to maintain independent living could be more relevant to people with knee and/or hip OA, particularly those with longer disease durations.

The strengths of our study include the random, population-based, longitudinal cohort with more than ten years follow-up, adoption of a preference-based MAUI for the first time to assess HRQoL for Australians with OA, and the investigation of the impact of OA on various health dimensions of people with OA. The difference on HSUVs between those with and without OA was the mean difference of these two groups from a representative, population-based cohort, which means this difference is representative for the comparison of Australian population with

and without OA which provides a broad indication of the nature and extent of the mean HRQoL impacts of OA and can be used to generate the estimates of OA's disease burden in Australia and to predict the long-term disease outcomes including life expectancy and quality adjusted life years.

Our study was subject to several limitations. First, we used the patient-reported OA diagnosis (reported as doctor-diagnosed) instead of clinically defined OA data, which may introduce some misclassification bias. The key reason behind using self-reported OA diagnosis was study's focus on OA of all joint site with the aim of estimating the overall HRQoL impacts of OA for the cohort. Whilst TASOAC collected radiographic and pain symptom data, those were only collected for a select few individual joint sites (i.e., knee and hip) and the timing of collecting these data varied by affected OA joint sites (leading to further complexities in relation to using these data to construct uncertainty intervals to quantify the sensitivity of the results to the uncertainties facing exposure definition). Notably, as the TASOAC database already had low number of participants completing all four waves of the survey, relying on clinical data would substantially reduce the sample sizes further, so no meaningful conclusions could be drawn (Supplement 5.5). However, this may have a minimal impact on our results as the accuracy of self-reported OA and its adequate agreement with clinically assessed OA has been demonstrated in the past (25-27). Second, we lacked information on duration or severity of OA, treatment histories and severity of comorbidities which may have substantial impacts on OA patients' HRQoL, so could not adjust for these factors in our model. Because we did not have data on treatment histories, our study was unable to identify the HRQoL impact of various OA treatments. Therefore, our HSU estimates are not helpful in the cost-effectiveness analyses of individual OA treatments. However, as our data provided a broad indication of the nature and extent of the mean HRQoL difference between people with and without OA, the estimates can be used in the health economic models to estimate OA's overall disease burden and to predict long-term disease outcomes (e.g.: life expectancy and quality adjusted life years) of Australians with OA. Third, due to small numbers, we could not investigate the difference in HSUVs between each OA joint sites of neck, back, hands, shoulders, and feet, so all non-load-bearing OA joint sites were grouped into a single category. Fourth, the TASOAC study was only conducted for older people living in Tasmania, with subsequent implications for generalizability of our findings, particularly to non-Caucasian and younger people with OA. Fifth, the AQoL-4D algorithm used to calculate HSUs was generated from an Australian general population which may not perfectly reflect the preference/norms of an OA population;

therefore it might not perfectly fit in the population with OA. However, use of general population norms is a standard practice. Additionally, some patients without OA at baseline may have developed OA over time, which may have consequences for their HRQoL. However, our analyses did not consider this possibility as TASOAC did not collect information on OA development at the seven sites during the follow-up period. To provide an indication of the impact of OA's incidence after baseline on our results, additional investigation was provided in Supplement 5.5, which showed that our results were likely underestimated the HRQoL impact of OA.

Conclusion

OA negatively impacts HRQoL, with differences remaining in the long term and after adjustment for possible confounders. Psychological wellness, independent living and social relationships were the key health dimensions driving the reduction in HSUVs in people with OA. Whilst psychological wellness was equally impacted for both knee and/or hip OA and other types of OA, independent living and social relationships scores were substantially lower for people with knee and/or hip OA and the effect on the former increased over time.

Authors' contributions

GJ designed the TASOAC study that provided the data for our analyses. TZ and HA conceived and designed the initial concept and method of this manuscript, DA, TW, AP and BG modified and improved the original concept and methods. TZ accomplished the setting up of database and performed necessary data analyses and drafted the manuscript. All authors were involved in the manuscript review and final approval.

Funding and conflicts

No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this article. There are no conflicts of interest.

Acknowledgements

We would like to thank all participants of TASOAC project (aimed at identifying the environmental, genetic, and biochemical factors associated with the development and progression of osteoarthritis at multiple sites) who provided the valuable data that informed the comprehensiveness of this paper. We would also like to thank all medical practitioners who were associated with TASOAC study. Finally, we thank to Mr Petr Otahal (Statistical Officer,

Menzies Institute for Medical Research, University of Tasmania) for his help with the statistical analysis of this study.

5.7 References

1. Wilson R, Blakely T, Abbott JH. Radiographic knee osteoarthritis impacts multiple dimensions of health-related quality of life: data from the Osteoarthritis Initiative. Rheumatology. 2018;57(5):891-9.

2. Ahmad H, Taylor BV, van der Mei I, Colman S, O'Leary BA, Breslin M, et al. The impact of multiple sclerosis severity on health state utility values: Evidence from Australia. Multiple Sclerosis Journal. 2017;23(8):1157-66.

3. Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the Measurement of Health Utilities in Published Cost-Utility Analyses. Value in Health. 2006;9(4):213-8.

4. Stevens KJ. How Well Do the Generic Multi-attribute Utility Instruments Incorporate Patient and Public Views Into Their Descriptive Systems? The patient. 2016;9(1):5-13.

5. Richardson J, McKie J, Bariola E. Review and critique of health related multi attribute utility instruments: Monash University, Business and Economics, Centre for Health Economics; 2011.

6. Rabenda V, Manette C, Lemmens R, Mariani A-M, Struvay N, Reginster J-Y. Prevalence and impact of osteoarthritis and osteoporosis on health-related quality of life among active subjects. Aging Clinical and Experimental Research. 2007;19(1):55-60.

7. Dominick KL, Ahern FM, Gold CH, Heller DA. Health-related quality of life among older adults with arthritis. Health and Quality of life Outcomes. 2004;2(1):5.

8. Ackerman IN, Bucknill A, Page RS, Broughton NS, Roberts C, Cavka B, et al. The substantial personal burden experienced by younger people with hip or knee osteoarthritis. Osteoarthritis Cartilage. 2015;23(8):1276-84.

9. Balogun S, Winzenberg T, Wills K, Scott D, Jones G, Aitken D, et al. Prospective Associations of Low Muscle Mass and Function with 10-Year Falls Risk, Incident Fracture and Mortality in Community-Dwelling Older Adults. The journal of nutrition, health & aging. 2017;21(7):843-8.

10. Centre for Health Economics MU. Assessment of Quality of Life [Available from: https://www.aqol.com.au/index.php/aqol-publications#4D.

11. James PT, Leach R, Kalamara E, Shayeghi M. The Worldwide Obesity Epidemic. Obesity Research. 2001;9(S11):228S-33S.

12. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Australian and New Zealand journal of public health. 2005;29(2):136-42.

13. Hawthorne G, Korn S, Richardson J. Population norms for the AQoL derived from the 2007 Australian National Survey of Mental Health and Wellbeing. Aust N Z J Public Health. 2013;37(1):7-16.

14. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging clinical and experimental research. 2005;17(4):255-63.

15. Slatkowsky-Christensen B, Mowinckel P, Loge JH, Kvien TK. Health-related quality of life in women with symptomatic hand osteoarthritis: A comparison with rheumatoid arthritis patients, healthy controls, and normative data. Arthritis care & research. 2007;57(8):1404-9.

16. Bergin SM, Munteanu SE, Zammit GV, Nikolopoulos N, Menz HB. Impact of first metatarsophalangeal joint osteoarthritis on health-related quality of life. Arthritis care & research. 2012;64(11):1691-8.

17. Swift A. Osteoarthritis 3: impact on patients. . Nursing Times [online]. 2012 108: online only.

18. Siviero P, Veronese N, Smith T, Stubbs B, Limongi F, Zambon S, et al. Association Between Osteoarthritis and Social Isolation: Data From the EPOSA Study. J Am Geriatr Soc. 2020;68(1):87-95.

19. Ethgen O, Vanparijs P, Delhalle S, Rosant S, Bruyère O, Reginster J-Y. Social support and health-related quality of life in hip and knee osteoarthritis. Quality of Life Research. 2004;13(2):321-30.

20. Ferreira V, Sherman A. The relationship of optimism, pain and social support to well-being in older adults with osteoarthritis. Aging and Mental Health. 2007;11(1):89-98.

21. Krol B, Sanderman R, Suurmeijer TP. Social support, rheumatoid arthritis and quality of life: concepts, measurement and research. Patient education and counseling. 1993;20(2-3):101-20.

22. Leyland KM, Hart DJ, Javaid MK, Judge A, Kiran A, Soni A, et al. The natural history of radiographic knee osteoarthritis: a fourteen-year population-based cohort study. Arthritis and rheumatism. 2012;64(7):2243-51.

23. Ling SM, Xue QL, Simonsick EM, Tian J, Bandeen-roche K, Fried LP, et al. Transitions to mobility difficulty associated with lower extremity osteoarthritis in high functioning older women: Longitudinal data from the Women's Health and Aging Study II. Arthritis care & research. 2006;55(2):256-63.

24. Miller ME, Rejeski WJ, Messier SP, Loeser RF. Modifiers of change in physical functioning in older adults with knee pain: the Observational Arthritis Study in Seniors (OASIS). Arthritis care & research. 2001;45(4):331-9.

25. Peeters GMEE, Alshurafa M, Schaap L, de Vet HCW. Diagnostic accuracy of self-reported arthritis in the general adult population is acceptable. Journal of Clinical Epidemiology. 2015;68(4):452-9.

26. Parsons C, Clynes M, Syddall H, Jagannath D, Litwic A, van der Pas S, et al. How well do radiographic, clinical and self-reported diagnoses of knee osteoarthritis agree? Findings from the Hertfordshire cohort study. Springerplus. 2015;4:177-.

27. March LM, Schwarz JM, Carfrae BH, Bagge E. Clinical validation of self-reported osteoarthritis. Osteoarthritis Cartilage. 1998;6(2):87-93.

5.8 Supplements

Supplement 5A: The publication of "Chapter 5: Cross-sectional and temporal differences in health-related quality of life of people with and without OA: a ten-year prospective study"

RHEUMATOLOGY

Rheumatology 2021;00:1–8
doi:10.1093/rheumatology/keaa787

Original article

Cross-sectional and temporal differences in health-related quality of life of people with and without osteoarthritis: a 10-year prospective study

Ting Zhao [1], Hasnat Ahmad[1], Tania Winzenberg[1], Dawn Aitken[1], Barbara de Graaff[1], Graeme Jones [1] and Andrew J. Palmer[1,2]

Abstract

Objective. To describe the impact of OA on health-related quality of life (HRQoL) in the forms of health state utilities (HSUs) and health-dimension scores, and to compare the longitudinal changes in HRQoL for people with and without OA, using an Australian population-based longitudinal cohort.

Methods. Participants of the Tasmanian Older Adult Cohort with data on OA diagnosis and HRQoL were included [interviewed at baseline ($n = 1093$), 2.5 years ($n = 871$), 5 years ($n = 760$) and 10 years ($n = 562$)]. HRQoL was assessed using the Assessment of Quality of Life four-dimensions and analysed using multivariable linear mixed regressions.

Results. Compared with participants without OA, HSUs for those with OA were 0.07 (95% confidence interval: 0.09, 0.05) units lower on average over 10 years. HSUs for participants with knee and/or hip OA were similar to those with other types of OA at the 2.5 year follow-up and then diverged, with HSUs of the former being up to 0.09 units lower than the latter. Those with OA had lower scores for psychological wellness, independent living and social relationships compared with those without OA. Independent living and social relationships were mainly impacted by knee and/or hip OA, with the effect on the former increasing over time.

Conclusion. Interventions to improve HRQoL should be tailored to specific OA types, health dimensions, and times. Support for maintaining psychological wellness should be provided, irrespective of OA type and duration. However, support for maintaining independent living could be more relevant to knee and/or hip OA patients living with the disease for longer.

Key words: osteoarthritis, HRQoL, HSUs, the Assessment of Quality of Life four-dimensions

Rheumatology Key messages

- Osteoarthritis negatively impacts multiple facets of HRQoL, but with varying intensity and timing.
- Interventions to improve HRQoL should be tailored to specific osteoarthritis types, health dimensions, and times.

Introduction

OA is one of the most common chronic joint diseases, affecting knees, hips and the small joints of the hands. OA is characterized by joint pain, stiffness, swelling, loss of function, and disability, which in turn negatively impacts individuals' health-related quality of life (HRQoL) [1]. HRQoL focuses on the impact of physical and psychosocial health status on quality of life. It can be reflected as an overall index defined by health state utility (HSU) or scores for different dimensions of health. HSUs measure the strength of preference for a given health state and are represented as a number between 0 (death) and 1 (optimal health). Health states worse than death may exist, with negative HSUs assigned [2].

The indirect measurement of HSUs using multi-attribute utility instruments (MAUIs) such as the EuroQoL five-dimensions (EQ-5D), Assessment of Quality of Life

[1]The Health Economics Research Unit, Menzies Institute for Medical Research, University of Tasmania, Hobart and [2]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

Submitted 29 June 2020; accepted 15 October 2020

Correspondence to: Andrew J. Palmer, Menzies Research Institute Tasmania, University of Tasmania, Medical Science 1 Building, 17 Liverpool St, Hobart, Tasmania 7000, Australia
E-mail: Andrew.palmer@utas.edu.au

CLINICAL SCIENCE

four-dimensions (AQoL-4D) and Short-Form-6D (SF-6D) has become increasingly popular [3]. A MAUI is a set of questions and response categories that seek to describe a person's health, and an accompanying formula or a set of weights elicited from a sample of the general population for converting responses into HSUs [4, 5]. In addition to HSUs, some MAUIs (e.g.: AqoL-4D) can also provide summary scores for specific health dimensions (e.g. physical dimension and psychosocial dimensions).

Studies assessing the impact of OA on HRQoL have mostly been undertaken in Europe and the USA, and are cross-sectional in nature [1, 6, 7]. There is a paucity of data on HRQoL impacts of OA in Australia. For instance, a 2015 Australian study estimated the impacts of knee and/or hip OA on HRQoL, using data from a multicentre, cross-sectional survey [8]. While this added to our understanding of knee and/or hip OA impacts on HRQoL of younger people, it used cross-sectional data from a small number ($n = 21$) of community participants who had knee and hip OA. Furthermore, no study has investigated longitudinal changes in the mean HSUs or individual health-dimension scores of people with OA compared with those without OA. Our study aims to extend the scope of previous research by investigating the cross-sectional and longitudinal differences in HRQoL in the forms of HSUs and health-dimension scores of people with OA compared with those without OA using an Australian population-based cohort.

Methods

Study design

The Tasmanian Older Adult Cohort (TASOAC) is a prospective, population-based study aimed at identifying the environmental, genetic and biochemical factors associated with the development and progression of OA at multiple sites (knee, hip, hand and spine). Further details about the TASOAC study have been described elsewhere [9]. Participants aged 50–80 years in 2002 were selected from the electoral roll in Southern Tasmania, using sex-stratified simple random sampling without replacement. Participants were excluded if they resided in an aged care facility or were unable to have a knee MRI scan. A total of 1099 adults (response rate = 57%) consented to participate in the study. The research was approved by the Southern Tasmanian Health and Medical Human Research Ethics Committee, and written informed consent to participate in the study was obtained from all participants. Participants completed questionnaires and attended necessary clinical tests (not reported in this study). Data were collected using interview-type questionnaires administered during four periods: baseline (from February 2002 to September 2004), 2.5 years (range 1.4 – 4.8 years) later, 5 years (range 3.6 – 6.9 years) later and 10 years later (range 9 – 13 years).

Measurement of health-related quality of life (HRQoL)

HRQoL was assessed at baseline, 2.5, 5 and 10 years using the AQoL-4D questionnaire, which consists of 12 items covering four health dimensions, each with three items and four response levels (0, 1, 2 and 3). The four dimensions and their corresponding items are: (i) independent living (self-care, activities of daily living, and mobility), (ii) physical senses (sight, hearing and communication), (iii) social relationships (social isolation, relationships and family role) and (iv) psychological wellbeing (sleep, anxiety and pain). The items were combined according to the AQoL-4D algorithm to calculate dimension scores and an overall index of HSUs [10].

Diagnosis of OA

At baseline, participants were asked, 'Have you had been told by a doctor that you have OA at any of these sites?' Seven sites were listed: neck, back, hands, shoulders, hips, knees and feet. Participants were given the choice between answering 'yes' or 'no', and all participants answering 'yes' were considered to have been diagnosed with OA.

Categorization of OA patients

We first categorized our sample into two groups based on their baseline OA diagnosis: (1) without OA (no joint site diagnosed with OA); (2) with OA (one or more of the seven joint sites diagnosed with OA). We further classified OA participants into two groups based on their OA types: (2a) participants with knee and/or hip OA (at least one of the sites of knee and hip diagnosed with OA); and (2b) other types of OA (at least one of the five joint sites other than knee and hip diagnosed with OA).

Other characteristics

Data on sex and date of birth were collected at baseline. Weight was measured to the nearest 0.1 kg (with no shoes/socks/bulky clothing/headwear) using a single pair of calibrated electronic scales (Seca Delta Model 707). Height was measured to the nearest 0.1 cm (barefooted) using a stadiometer. Weight and height were collected at all time points. BMI was calculated as $\{\text{weight (kg)}/[\text{height (m)}]^2\}$, and participants were grouped into three categories of BMI [normal weight ($\text{BMI} < 25.0\ \text{kg/m}^2$), overweight ($25.0\ \text{kg/m}^2 \leq \text{BMI} \leq 29.9\ \text{kg/m}^2$) and obese ($\text{BMI} \geq 30.0\ \text{kg/m}^2$)], following the WHO criteria [11]. Education data was collected at baseline, and participants were grouped into three categories [low education (having no formal qualification), middle education (holders of school or intermediate certificate, higher school or leaving certificate, trade/apprenticeship certificate/diploma) or high education (holders of university degree/higher degree)], based on their highest education levels. Employment status was collected at all time points, and participants were categorized into five groups [1, employed/self-employed

(part-time/full-time); 2, retired; 3, unemployed; 4, disability pensioners; 5, others (including people engaged in home duties, students, and those receiving a sole parent pension)]. Presence or absence of 10 comorbidities were assessed at all time points by asking, 'Have you ever been diagnosed by a doctor as having any of the following: diabetes, heart attack, hypertension, thrombosis, asthma, bronchitis/emphysema, osteoporosis, hyperthyroidism, hypothyroidism, or RA?' The number of comorbidities was then calculated as the total number of diagnosed chronic conditions at each time point.

Statistical analysis

We summarized the demographic and other features of our sample using descriptive statistics. The difference in HSUs between groups and over time was evaluated based on both the statistical and clinical significance. Statistical significance was set as a *P*-value $\leq$0.05 (two-tailed), and the clinical significance followed the minimum clinically important difference (MID) in the AQoL-4D HSUs of 0.06 for the Australian population [12]. Linear mixed model (LMM) with random intercepts by patient identification number (ID) (the variable designating the clusters) using maximum likelihood estimation was conducted to estimate the impacts of OA on HRQoL (i.e.: 1, HSUs; 2, independent living; 3, physical senses; 4, social relationships; and 5, psychological wellness) over 10 years. We estimated two sets of separate LMM regressions (each comprising 5 regression equations): first set of regressions evaluated the HRQoL differences between those with and without OA; and the second one compared these differences among participants with knee and/or hip OA, those with other types of OA and those without OA (Supplementary Data S1, available at *Rheumatology* online). We adjusted the models for age (50–59 = 0, 60–69 = 1, 70–80 = 2), sex (male = 0, female = 1), BMI (normal = 0, overweight = 1, obese = 2), education (low = 0, middle = 1, high = 2), employment status (employed = 0, retired = 1, unemployed = 2, disability pensioner = 3, others = 4) and number of comorbidities (0, 1, 2, $\geq$3). Regression estimates and 95% confidence interval (CI) of the fixed effects are reported in the tables. All statistical analyses were performed using STATA (version16.0, StataCorp, College Station, Texas, USA).

Results

OA diagnosis

A total of 1093 participants with complete information on OA diagnosis and HRQoL assessments at baseline were included in the analysis. Of these, 871 (80%), 760 (70%) and 562 (51%) completed the 2.5, 5- and 10-year follow-up assessments. Participants lost to follow-up at 10 years were more likely to report presence of knee and/or hip OA, be aged between 70 and 80 years, and have lower levels of education at baseline than those completing 10-year follow-up (Supplementary Table S1, available at *Rheumatology* online). From Table 1, 36% (n = 398) of participants were diagnosed with some form of OA at baseline (203 with knee and/or hip OA, 195 with types of OA other than hip and/or knee). The proportion of those who reported OA at baseline was similar at all four time points.

Characteristics of the study population

Table 2 shows the baseline characteristics of participants by OA diagnosis. Compared with those without OA, the OA group had higher proportions of females (61% *vs* 46%), participants aged $\geq$60 (65% *vs* 52%), and participants with $\geq$2 comorbidities (42% *vs* 28%), and a lower proportion of employed participants (30% *vs* 45%). Distributions of BMI and education were similar between the two groups. Compared with those with other types of OA, the knee and/or hip OA group had a higher proportion of obesity (36% *vs* 25%), retired participants (45% *vs* 38%) and participants with $\geq$3 comorbidities (24% *vs* 18%).

Linear mixed-model regression of HSUs

Table 3 shows the results of LMMs, comparing HSUs and health-dimension scores between those with and without OA. Compared with participants without OA, those with OA had clinically important difference of 0.07 (95% CI: 0.09, 0.05) units lower HSUs at all time points. The mean baseline psychological wellness score of participants with OA was 0.04 (95% CI: 0.05, 0.03) units lower than those without OA, and this difference remained stable till the 10-year follow-up period. The mean baseline independent living score of participants with OA was 0.02 (95% CI: 0.03, 0.01) units lower than those without OA, and this difference increased to 0.04 units by the 10-year follow-up. The mean baseline social relationship score of participants with OA was 0.02 (95% CI: −0.03, 0.00) units lower than those without OA. Supplementary Figure S1, available at *Rheumatology* online, provides the estimated marginal means of HRQoL scores over time for participants with and without OA.

Table 4 shows the results of LMMs, comparing HSUs and health-dimension scores of those with knee and/or hip OA, and of those with other types of OA to those without OA. Baseline HSUs of people with knee and/or hip OA and with other types of OA were lower than those without OA, with a mean clinically important difference of 0.09 (95% CI: 0.12, 0.06) and statistically significant difference of 0.05 (95% CI: 0.07, 0.02) units, respectively. For those with knee and/or hip OA, baseline scores on the health dimensions of independent living [0.04 (95% CI: 0.05, 0.02) units], social relationships [0.03 (95% CI: 0.05, 0.02) units] and psychological wellness [0.05 (95% CI: 0.07, 0.04) units] were lower; for those with other types of OA, the only difference was in psychological wellness [0.03 (95% CI: 0.05, 0.02) units] score. While the mean HSU difference between those with other types of OA and those without OA did not change over time, participants with knee and/or hip OA recorded a further reduction of 0.04 units at the 5-year

TABLE 1. The distribution of baseline OA diagnosis at each time point

Time point	Participants with OA			Participants without OA	All participants
	Knee and/or hip OA *n* (%)	Other types of OA *n* (%)	Any OA *n* (%)	No OA *n* (%)	Any OA + No OA *n* (%)
Baseline	203 (19)	195 (18)	398 (36)	695 (64)	1093 (100)
2.5 years	149 (17)	155 (18)	304 (35)	567 (65)	871 (100)
5 years	127 (17)	142 (19)	269 (35)	491 (65)	760 (100)
10 years	82 (15)	109 (19)	191 (34)	371 (66)	562 (100)

Knee and/or hip OA included at least one of the sites of knee and hip diagnosed with OA; other types of OA included at least one of the five joint sites other than knee and hip diagnosed with OA.

TABLE 2. The baseline characteristics of participants by OA diagnosis

Variables	Without OA	With OA	Knee and/or hip OA	Other types of OA
Sex (*n*, %)				
Male	377 (54)	157 (39)	82 (40)	75 (38)
Female	318 (46)	241 (61)	121 (60)	120 (62)
Age (*n*, %)				
50–59	330 (48)	139 (35)	72 (36)	67 (34)
60–69	244 (35)	154 (39)	80 (39)	74 (38)
70–80	121 (17)	105 (26)	51 (25)	54 (28)
BMI (*n*, %)				
Normal (BMI $<$ 25. 0 kg/m^2)	196 (28)	120 (30)	53 (26)	67 (34)
Overweight (25.0 kg/m^2 $<$BMI $<$ 29.9 kg/m^2)	320 (46)	156 (39)	76 (37)	80 (41)
Obesity (BMI $\geq$ 30.0 kg/m^2)	179 (26)	122 (31)	74 (36)	48 (25)
Education[a] (*n*, %)				
Low	104 (15)	75 (19)	39 (19)	36 (18)
Middle	509 (73)	280 (70)	141 (69)	139 (71)
High	81 (12)	43 (11)	23 (11)	20 (10)
Employment (*n*, %)				
Employed	310 (45)	121 (30)	58 (29)	63 (32)
Retired	251 (36)	166 (42)	92 (45)	74 (38)
Unemployment	15 (2)	4 (1)	1 0	3 (2)
Disability pension	30 (4)	42 (11)	23 (11)	19 (10)
Others[b]	89 (13)	65 (16)	29 (14)	36 (18)
Number of comorbidities (*n*, %)				
0	277 (40)	98 (25)	50 (25)	48 (25)
1	228 (33)	133 (33)	59 (29)	74 (38)
2	115 (17)	83 (21)	45 (22)	38 (20)
$\geq$3	75 (11)	84 (21)	49 (24)	35 (18)

Other types of OA included OA at sites of neck, back, hands, shoulders and feet. [a]Education categorization was based on participants' highest education levels: low education included those having no formal qualification; middle education included holders of school or intermediate certificate, higher school or leaving certificate, or a trade/apprenticeship certificate/diploma; and high education included holders of university degree/higher degree. [b]Others employment included those participants engaged in home duties, students and those receiving the sole parent pension.

follow-up, with an additional 0.03 units reduction in their independent living score over the same period. The physical senses dimension score was similar across OA groups. The HSUs for the two OA subgroups stayed similar till the 2.5-year follow-up, then widened such that HSUs of knee and/or hip OA were up to 0.09 units lower than those for other types of OA at the 5-year follow-up. Supplementary Figure S2, available at *Rheumatology* online, provides the estimated marginal means of HRQoL scores over time for the two OA groups and those without OA.

Discussion

This is the first long-term prospective study to investigate the cross-sectional and temporal differences in

Table 3. Estimates of linear mixed models for HRQoL scores comparing those with and without OA

	HSUs	Independent living	Social relationships	Physical senses	Psychological wellness
OA diagnosis					
Without OA	Reference	Reference	Reference	Reference	Reference
With OA	−0.07 (−0.09, −0.05)*	−0.02 (−0.03, −0.01)*	−0.02 (−0.03, 0.00)*	−0.01 (−0.02, 0.00)	−0.04 (−0.05, −0.03)*
Time points					
Baseline	Reference	Reference	Reference	Reference	Reference
2.5 years	−0.01 (−0.03, 0.00)	0.00 (−0.01, 0.00)	−0.02 (−0.02, −0.01)*	0.00 (−0.01, 0.00)	0.00 (−0.01, 0.01)
5 years	0.00 (−0.01, 0.02)	0.00 (−0.01, 0.01)	0.00 (−0.01, 0.01)	0.00 (−0.01, 0.01)	0.00 (0.00, 0.01)
10 years	−0.02 (−0.04, 0.00)	−0.02 (−0.03, −0.01)*	0.00 (−0.01, 0.02)	−0.01 (−0.02, 0.00)*	0.00 (−0.01, 0.01)
OA diagnosis×time points					
OA×2.5 years	0.01 (−0.01, 0.03)	0.00 (−0.02, 0.01)	0.01 (0.00, 0.03)	0.00 (−0.01, 0.01)	0.01 (−0.01, 0.02)
OA×5 years	−0.01 (−0.03, 0.01)	−0.01 (−0.03, 0.00)	−0.01 (−0.02, 0.01)	0.00 (−0.01, 0.01)	0.00 (−0.01, 0.01)
OA×10 years	−0.01 (−0.04, 0.02)	−0.02 (−0.04, 0.00)*	−0.01 (−0.03, 0.00)	0.01 (−0.01, 0.02)	0.01 (−0.01, 0.02)
Intercept	0.77 (0.74, 0.80)	1.00 (0.99, 1.02)	0.94 (0.92, 0.96)	0.90 (0.89, 0.92)	0.89 (0.87, 0.91)

HSUs = health state utilities, HRQoL = health-related quality of life. Other types of OA include OA of the neck, back, hands, shoulders and feet. All the models were adjusted for age, sex, education, BMI, employment and number of comorbidities. OA diagnosis×time points indicates the interaction effects between OA diagnosis and the time of follow-up. Asterisk (*) indicates statistical significance (*P*-value < 5%).

Table 4. Estimates of linear mixed models for HRQoL scores comparing participants with different types of OA

	HSUs	Independent living	Social relationships	Physical senses	Psychological wellness
OA diagnosis					
Without OA	Reference	Reference	Reference	Reference	Reference
Knee and/or hip OA	−0.09 (−0.12, −0.06)*	−0.04 (−0.05, −0.02)*	−0.03 (−0.05, −0.02)*	−0.01 (−0.02, 0.01)	−0.05 (−0.07, −0.04)*
Other types of OA	−0.05 (−0.07, −0.02)*	−0.01 (−0.02, 0.01)	0.00 (−0.02, 0.02)	−0.01 (−0.02, 0.01)	−0.03 (−0.05, −0.02)*
Time points					
Baseline	Reference	Reference	Reference	Reference	Reference
2.5 years	−0.01 (−0.03, 0.00)	0.00 (−0.01, 0.00)	−0.02 (−0.03, −0.01)*	0.00 (−0.01, 0.00)	0.00 (−0.01, 0.01)
5 years	0.00 (−0.01, 0.02)	0.00 (−0.01, 0.01)	0.00 (−0.01, 0.01)	0.00 (−0.01, 0.01)	0.00 (0.00, 0.01)
10 years	−0.02 (−0.04, −0.01)*	−0.02 (−0.03, −0.01)*	0.00 (−0.01, 0.02)	−0.01 (−0.02, 0.00)*	0.00 (−0.01, 0.01)
OA diagnosis×time points					
Knee and/or hip OA×2.5 years	0.01 (−0.02, 0.04)	−0.01 (−0.03, 0.01)	0.01 (−0.01, 0.03)	0.00 (−0.01, 0.01)	0.01 (−0.01, 0.03)
Knee and/or hip OA×5 years	−0.04 (−0.07, −0.01)*	−0.03 (−0.05, −0.01)*	0.00 (−0.02, 0.02)	−0.01 (−0.03, 0.00)	0.00 (−0.02, 0.01)
Knee and/or hip OA×10 years	−0.03 (−0.06, 0.01)	−0.03 (−0.05, −0.01)*	0.00 (−0.03, 0.02)	−0.01 (−0.02, 0.01)	0.00 (−0.02, 0.02)
Other types OA×2.5 years	0.01 (−0.01, 0.04)	0.00 (−0.02, 0.02)	0.01 (−0.01, 0.03)	0.01 (−0.01, 0.02)	0.00 (−0.01, 0.02)
Other types OA×5 years	0.01 (−0.02, 0.04)	0.00 (−0.02, 0.02)	−0.01 (−0.03, 0.01)	0.02 (0.00, 0.03)	0.00 (−0.01, 0.02)
Other types OA×10 years	0.00 (−0.03, 0.03)	−0.01 (−0.03, 0.01)	−0.02 (−0.05, 0.00)	0.02 (0.00, 0.03)	0.01 (−0.01, 0.03)
Intercept	0.77 (0.74, 0.80)	1.00 (0.98, 1.02)	0.94 (0.92, 0.96)	0.90 (0.89, 0.92)	0.89 (0.87, 0.91)

HSUs = health state utilities, HRQoL = health-related quality of life. Other types of OA include OA of the neck, back, hands, shoulders and feet. All the models were adjusted by age, sex, education, BMI, employment and number of comorbidities. OA diagnosis×time points indicates the interaction effects between OA diagnosis and the time of follow-up. Asterisk (*) indicates statistical significance (*P*-value < 5%).

HRQoL in the forms of HSUs and health-dimension scores of people with and without OA using a population-based cohort. Participants with OA had a clinically important difference of 0.07 (95% CI: 0.09, 0.05) units lower HSUs over 10 years compared with participants without OA. The main dimensions impacted by OA were psychological wellness, independent living, and social relationships. While psychological wellness was equally impacted for both knee and/or hip OA and other types of OA, independent living and social

relationships scores were substantially lower for people with knee and/or hip OA, and the effect on the former increased over time. Appropriate and timely support to maintain independent living, social relationships and psychological wellness of OA patients may therefore have the potential to minimize the negative HRQoL impacts of OA.

The mean HSU of OA participants was 0.07 units lower than for people without OA over 10 years follow-up, suggesting that the difference is likely to be clinically important, because it exceeds the minimum clinically important difference (0.06) of the Australian population. Our results are similar to the findings from the AQoL-4D-based nationally representative cross-sectional Australian National Survey of Mental Health and Wellbeing, which found a 0.07-unit difference in HSUs between those with and without arthritis [13]. However, no suitable published data were available to compare with our estimates of longitudinal changes in HSUs, so further research to confirm our findings in other populations would be helpful.

Our study is the first to compare the difference in HSUs of knee and/or hip OA and other (relatively understudied) types of OA, and to track this difference over a period of 10 years. Our findings indicate that HSUs in people with knee and/or hip OA were 0.04 units lower compared with those with other types of OA for up to 2.5 years follow-up, and the HSU gap widened (up to 0.09 units) thereafter. This was driven predominantly by the increased impacts of knee and/or hip OA on independent living over time. We therefore recommend the use of interventions that are tailored to specific OA types and stages of disease process to maintain patients' HSUs.

A further important finding is that the mean HSU of participants with other types of OA was 0.05 (95% CI: 0.07, 0.02) units lower than for people without OA. Previous research has mostly focused on the HRQoL consequences of knee and/or hip OA, with few studies exploring the HRQoL consequences of OA at other joint sites, including hand, feet, shoulder, neck and back [1, 8, 14]. In a cross-sectional study, patients with hand OA had 0.13 units lower mean HSU (measured with the SF-6D) than healthy controls [15]. In another cross-sectional study, the physical function score of those with feet OA was significantly lower compared with those without OA; however, the HSU was not assessed in this study [16]. Given our results and these other published data, future researchers are recommended to give more attention to the HRQoL impact of OA of these joint sites.

The health dimension most negatively impacted by OA was psychological wellness, and hence this is the key overall driver of lower HSUs in patients with OA. Upon investigating the impact of OA on health dimensions by OA subgroups, both knee and/or hip OA and other types of OA had a noticeable impact on psychological wellness. Our results are consistent with previous findings showing worse scores in pain, sleeping quality, and mental health (the key determinants of psychological wellness) in people with OA [7]. We further found that the impact on psychological wellness was sustained over time in both OA subgroups. As psychosocial wellness is impacted irrespective of the type of OA and the time of assessment, we expect that the introduction and promotion of interventions/preventions to effectively manage the psychological wellness (e.g.: pain, sleep, and anxiety/depression) of people with all types of OA throughout the disease process is important to minimize the negative HSU impact of OA.

Knee and/or hip OA had a noticeable (0.03 units) impact on social relationships at baseline. This aligns with previous findings suggesting that people with knee or hip OA may have difficulties keeping or developing their social network due to pain and disability that obstruct access to social support [17, 18]. As social supports play a critical role in improving HRQoL [19–21], we recommend improved access to social support for people with OA, especially those with knee and/or hip OA.

There was a noticeable impact of knee and/or hip OA on independent living (0.04 units) at baseline, and this impact increased over time. This is expected, as the knee and hip are the main load-bearing joints, with huge potential to impact mobility, and consistent with the chronic progressive nature of knee and/or hip OA [22]. Our results are supported by previous literature suggesting the increased likelihood of people with knee and/or hip OA developing difficulty over time with lower extremity tasks, activities of daily living [23], and self-reported functioning [24]. Increased and targeted support to maintain independent living of people with knee and/or hip OA, particularly those who have been living with the disease for longer, should therefore be on the agenda for healthcare providers and policy makers.

As OA impacts each health dimension with different intensity and timing, so interventions to improve HRQoL may need to be tailored to specific OA types and health dimensions, and the intensity of the interventions should be tailored over time. Such interventions might include those to manage psychological wellness targeting at pain, sleep issues, anxiety/depression, maintaining social relationships and independent living. Support for maintaining psychological wellness of people with OA should be provided, irrespective of the type and disease duration. However, support for maintaining independent living could be more relevant to people with knee and/or hip OA, particularly those with longer disease durations.

The strengths of our study include the random, population-based, longitudinal cohort with more than 10 years of follow-up, adoption of a preference-based MAUI for the first time to assess HRQoL for Australians with OA, and the investigation of the impact of OA on various health dimensions of people with OA.

Our study was subject to several limitations. First, we used the patient-reported OA diagnosis (reported as doctor-diagnosed) instead of clinically defined OA data, which may introduce some misclassification bias. However, this may have a minimal impact on our results as the accuracy of self-reported OA and its adequate

agreement with clinically assessed OA has been demonstrated in the past [25–27]. Second, we lacked information on duration or severity of OA, treatment histories and severity of comorbidities, which may have substantial impacts on OA patients' HRQoL, so we could not adjust for these factors in our model. Third, due to small numbers, we could not investigate the difference in HSUs between the OA joint sites of neck, back, hands, shoulders and feet, so all non-load-bearing OA joint sites were grouped into a single category. Fourth, the TASOAC study was only conducted for older people living in Tasmania, with subsequent implications for generalizability of our findings, particularly to non-Caucasian and younger people with OA. Fifth, the HSU norm of AQoL-4D was generated from an Australian general population; therefore, it might not perfectly fit in the population with OA. However, use of general population norms is a standard practice. Additionally, some patients without OA at baseline may have developed OA over time, which may have consequences for their HRQoL. However, our analyses did not consider this possibility because TASOAC did not collect information on OA development at the seven sites during the follow-up period.

Conclusion

OA negatively impacts HRQoL, with differences remaining in the long term and after adjustment for possible confounders. Psychological wellness, independent living and social relationships were the key health dimensions driving the reduction in HSUs in people with OA. Appropriate and timely support to maintain these dimensions may therefore have the potential to substantially improve OA patients' HRQoL. Based on our findings, support for maintaining psychological wellness should be provided irrespective of OA type and disease duration. However, support for maintaining independent living could be more relevant to knee and/or hip OA patients who have been living with the disease for longer.

Acknowledgements

We would like to thank all participants of the TASOAC project (aimed at identifying the environmental, genetic and biochemical factors associated with the development and progression of OA at multiple sites), who provided the valuable data that informed the comprehensiveness of this paper. We would also like to thank all medical practitioners who were associated with the TASOAC study. Finally, we thank Mr Petr Otahal (Statistical Officer, Menzies Institute for Medical Research, University of Tasmania) for his help with the statistical analysis of this study. G.J. designed the TASOAC study that provided the data for our analyses. T.Z. and H.A. conceived and designed the initial concept and method for the present study. D.A., T.W., A.P. and B.G. modified and improved the original concept and methods. T.Z. accomplished the setting up of the database, performed necessary data analyses and drafted the manuscript. All authors were involved in review of the manuscript and its final approval.

Funding: No specific funding was received from any funding agency in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: The authors have declared no conflicts of interest.

Data availability statement

The data that was generated from this study will not be deposited in a public repository due to privacy and consent restrictions. De-identified data can be made available from the corresponding author on reasonable request, subject to a data sharing agreement.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Wilson R, Blakely T, Abbott JH. Radiographic knee osteoarthritis impacts multiple dimensions of health-related quality of life: data from the Osteoarthritis Initiative. Rheumatology 2018;57:891–9.

2 Ahmad H, Taylor BV, van der Mei I *et al.* The impact of multiple sclerosis severity on health state utility values: evidence from Australia. Mult Scler 2017;23:1157–66.

3 Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the measurement of health utilities in published cost-utility analyses. Value Health 2006;9:213–8.

4 Stevens KJ. How well do the generic multi-attribute utility instruments incorporate patient and public views into their descriptive systems? Patient 2016;9:5–13.

5 Richardson J, McKie J, Bariola E. Review and critique of health related multi attribute utility instruments: Monash University, Business and Economics, Centre for Health Economics; 2011. https://scholar.google.com.au/scholar?q=Review+and+critique+of+health+related+multi+attribute+utility+instruments:+Monash+University&hl=en&as_sdt=0&as_vis=1&oi=scholart.

6 Rabenda V, Manette C, Lemmens R *et al.* Prevalence and impact of osteoarthritis and osteoporosis on health-related quality of life among active subjects. Aging Clin Exp Res 2007;19:55–60.

7 Dominick KL, Ahern FM, Gold CH, Heller DA. Health-related quality of life among older adults with arthritis. Health Qual Life Outcomes 2004;2:5.

8 Ackerman IN, Bucknill A, Page RS *et al.* The substantial personal burden experienced by younger people with hip or knee osteoarthritis. Osteoarthritis Cartilage 2015;23:1276–84.

9 Balogun S, Winzenberg T, Wills K *et al.* Prospective associations of low muscle mass and function with 10-year falls risk, incident fracture and mortality in community-dwelling older adults. J Nutr Health Aging 2017;21:843–8.

10 Centre for Health Economics MU. Assessment of Quality of Life. https://www.aqol.com.au/index.php/aqolinstruments?id=54 (9 December 2020, date last accessed).

11 James PT, Leach R, Kalamara E, Shayeghi M. The worldwide obesity epidemic. Obes Res 2001;9:228S–33S.

12 Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Aust N Z J Public Health 2005;29:136–42.

13 Hawthorne G, Korn S, Richardson J. Population norms for the AQoL derived from the 2007 Australian National Survey of Mental Health and Wellbeing. Aust N Z J Public Health 2013;37:7–16.

14 Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clin Exp Res 2005;17:255–63.

15 Slatkowsky-Christensen B, Mowinckel P, Loge JH, Kvien TK. Health-related quality of life in women with symptomatic hand osteoarthritis: a comparison with rheumatoid arthritis patients, healthy controls, and normative data. Arthritis Rheum 2007;57:1404–9.

16 Bergin SM, Munteanu SE, Zammit GV, Nikolopoulos N, Menz HB. Impact of first metatarsophalangeal joint osteoarthritis on health-related quality of life. Arthritis Care Res 2012;64:1691–8.

17 Swift A. Osteoarthritis 3: impact on patients. Nursing Times [online]. 2012 108: online only. https://www.nursingtimes.net/clinical-archive/long-term-conditions/osteoarthritis-3-impact-on-patients-28-02-2012/ (2 December 2020, date last accessed).

18 Siviero P, Veronese N, Smith T, for the EPOSA Research Group *et al.* Association between osteoarthritis and social isolation: data from the EPOSA study. J Am Geriatr Soc 2020;68:87–95.

19 Ethgen O, Vanparijs P, Delhalle S *et al.* Social support and health-related quality of life in hip and knee osteoarthritis. Qual Life Res 2004;13:321–30.

20 Ferreira V, Sherman A. The relationship of optimism, pain and social support to well-being in older adults with osteoarthritis. Aging Ment Health 2007;11:89–98.

21 Krol B, Sanderman R, Suurmeijer TP. Social support, rheumatoid arthritis and quality of life: concepts, measurement and research. Patient Educ Couns 1993;20:101–20.

22 Leyland KM, Hart DJ, Javaid MK *et al.* The natural history of radiographic knee osteoarthritis: a fourteen-year population-based cohort study. Arthritis Rheum 2012;64:2243–51.

23 Ling SM, Xue QL, Simonsick EM *et al.* Transitions to mobility difficulty associated with lower extremity osteoarthritis in high functioning older women: longitudinal data from the Women's Health and Aging Study II. Arthritis Care Res 2006;55:256–63.

24 Miller ME, Rejeski WJ, Messier SP, Loeser RF. Modifiers of change in physical functioning in older adults with knee pain: the Observational Arthritis Study in Seniors (OASIS). Arthritis Care Res 2001;45:331–9.

25 Peeters GMEE, Alshurafa M, Schaap L, de Vet HCW. Diagnostic accuracy of self-reported arthritis in the general adult population is acceptable. J Clin Epidemiol 2015;68:452–9.

26 Parsons C, Clynes M, Syddall H, the EPOSA research group *et al.* How well do radiographic, clinical and self-reported diagnoses of knee osteoarthritis agree? Findings from the Hertfordshire cohort study. Springerplus 2015;4:177.

27 March LM, Schwarz JM, Carfrae BH, Bagge E. Clinical validation of self-reported osteoarthritis. Osteoarthritis Cartilage 1998;6:87–93.

Supplement 5.1

The equation of LMM consists of two levels.

Level 1 Equation

$Y_{ij} = \pi_{0i} + \pi_{1ij}TIMEPOINT_{ij} + \varepsilon_{ij}$

Level 2 Equation

$\pi_{0i} = \beta_{00} + \beta_{01}OAdiagnose + \beta_{02}Age + \beta_{03}Sex + \beta_{04}BMI + \beta_{05}Education + \beta_{06}Employment + \beta_{07}Comorbidity + \gamma_{0i}$

$\pi_{1ij} = \beta_{10j} + \beta_{11j}OAdiagnose + \gamma_{1ij}$

Note: Y_{ij} denotes individual i's HRQoL score at time j where j = 1, 2, 3, or 4 represents baseline, 2.5 years, 5 years and 10 years, respectively.

At level 1, only time-varying covariate (TIMEPOINT) is included, and it indicates each individual's growth trajectory of outcome measure (π_{1i}) by releasing its random effect in level 2 (γ_{1i}). On the other hand, each individual's response at baseline (π_{0i}) is allowed to differ by releasing its random effect in level 2 (γ_{0i}). At level 2, the regression coefficients (β_{01}–β_{07}) represents that each participants' initial status (intercept) will be associated with their covariates. Importantly, the regression coefficients (β_{11j}) indicate the interaction terms of time by the OA diagnose which represents that the impact of time on outcome (π_{1i}) will be associated with patients' OA diagnose. That is, the changes in HRQoL (time effect) are significantly different for a given subgroup when a significant interaction term appears.

As listed in the above equations, we adjusted for fixed effects for timepoint, OA diagnose, age, sex, BMI, education, employment, and comorbidity counts, and two-way interaction terms of time by OA diagnose as well as random effects for the intercept (baseline score) and time effect (linear slope). The random effect for the intercept was estimated, indicating that the baseline score changed across subjects. The random effect for the time effect was also estimated, indicating that each subject has his/her own growth trajectory. In terms of the random-effects covariance–variance matrix, we assumed an unstructured covariance with no specific form. In contrast to the traditional statistical methods, LMM thus allows that the baseline score and the time effect to change across subjects.

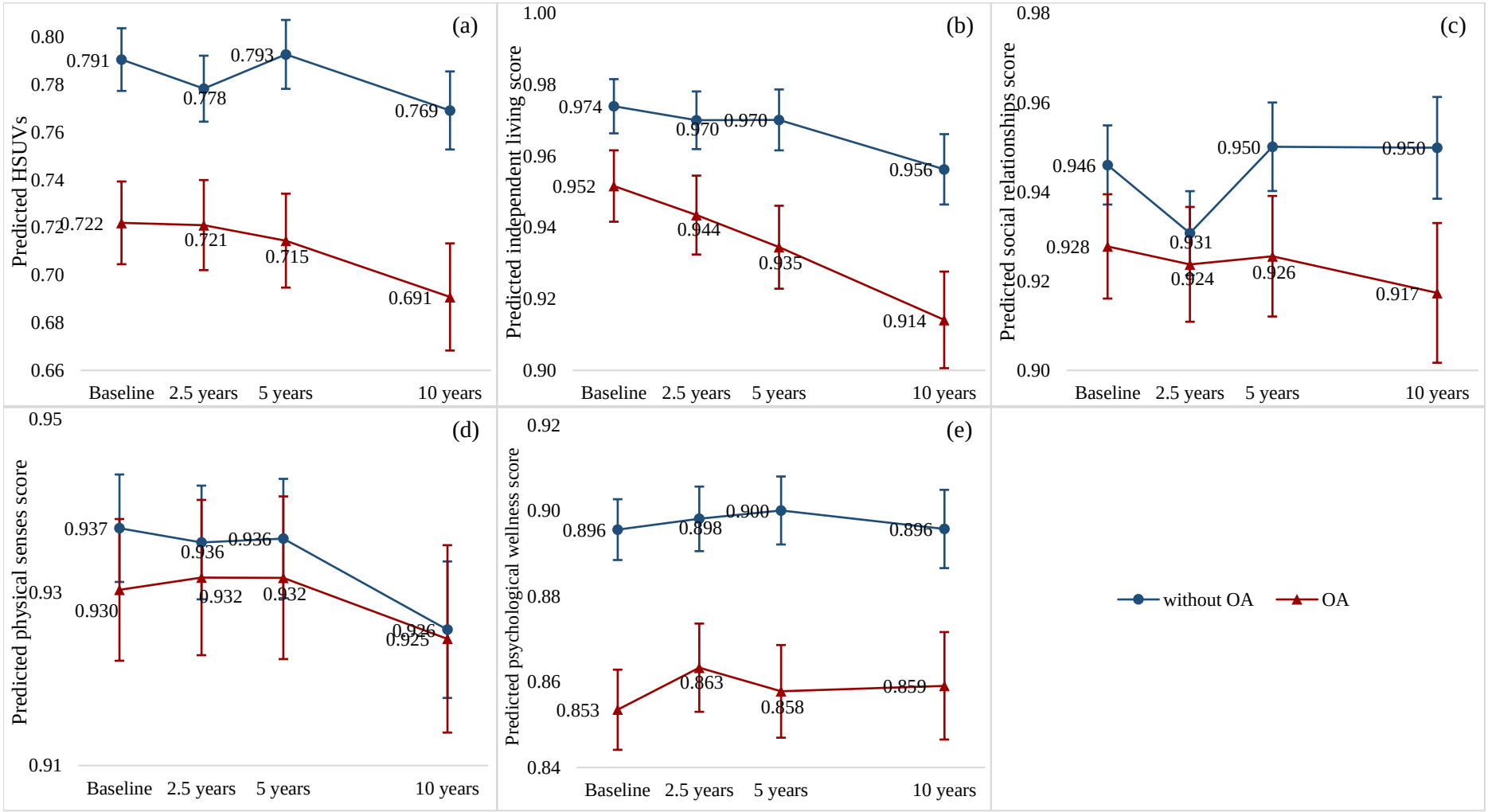
Supplement 5.2

The baseline characteristics of participants who were lost to follow-up at 10-years compared to those who completed the 10-year follow-up

	Lost to follow-up		Completing Follow-up	
Variables	n	%	n	%
Sex				
Male	256	47.9	280	49.6
Female	278	52.1	284	50.4
Age				
50-59	193	36.1	279	49.5
60-69	180	33.7	219	38.8
70-80	161	30.1	66	11.7
BMI				
Normal	151	28.3	167	29.6
Overweight	219	41	259	45.9
Obesity	164	30.7	138	24.5
Education				
No	108	20.2	72	12.8
Middle	372	69.7	421	74.8
High	54	10.1	70	12.4
OA diagnosis				
No OA	325	61.1	372	66.1
Knee and/or hip OA	121	22.7	82	14.6
Other types of OA	86	16.2	109	19.4

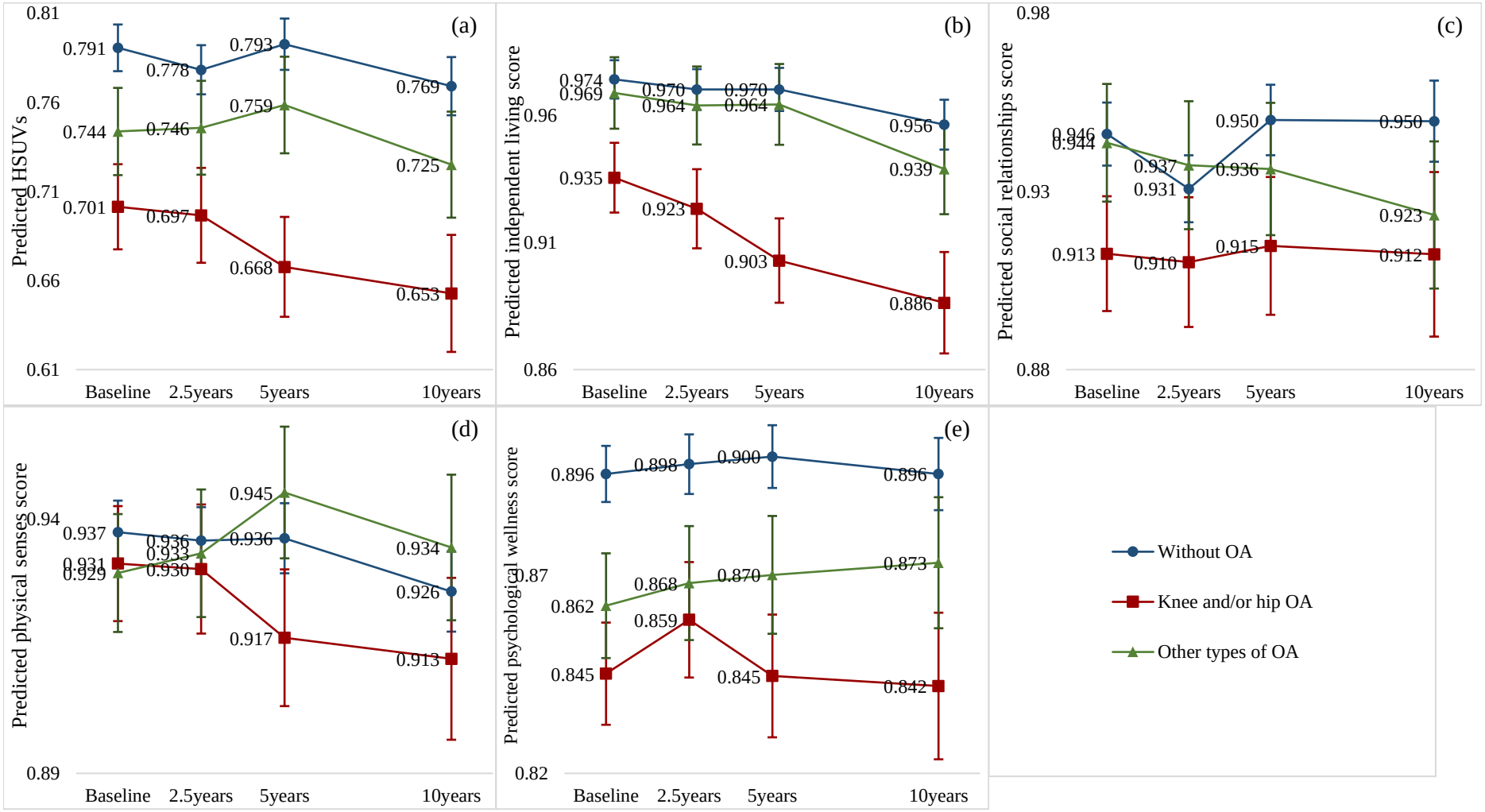
Note: OA=osteoarthritis, BMI=body mass index, other types of OA include OA at the neck, back, hands, shoulders and feet.

Supplement 5.3



The predicted HRQoL with 95% confidence interval over time for participants with and without OA. (a) predicted HSUVs, (b) predicted independent living score, (c) predicted social relationships score, (d) predicted physical senses score, (e) predicted psychological wellness score. OA=osteoarthritis. HRQoL=health-related quality of life. HSUVs=health state utility values.

Supplement 5.4



The predicted HRQoL over time for No OA and each OA type. (a) predicted HSUVs, (b) predicted independent living score, (c) predicted social relationships score, (d) predicted physical senses score, (e) predicted psychological wellness score. OA=osteoarthritis. HRQoL=health-related quality of life. HSUVs=health state utility values.

Supplement 5.5 The rational of using self-reported OA diagnosis and the potential impact of OA incidence on results

The self-reported OA is not a perfect measure of OA's prevalence and, hence, may impact study results due to the associated misclassification bias. However, a strong agreement between the self-reported OA and clinical OA measures (81%) have been reported in the past, so we don't expect a drastically huge HRQoL impacts due to misclassification bias in our study. The key reason behind using self-reported OA diagnosis in this study was the study's focus on OA of all joint sites with the aim of estimating the overall HRQoL impacts of OA for the cohort. Whilst TASOAC collected radiographic and pain symptom data, those were only collected for a select few individual joint sites (i.e., knee and hip) and the timing of collecting these data varied by affected OA joint sites (leading to further complexities in relation to using these data to construct uncertainty intervals to quantify the sensitivity of the results to the uncertainties facing exposure definition). Notably, as the TASOAC database already had low number of participants completing all four waves of the survey, relying on clinical data would substantially reduce the sample sizes further, so no meaningful conclusions could be drawn. Table 5.5.1 below provides further specific details about these data.

As shown in Table 5.5.1, the pain WOMAC score for knee and hip and X-ray imaging for hand, knee and hip were collected at some of the timepoints, and only for subgroups of the study participants. Previous literature on the effect of OA definitions on prevalence and incidence estimates has suggested that OA incidence based on alternative definitions can be seriously misleading and is very unlikely to provide valid estimates of the incidence rate (REF). Therefore, to estimate OA incidence using self-reported OA diagnosis at baseline and other OA definitions at follow-up was not conducted.

Table 5.5.1 Summary of OA symptoms and radiographic data availability in the TASOAC

Measure	Baseline	2.5-years	5-years	10 years
Knee pain-WOMAC scale	√	√	√	√
Hip pain-WOMAC scale		√	√	
X-ray				
hand	√			√
hip	√ (n=939 left) (n=1009 right)			√ (n=500 left & right)

knee	√ (n=1017 left) (n=1019 right)	√ (n=392 right only)	√ (n=500 left & right)

In the TASOAC study, self-reported OA was collected at baseline for n=1,093 and 2.5-years follow-up for n=873 (Table 5.5.2). To provide an indication of the impact of OA's incidence after baseline on our results, we have used these data to get some indication of the impact of unmeasured confounders (i.e., the incidence of OA) on our outcome (i.e., HSUVs).

Table 5.5.2 Self-reported OA diagnosis at baseline and 2.5-year follow-up

	OA	No OA	Total
Baseline	695 (64%)	398 (36%)	1,093
2.5 years	529 (61%)	344 (39%)	873

As a chronic, non-curable disease, it is sensible to assume that people who reported having OA at a previous timepoint will have OA in the following timepoints. However, among those who's self-reported OA data were available at both the baseline and 2.5-years follow-up (n=869), n=72 (8%) participants with self-reported OA at baseline reported having no OA at 2.5-years follow-up (which is unlikely to be correct). Furthermore, n=111 (13%) participants who reported no OA at baseline reported having OA at 2.5-years follow-up (Table 5.5.3).

Table 5.5.3 TASOAC participants' OA diagnosis self-reports at baseline and 2.5 years follow-up

		2.5-years		
		No OA	OA	Total
Baseline	No OA	454	111	565
	OA	72	232	304
	Total	526	343	869

Note: OA=osteoarthritis.

Based on these data, the annual self-reported OA incidence for TASOAC cohort is estimated at 7.9% (=111÷565÷2.5). We also compared the HSUVs of baseline and 2.5-years follow-up for patients' sub-groups of Table 5.5.3 and found that for those who reported having no OA at baseline and 2.5-years follow-up (n=454), the HSUVs reduced 0.02 units from baseline to 2.5 years later, for those who reported no OA at baseline and having OA at 2.5-years follow-up, HSUVs reduced 0.03 units during the same period (Table 5.5.4). As the self-reported OA diagnosis was only collected at baseline and 2.5-years follow-up, we were not able to calculate how the incidence of self-reported OA impacted the exposure and outcome of Study 3 for the full study duration of 10 years. From the results of our additional investigation, we can say that Study 3 has likely underestimated the HRQoL impact of OA.

Table 5.5.4 HSUVs of TASOAC participants based on their self-reported OA at baseline and 2.5 years follow-up

	Variable	n	Mean	SD	Difference
'No OA' to 'No OA'					
	HSUV at baseline	454	0.812	0.16	0.02
	HSUV at 2.5 years	452	0.794	0.18	
'No OA' to 'OA'					
	HSUV at baseline	111	0.782	0.15	0.03
	HSUV at 2.5 years	111	0.756	0.17	
'OA' to 'No OA'					
	HSUV at baseline	72	0.755	0.20	-0.02
	HSUV at 2.5 years	72	0.770	0.18	
'OA' to 'OA'					
	HSUV at baseline	232	0.716	0.19	0.02
	HSUV at 2.5 years	231	0.700	0.21	

Note: OA=osteoarthritis, HSUV=health-state utility value, SD=standard deviation.

Chapter 6: The impact of comorbidities on health-related quality of life of people with osteoarthritis over ten years

6.1 Preface

Chapter 6 presents Study 4 which comprehensively investigates the impact of comorbidity count and comorbidity patterns on long-term changes in HSUVs and individual health-dimension scores of OA people, using Australian data and Assessment of Quality of Life (AQoL)-4D instrument. An important aim of this chapter is to guide the adjustment of HSUVs inputs for comorbidity numbers and patterns in the future OA health economic models of Australians and similar populations with alternative comorbidity profiles.

The text in Chapter 6 has been published in *Rheumatology* (Zhao, Ting, Tania Winzenberg, Dawn Aitken, Barbara de Graaff, Hasnat Ahmad, Graeme Jones, and Andrew J. Palmer. "The impact of comorbidities on health-related quality of life of people with osteoarthritis", Rheumatology (2021) (Supplement 5A).

6.2 Abstract

Objective: To investigate the impact of total number and patterns of comorbidities on health-related quality of life (HRQoL) and identify the most prevalent and influential comorbidity patterns in people with osteoarthritis (OA) over ten years.

Methods: Participants from the Tasmanian Older Adult Cohort aged 50-80 years, with self-reported OA and data on comorbidities and HRQoL were included. Participants were interviewed at baseline (n=398), 2.5-years (n=304), 5-years (n=269) and 10-years (n=191). Data on the self-reported presence of 10 chronic comorbidities were collected at baseline. HRQoL was assessed using the Assessment of Quality of Life-4-Dimensions. The long-term impacts of the number and of the nine most prevalent combinations of cardiovascular (CVD), non-OA musculoskeletal (Ms), metabolic, and respiratory comorbidities on HRQoL over ten years were analysed using linear mixed regressions.

Results: Compared with comorbidity-free OA participants, the health state utility (HSU) of those with 2 or ≥3 comorbidities were respectively -0.07 and -0.13 units lower over ten years, largely driven by reduced scores for independent living, social relationships and psychological

wellness. Comorbidity patterns including 'CVD+Ms' were most influential, and associated with up to 0.13 units lower HSU, mostly through negative impacts on independent living (up to -0.12), psychological wellness (up to -0.08) and social relationship (up to -0.06).

Conclusion: Having more comorbidities negatively impacted OA patients' long-term HRQoL. OA patients with CVD and non-OA musculoskeletal conditions had the largest HSU impairment, therefore optimal management and prevention of these conditions may yield improvements in OA patients' HRQoL.

6.3 Introduction

Osteoarthritis (OA) is a chronic and progressive condition that leads to substantial socioeconomic burden and low individual health-related quality of life (HRQoL) (1, 2). People with OA are more likely to have comorbidities and lower HRQoL than people without OA, with a systematic review reporting a prevalence of any comorbidity of 67% in people with OA compared with 56% in those without (3). Multi-attribute utility instruments (MAUIs) such as the Assessment of Quality of life-4-Dimensions (AQoL-4D) can be used to measure HRQoL. Health dimension scores can be combined according to the AQoL-4D algorithm to generate a weighted index of the health state utility (HSU)─measuring the strength of an individual or society's preference for a given health state (4).

Comorbidities have a negative impact on HRQol of people with OA, with most previous studies using the total number of comorbidities (comorbidity count) as the comorbidity measure (5-8). Whilst these studies have improved our understanding of the potentially additive impact, their reliance on an implicit and rather unrealistic assumption of a uniform HRQoL impact of all comorbidities (both individually as well as in combination with each other) is a limitation. In reality, comorbidities vary considerably between studies due to differences in data sources and study populations (5-8). Moreover, the impact of specific conditions or combinations of conditions may be caused by each condition(s)' specific pathophysiology, symptoms, and treatments that may involve synergistic effects rather than additive effects on HRQoL (9-11). Thus, it is also important to consider the HRQoL impacts of different patterns (i.e. types and combinations) of comorbidities in OA to improve on the existing count method.

Additionally, the long-term effects of comorbidities on HRQoL of people with OA are not well researched. To date, only a single study has reported that comorbidities are associated with increased long-term disability in people with OA (12), but none have investigated effects on

HSUs and individual health dimensions. Knowing how comorbidities impact OA patients over the longer term could help guide decision-making of clinicians and people with OA with regards to prioritising management choices for comorbidities to optimise HRQoL.

Our study aims to fill these major evidence gaps by investigating the impact of total number and patterns of comorbidities on HRQoL (in terms of HSUs and health dimension scores) in people with OA over ten years. It will also identify the most prevalent and influential comorbidity pattern in OA patients.

6.4 Methods

6.4.1 Study Design

Participants came from the Tasmanian Older Adult Cohort (TASOAC), a prospective, population-based longitudinal study aimed at identifying the environmental, genetic and biochemical factors associated with the development and progression of OA at multiple joint sites (e.g.: knee, hip, hand, and spine). Participants aged 50–80 years in 2002 were selected from the electoral roll in Southern Tasmania, using sex-stratified simple random sampling without replacement. Participants were excluded if they resided in an aged care facility or were unable to have a knee MRI scan. The research was approved by the Southern Tasmanian Health and Medical Human Research Ethics Committee and written informed consent to participate in the study was obtained from all participants. Only those who reported having doctor-diagnosed OA were included in our study. Data were collected using interview-administered questionnaires at 4 periods: baseline (from February 2002 to September 2004, n=398), 2.5 years later (range 1.4−4.8 years, n=304), 5 years later (range 3.6−6.9 years, n=269), and 10 years later (range 9−13 years, n=191).

6.4.2 Diagnosis of Osteoarthritis

At baseline, participants were asked “Have you had been told by a doctor that you have osteoarthritis at any of these sites”. Seven sites were listed: 1) neck, 2) back, 3) hands, 4) shoulders, 5) hips, 6) knees, and 7) feet. Participants chose "yes" or "no" for each site. Participants answering “yes” to at least one site were considered as diagnosed with OA and were included in our study. We categorised participants as having knee and/or hip OA if at least one of the sites of knee and hip (individually or in combination with other joint sites) were indicated, and as having other types of OA if at least one of the five joint sites (other than knee and hip) were indicated.

6.4.3 Comorbidity counts and patterns

Data on comorbidities were collected at baseline by asking “Have you ever been diagnosed by a doctor as having any of the following”. Ten chronic conditions were listed: 1) diabetes, 2) heart attack (a coronary, coronary occlusion, coronary thrombosis, myocardial infarction), 3) hypertension (high blood pressure), 4) thrombosis, 5) asthma, 6) bronchitis/emphysema, 7) osteoporosis, 8) hyperthyroidism, 9) hypothyroidism, and 10) rheumatoid arthritis. Participants were given the choice between answering "yes" or "no" to each comorbid condition. Those answering “yes” were considered as having that comorbidity. The total number of comorbidities of each participant were calculated by summing up the “yes” answers for each comorbidity and were classified into 4 groups: 0, 1, 2, and ≥3.

Based on the type of individual comorbid conditions, patients were then categorised into four broad comorbidity groups: 1) cardiovascular disease (CVD) group which included heart attack, hypertension, thrombosis; 2) non-OA musculoskeletal disease (Ms) group, including osteoporosis and rheumatoid arthritis; 3) metabolic disease (Me) group, including hyperthyroidism, hypothyroidism and diabetes; and 4) respiratory disease (Re) group, including asthma and bronchitis/emphysema.

Sixteen potential comorbidity patterns were identified based on the absence or presence of the above-mentioned four comorbidity types individually or in combination with each other (i.e.: comorbidity-free OA participants [control group]; OA participants with one of the four [CVD, Ms, Me, or Re] comorbidity types—4 possible patterns; OA participants with any two comorbidity types—6 patterns; OA participants with any three comorbidity types—4 patterns; and OA participants with all four comorbidity types—1 pattern).

Measurement of Health-related Quality of life

HRQoL was assessed at baseline, 2.5, 5 and 10 years using the well validated Assessment of Quality of life 4-Dimensions (AQoL-4D) questionnaire, with utility weights derived from an Australian population sample (13). AQoL-4D consists of 12 items covering four dimensions. The four dimensions and their corresponding items are: 1) independent living (self-care, activities of daily living, and mobility), 2) physical senses (sight, hearing and communication), 3) social relationships (social isolation, relationship, and family role) and 4) psychological wellbeing (sleep, anxiety, and pain). The scores for items in each dimension were transformed and calculated to provide dimension scores and an overall index of HSU using the AQoL-4D

algorithm (Supplement 6.1) (4). HSU and dimensional scores range from 0.00 to 1.00, with higher score indicating better HRQoL.

6.4.4 Other Characteristics

Sex and date of birth were collected at baseline; and weight and height were collected at all periods. Weight was measured to the nearest 0.1 kg (with no shoes/socks/bulky clothing/headwear) using a single pair of calibrated electronic scales (Seca Delta Model 707). Height was measured to the nearest 0.1 cm (barefooted) using a stadiometer. Body mass index (BMI) was calculated as [weight (kg)/ (height $(m)^2$], and participants were grouped into three categories of BMI (normal weight [BMI $<25.0 kg/m^2$], overweight [BMI$\geq$25.0kg/$m^2\leq$29.9kg/m^2] and obese [BMI$\geq$30.0kg/m^2]) following the WHO criteria (14). Data on disability severity was collected using the Health Assessment Questionnaire Disability Index (HAQ-DI) (15) at all periods. This has eight sections (dressing, arising, eating, walking, hygiene, reach, grip, and activities) with 20 items. The HAQ-DI score (ranging between 0 and 3) was calculated and categorized into two groups: no to mild disability (HAQ-DI levels 0-1) and moderate to very severe disability (HAQ-DI levels >1) (15).

6.4.5 Statistical analysis

We summarised the demographic and other features of our sample using descriptive statistics. Linear mixed regressions were conducted to estimate the impact of baseline comorbidity status (i.e.: total number of comorbidities and comorbidity patterns) on HRQoL (in terms of HSUs and individual health-dimension scores) of people with OA over ten years, with adjustment for age (50-59=1, 60-69=2, 70-79=3, ≥80=4), sex (male=1, female=2) and time of follow-up. In order to produce estimates with a meaningful level of precision, comorbidity patterns with less than 10 participants were excluded from the regression analysis so a total number of 9 most prevalent comorbidity patterns that were assessed for their influence on HRQoL. Subgroup analyses were conducted to evaluate whether the HRQoL impacts of (the number and patterns of) comorbidities vary between patients with knee and/or hip OA and other types of OA. The difference in HSUs of participants with different numbers and patterns of comorbidity were evaluated based on both statistical and clinical significance. Statistical significance was set as a P-value ≤0.05 (two-tailed) and results were considered clinically significant if they met or exceeded the minimum clinically important difference (MID) in the AQoL-4D HSUs of 0.06 (95% confident interval [CI]: 0.03−0.08) units for the Australian population (16). Regression estimates and 95% confidence interval (CI) of the effects of comorbidity status on HRQoL

were reported, which represent the average difference in HRQoL (HSU and dimensional) scores of each comorbidity category of OA patients compared with the reference (no comorbidity) group over ten years. All analyses were conducted in STATA (STATA 15.1, StataCorp, College Station, Texas, USA).

6.5 Results

398 participants with OA were included in our analyses (Table 6.1). The mean age (standard deviation, SD) at baseline was 64.0 (7.6) years. From Table 6.1, 39.4% (n=157) of participants were female, 69.8% (n=278) were overweight or obese, and 51.0% (n=203) had knee and/or hip OA. The median (interquartile range) number of comorbidities for OA participants at baseline was 1 (1-2). Comorbidity counts ranged between 0 and 7, with 24.6% (n=98) participants having no comorbidities at baseline, and 33.4% (n=133), 20.9% (n=83) and 21.1% (n=84) having one, two, and three or more comorbidities, respectively. The mean (SD) HSU at baseline was 0.71 (0.20). The mean HSUs at each timepoint and by comorbidity status (i.e.: total number of comorbidities and comorbidity patterns) are shown in Supplement Figures 6.1-6.3.

Table 6.1. The demographic and other characteristics of participants at baseline (N=398)

Variable	OA participants (N=398)
Sex, n (%)	
Female	157 (39.4)
Male	241 (60.6)
Age (years), n (%)	
50-60	139 (34.9)
60-70	154 (38.7)
70-80	105 (26.4)
BMI, n (%)	
Normal (BMI<25.0kg/m^2)	120 (30.2)
Overweight (25.0kg/m^2≤ BMI ≤29.9kg/m^2)	156 (39.2)
Obese (BMI≥30.0kg/m^2)	122 (30.7)
HAQ-DI, n (%)	
No to mild (HAQ-DI 0-1)	349 (87.7)
Moderate to very severe (HAQ-DI >1)	49 (12.3)
OA affected join sites*, n (%)	
Knee and/or hip OA	203 (51.0)
Other types of OA	195 (49.0)
Number of comorbidities[#], n (%)	
0	98 (24.6)
1	133 (33.4)
2	83 (20.9)
≥3	84 (21.1)

AQoL-4D's HSUs, mean (SD)	0.71 (0.20)
AQoL-4D's independent living score, mean (SD)	0.94 (0.13)
AQoL-4D's social relationships score, mean (SD)	0.92 (0.13)
AQoL-4D's physical senses score, mean (SD)	0.93 (0.09)
AQoL-4D's psychological wellness score, mean (SD)	0.85 (0.10)

Note: *OA affected joint sites were based on self-reported OA diagnosis, Other types of OA included OA at joint sites of hand, shoulder, feet, neck and back. #Number of comorbidities was based on the self-reported comorbidities. OA=osteoarthritis, BMI=body mass index, AQoL-4D=Assessment of Quality of Life-4-Dimensions, HSUs=health state utilities, SD=standard deviation, HAQ-DI=health assessment questionnaire-disability index.

As shown in Table 6.2, 16 comorbidity patterns were identified at baseline, nine of which were seen in 10 or more participants. Among those with comorbidities, the most frequent comorbidity patterns were CVD (20.9%), 'CVD+Ms' (10.6%) and Ms (9.5%).

Table 6.2. Number of participants by comorbidity patterns at baseline (N=398)

Comorbidities type #	**n (%)**
No comorbidities	**98 (24.6)**
Cardiovascular Disease (CVD) only	**83 (20.9)**
Non-OA musculoskeletal Disease (Ms) only	**38 (9.5)**
Metabolic (Me) only	7 (1.8)
Respiratory Difficulties (Re) only	**25 (6.3)**
CVD+Ms	**42 (10.6)**
CVD+Me	8 (2.0)
CVD+Re	**23 (5.8)**
Ms+Me	3 (0.8)
Ms+Re	**12 (3.0)**
Me+Re	6 (1.5)
CVD+Ms+Me	**13 (3.3)**
CVD+Ms+Re	**23 (5.8)**
CVD+Me+Re	5 (1.3)
Ms+Me+Re	3 (0.8)
CVD+Ms+Me+Re	9 (2.3)

#N=16 comorbidities combinations were found.
Note: Cardiovascular diseases (CVD) included Heart attack, Hypertension and Thrombosis; Non-OA musculoskeletal diseases (Ms) included Osteoporosis and Rheumatoid arthritis; Metabolic diseases (Me) included Diabetes, Hyperthyroidism and Hypothyroidism; Respiratory diseases (Re) included Asthma and Emphysema. Comorbidity patterns with more than 10 participants are given in bold.

Table 6.3 shows the impacts of total number of comorbidities on HSU and health dimension scores over ten years. The adjusted difference in HSU between participants with zero and one comorbidity was neither statistically significant nor clinically important over ten years. However, participants with two comorbidities recorded clinically meaningful and statistically

significant reductions in mean HSU (-0.07 [95% CI: -0.12, -0.02]) over ten years compared to comorbidity-free participants. Participants with two comorbidities had significantly lower scores on independent living (-0.06 [95% CI: -0.09, -0.03] units), social relationships (-0.03 [95% CI: -0.06, 0.00] units), and psychological wellness (-0.03 [95% CI: -0.05, 0.00] units) compared to those without comorbidity.

Participants with three or more comorbidities recorded the largest -0.13 (95% CI: -0.19, -0.08) units reduction in mean HSU over ten years compared to those without comorbidity. This reduction is both clinically important and statistically significant. All four health dimensions were also substantially impacted in this group. Results were similar for the two sub-groups of affected joint sites (i.e.: knee and/or hip joints; and other joints) (Supplement 6.2).

Table 6.4 shows the impacts of comorbidity patterns on HSU and health dimension scores over ten years. Having any single comorbidity was not associated with statistically significant reduction in HSUs compared to having none. Of the combinations of two comorbidities, 'CVD+Ms' was associated with the largest reductions in HSUs (-0.13 [95% CI: -0.20, -0.07]), driven by the impacts on independent living (-0.08 -[0.12, -0.04]), social relationships (-0.06 [-0.10, -0.03]) and psychological wellness (-0.04 -[0.08, -0.01]). The impacts on independent living and psychological wellness became greater when Me or Re was combined with 'CVD+Ms'. Results were similar for the two sub-groups of affected joint sites (i.e.: knee and/or hip joints; and other joints) (Supplement 6.3).

Table 6.3. The impacts of total number of comorbidities on OA participants' HSUs and health dimension scores over ten years

Number of comorbidities	HSUs	Independent living	Social relationships	Physical senses	Psychological wellness
0 (n=98)	Reference	Reference	Reference	Reference	Reference
1 (n=133)	-0.03 (-0.07, 0.02)	-0.02 (-0.05, 0.01)	-0.01 (-0.04, 0.01)	0.00 (-0.02, 0.02)	-0.01 (-0.04, 0.01)
2 (n=83)	**-0.07 (-0.12, -0.02)**	**-0.06 (-0.09, -0.03)**	**-0.03 (-0.06, 0.00)**	0.00 (-0.02, 0.02)	**-0.03 (-0.05, 0.00)**
≥3 (n=84)	**-0.13 (-0.19, -0.08)**	**-0.10 (-0.13, -0.06)**	**-0.04 (-0.07, -0.01)**	**-0.03 (-0.05, 0.00)**	**-0.06 (-0.09, -0.03)**

Note: Regressions were adjusted by age, sex, and follow-up time. Data were presented as mean difference (95% confidence interval). OA=osteoarthritis, HSUs=health state utilities. The reported beta coefficients of each category of comorbidity status represent average difference in HRQoL (HSU and dimensional) scores of each OA subgroup compared with the reference (no comorbidity) group over ten years. No mathematical correction was made for multiple comparisons

Table 6.4. The impacts of comorbidity patterns on OA participants' HSUs and health dimension scores over ten years

Comorbidity patterns	HSUs	Independent living	Social relationships	Physical senses	Psychological wellness
No comorbidity (n=98)	Reference	Reference	Reference	Reference	Reference
CVD only (n=83)	-0.04 (-0.09, 0.01)	-0.03 (-0.06, 0.00)	-0.02 (-0.05, 0.01)	0.00 (-0.02, 0.02)	-0.02 (-0.04, 0.01)
Ms only (n=38)	-0.05 (-0.12, 0.01)	-0.02 (-0.06, 0.02)	**-0.04 (-0.08, 0.00)**	0.01 (-0.01, 0.04)	-0.02 (-0.06, 0.01)
Re only (n=25)	0.01 (-0.07, 0.08)	-0.01 (-0.06, 0.04)	0.02 (-0.02, 0.06)	0.00 (-0.03, 0.03)	0.00 (-0.04, 0.04)
CVD+Ms (n=42)	**-0.13 (-0.20, -0.07)**	**-0.08 (-0.12, -0.04)**	**-0.06 (-0.10, -0.03)**	-0.02 (-0.05, 0.01)	**-0.04 (-0.08, -0.01)**
CVD+Re (n=23)	-0.07 (-0.15, 0.01)	**-0.08 (-0.13, -0.03)**	-0.01 (-0.06, 0.04)	0.01 (-0.03, 0.04)	-0.03 (-0.08, 0.01)
Ms+Re (n=12)	-0.04 (-0.15, 0.07)	-0.05 (-0.12, 0.02)	-0.01 (-0.07, 0.06)	-0.03 (-0.07, 0.02)	0.01 (-0.05, 0.06)
CVD+Ms+Me (n=13)	**-0.13 (-0.23, -0.02)**	**-0.12 (-0.19, -0.06)**	-0.03 (-0.09, 0.04)	0.00 (-0.04, 0.04)	**-0.07 (-0.13, -0.02)**
CVD+Ms+Re (n=23)	**-0.13 (-0.21, -0.05)**	**-0.10 (-0.16, -0.05)**	-0.03 (-0.07, 0.02)	0.00 (-0.04, 0.03)	**-0.08 (-0.12, -0.04)**

Note: Regression models were adjusted for age, sex, and follow-up time. Data were presented as mean difference (95% confidence interval). Combination groups with less than 10 respondents at baseline were excluded from modelling, leaving 9 combination groups in the regressions. HSUs= health state utilities; OA=osteoarthritis; CVD=cardiovascular disease (including Heart attack, Hypertension and Thrombosis); Ms=non-OA musculoskeletal disease (including Osteoporosis and Rheumatoid arthritis); Me=metabolic disease (including Diabetes, Hyperthyroidism and Hypothyroidism); Re=respiratory disease including (Asthma and Emphysema). The reported beta coefficients of each category of comorbidity status represent average difference in HRQoL (HSU and dimensional) scores of each OA subgroup compared with the reference (no comorbidity) group over ten years. No mathematical correction was made for multiple comparisons.

6.6 Discussion

This study fills a major evidence gap by investigating the associations between the total number and patterns of comorbidities and OA-related HRQoL over ten years using a population-based cohort and identifying the most prevalent and influential comorbidity patterns in an OA population. Comorbidities were strongly associated with lower HRQoL in people with OA - compared with comorbidity-free OA participants, people with ≥ 2 comorbidities experienced a clinically and statistically significant reduction of up to 0.13 HSU units over ten years, largely driven by reduced scores on independent living (up to -0.10) and psychological wellness (up to -0.06). The commonest comorbidity combination was CVD and Ms conditions which was associated with largest HSU impairment, driven by negative impacts on independent living, social relationships and psychological wellness scores. These results highlight that cardiovascular and non-OA musculoskeletal conditions might be comorbidities that yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention. Our findings also emphasize the importance of adjusting for the number as well as patterns of comorbidities when assessing HRQoL in OA patients.

Comorbidities were common in our study, with CVD and MS being the most prevalent comorbid conditions. The mean age (64 years) and the proportion of participants affected by OA-related comorbidities (75%) in our sample was similar to that reported in a recently published systematic review (3). The distributions of comorbidity count were also similar between the two studies, with 33%, 21% and 21% of our participants having 1, 2 and ≥3 comorbidities, respectively (compared with 29%, 25% and 24% in the review). The most prevalent co-morbid condition in our participants was CVD (20.9%), and the most prevalent comorbidity combination was ‘CVD+Ms’ (10.6%). Once again, our results were consistent with previous studies that have identified the cardiovascular and non-OA musculoskeletal diseases as the two most commonly observed comorbid conditions in people with OA. This suggests that our findings can be generalised to OA populations in other geographic locations with enough degree of confidence.

Participants with two or more comorbidities recorded clinically meaningful and statistically significant reductions in mean HSUs over ten years compared to comorbidity-free participants. Whilst the cross-sectional HSU difference between people with OA with and without comorbidities (17), and the impact of the number of comorbidities on HSU of OA patients (5) have previously been investigated, our study provides the first data of which we are aware

describing impacts of OA-related comorbidities on long-term changes in HSU. As such, our findings can be instrumental in generating insights into the future health trajectories of OA patients with different comorbidity profiles.

Our study bridges another important evidence gap by assessing the nature and extent of the effects of various comorbidity patterns on the HRQoL in OA population, highlighting the importance of adjusting for the number as well as types and combinations of comorbidities. The investigation of combined effects is particularly relevant for people with OA who often have multiple comorbid conditions (3). We found the combination of 'CVD+Ms' was the most dominant predictor of HRQoL impairment, exhibiting additive as well as synergistic effects (as indicated by the -0.13 unit reduction in HSUs in this group, compared to -0.04 in the CVD only and -0.05 in the Ms only). The combination of 'CVD+Ms' also strongly impacted individual health dimensions with the resultant largest reduction in independent living scores, followed by social relationships and psychological wellness scores. The impacts on independent living and psychological wellness became greater when Me or Re were combined with 'CVD+Ms'. For OA with Ms only and with 'CVD+Re', the impairment was seen in social relationships and independent living, respectively. Our results show that as well as the number of comorbidities, the types and combinations of comorbidities vary in effect sizes and health dimensions influenced, and suggest that the CVD and Ms comorbid conditions might yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention. The underlying mechanisms behind the various impacts of comorbidity patterns on HSUs and health dimensions may be different (18) but was not examined in our study and this remains unclear for OA populations (3). A more in-depth investigation of the prevalence of comorbidity patterns, their precise effects (as well as the mechanism through which these effects on various health dimensions are realised) should therefore be on agenda for the future research.

Linear mixed regression models are recommended when outcomes are measured repeatedly overtime and are likely to be correlated within each patient (19). The ability of linear mixed regression in incorporating fixed effects (the factors assumed to have the same effect across many patients) and random effects (the factors likely to vary substantially from patient to patient) and accommodating unbalanced data patterns even when the missing values are not completely at random (19), also make this method appropriate for our data. Our study reported separate estimates of the impact of different numbers and/or combinations of comorbidities on HSUs among people with OA, highlighting the differences identified as significant at the 5%

level. Although given the large number of comparisons thus reported, the associated significance levels was not adjusted for multiple comparisons in our study due to two key reasons: first, because within this kind of studies we are more interested in the clinical significance of our results rather than their statistical significance. Second, because the number of hypotheses (n=13) to test the impact of numbers and patterns of comorbidities on HSUs in our study is not immensely large, and the expected number of significant results just due to chance error (0.05 x 13=0.65 [or <1]) is negligibly small. This essentially means that out of 13 hypotheses that we have tested, only 5% (<1) comparisons may appear significant just due to chance, which does not seem to be a big problem. Finally, as the multiple testing burden correction methods tend to be overly conservative, they can easily lead to a high rate of false negatives, which may result in a situation that is equally bad or even more so.

Our study has some limitations. First, the number of participants in some comorbidity groups were quite small, and groups having <10 participants were not included in analysis. Only small number of individuals had individual comorbidities, so we also could not identify the relative contribution of individual comorbidities within each comorbidity group. Future studies with larger sample sizes are needed to 1) validate our findings, 2) provide more information about alternative comorbidity patterns which could not be examined in this study, and 3) investigate how comorbidities in each category compare with each other. A possible method to achieve these goals is to consider the main effect plus interaction specification (i.e.: indicator variables for the presence or absence of each comorbidity type, plus their first- and higher-order interaction) in the future models when sufficiently large sample become available, which may generate precise estimates of the impact of each comorbidity and would allow estimates to be made for all possible comorbidity combinations. Second, participants were selected based on self-reports of doctor-diagnosed OA and comorbidities were also self-reported without medical record verification. While high agreement has been found between self-reported comorbidity measures and those sourced from patients' medical records (20, 21), there remains the potential for misclassification. For example, the prevalence of rheumatoid arthritis in TASOAC baseline was 12% (22) which is higher than the expected population prevalence (1). Third, we lacked information on duration, severity, and treatment histories of OA and associated comorbidities, which may have important consequences for OA patients' HRQoL. Fourth, some patients may have developed comorbidities during the follow-up period, but our analysis did not consider this possibility and used only the baseline comorbidity data. The self-repored comorbidity data were collected at all four timepoints in TASOAC, however, the key

reason that we are not able to estimate and then account for the incidence of comorbidities during follow-up periods was the inconsistency facing these data over time (Supplement 6.4). As people with OA are more likely to develop comorbidities than non-OA people overtime (23, 24), our results may underestimate the long-term impacts of comorbidities on HRQoL of OA participants (Supplement 6.4). Finally, whilst our results on comorbidity prevalence were generally consistent with previous studies that have identified the cardiovascular and non-OA musculoskeletal diseases as the two most observed comorbid conditions in people with OA, some degree of sample selection/recruitment bias may be present as the prevalence of comorbidity in TASOAC participants may differ from that of the underlying OA population, for example, people with more severe comorbidity may be less inclined to participate in the study, which would result in under-estimation of the HRQoL decrements attributable to comorbidity. Furthermore, the fact that more than 50% (207/398) of our baseline sample was lost to follow-up by 10-year (final) review, has potential to introduce further uncertainties to our results. Due to the unavailability of demographic and clinical data from the overall population of Australians with OA, we are unable to know if our baseline sample was representative of the overall population of Australians with OA, and how the non-representativeness of TASOAC sample at baseline (if any) would impact the study conclusions. However, based on the comparison of HSUVs of TASOAC participants completing all four waves of the survey (n=0.76) with HSUVs of those lost to follow-up at 10 years (0.67) (Supplement 6.5), we can say that our study has likely underestimated the true impacts of comorbidities on HRQoL of Australians with OA.

Conclusion

Comorbidities were negatively associated with OA patients' HRQoL over ten years, and the effect sizes and health dimensions influenced varied by numbers and patterns of comorbidities. The findings of the long-term impacts of comorbidities generate important insights into the future health trajectories of OA patients with different comorbidity profiles. Having three or more comorbidities was associated with 0.13, 0.10, and 0.06 unit reductions in HSU, independent living, and psychological wellness scores, respectively. 'CVD+Ms' was the most prevalent and influential comorbidity combination in this OA population, and impacted HSU (-0.13 units), independent living (-0.08 units), social relationships (-0.06 units) and psychological wellness (-0.04 units). These results suggest that cardiovascular and non-OA musculoskeletal conditions might be the comorbidities that yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention.

Authors' contributions

GJ designed the TASOAC study that provided the data for our analyses. TZ, AP and HA conceived and designed the initial concept and method of this manuscript, DA, TW, and BG helped in modifying and improving the original concept and methods. TZ accomplished the setting up of database and performed necessary data analyses and drafted the manuscript. All authors were involved in the manuscript review and final approval.

Funding

No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this article.

Acknowledgements

We would like to thank all participants of TASOAC project (aimed at identifying the environmental, genetic, and biochemical factors associated with the development and progression of osteoarthritis at multiple sites) who provided the valuable data that informed the comprehensiveness of this paper. We would also like to thank all medical practitioners who were associated with TASOAC study. Finally, we thank to Mr Petr Otahal (Statistical Officer, Menzies Institute for Medical Research, University of Tasmania) for his help with the statistical analysis of this study.

Data availability statement

The data that was generated from this study will not be deposited in a public repository due to privacy and consent restrictions. De-identified data can be made available from the corresponding author on reasonable request, subject to a data sharing agreement.

6.7 References

1. Ackerman IN, Pratt C, Gorelik A, Liew D. Projected burden of osteoarthritis and rheumatoid arthritis in Australia: a population-level analysis. Arthritis care & research. 2018;70(6):877-83.

2. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging clinical and experimental research. 2005;17(4):255-63.

3. Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in Osteoarthritis: A systematic review and meta-analysis of observational studies. Arthritis care & research. 2019.

4. Hawthorne G, Richardson J, Day N. Using the Assessment of Quality of life (AQoL) version 1. Melbourne: Centre for Health Program Evaluation. 2001.

5. Hosseini K, Gaujoux-Viala C, Coste J, Pouchot J, Fautrel B, Rat A-C, et al. Impact of co-morbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis. Best Practice & Research Clinical Rheumatology. 2012;26(5):627-35.

6. Geryk LL, Carpenter DM, Blalock SJ, DeVellis RF, Jordan JM. The impact of co-morbidity on health-related quality of life in rheumatoid arthritis and osteoarthritis patients. Clin Exp Rheumatol. 2015;33(3):366-74.

7. Wesseling J, Welsing PM, Bierma-Zeinstra SM, Dekker J, Gorter KJ, Kloppenburg M, et al. Impact of self-reported comorbidity on physical and mental health status in early symptomatic osteoarthritis: the CHECK (Cohort Hip and Cohort Knee) study. Rheumatology. 2013;52(1):180-8.

8. van Dijk GM, Veenhof C, Schellevis F, Hulsmans H, Bakker JP, Arwert H, et al. Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. BMC musculoskeletal disorders. 2008;9(1):95.

9. Fortin M, Dubois M-F, Hudon C, Soubhi H, Almirall J. Multimorbidity and quality of life: a closer look. Health and Quality of Life Outcomes. 2007;5(1):52.

10. Hunger M, Thorand B, Schunk M, Döring A, Menn P, Peters A, et al. Multimorbidity and health-related quality of life in the older population: results from the German KORA-Age study. Health and Quality of Life Outcomes. 2011;9(1):53.

11. Mujica-Mota RE, Roberts M, Abel G, Elliott M, Lyratzopoulos G, Roland M, et al. Common patterns of morbidity and multi-morbidity and their impact on health-related quality of life: evidence from a national survey. Quality of Life Research. 2015;24(4):909-18.

12. Ettinger WH, Davis MA, Neuhaus JM, Mallon KP. Long-term physical functioning in persons with knee osteoarthritis from NHANES I: effects of comorbid medical conditions. Journal of clinical epidemiology. 1994;47(7):809-15.

13. Hawthorne G, Richardson J, Osborne R, McNeil H. The Australian quality of life (AQoL) instrument: initial validation: Centre for Health Program Evaluation Melbourne, Vic.; 1997.

14. James PT, Leach R, Kalamara E, Shayeghi M. The worldwide obesity epidemic. Obesity research. 2001;9(S11):228S-33S.

15. Bruce B, Fries JF. The health assessment questionnaire (HAQ). Clinical and experimental rheumatology. 2005;23(5):S14.

16. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Australian and New Zealand journal of public health. 2005;29(2):136-42.

17. Tuominen U, Blom M, Hirvonen J, Seitsalo S, Lehto M, Paavolainen P, et al. The effect of co-morbidities on health-related quality of life in patients placed on the waiting list for total joint replacement. Health and quality of life outcomes. 2007;5(1):16.

18. Kriegsman DM, Deeg DJ, Stalman WA. Comorbidity of somatic chronic diseases and decline in physical functioning: the Longitudinal Aging Study Amsterdam. Journal of clinical epidemiology. 2004;57(1):55-65.

19. Detry MA, Ma Y. Analyzing repeated measurements using mixed models. Jama. 2016;315(4):407-8.

20. Baumeister H, Kriston L, Bengel J, Härter M. High agreement of self-report and physician-diagnosed somatic conditions yields limited bias in examining mental–physical comorbidity. Journal of Clinical Epidemiology. 2010;63(5):558-65.

21. Peeters GMEE, Alshurafa M, Schaap L, de Vet HCW. Diagnostic accuracy of self-reported arthritis in the general adult population is acceptable. Journal of Clinical Epidemiology. 2015;68(4):452-9.

22. Laslett LL, Quinn SJ, Winzenberg TM, Sanderson K, Cicuttini F, Jones G. A prospective study of the impact of musculoskeletal pain and radiographic osteoarthritis on health related quality of life in community dwelling older people. BMC Musculoskeletal Disorders. 2012;13(1):168.

23. Veronese N, Stubbs B, Solmi M, Smith T, Reginster J-Y, Maggi S. Osteoarthritis increases the risk of cardiovascular disease: data from the osteoarthritis initiative. The journal of nutrition, health & aging. 2018;22(3):371-6.

24. Rahman MM, Cibere J, Anis AH, Goldsmith CH, Kopec JA. Risk of type 2 diabetes among osteoarthritis patients in a prospective longitudinal study. International journal of rheumatology. 2014;2014.

6.8 Supplements

Supplement 6A: the publication of "The impact of comorbidities on health-related quality of life of people with osteoarthritis over ten years"

RHEUMATOLOGY

Rheumatology 2021;0:1–7
doi:10.1093/rheumatology/keab358

Original article

The impact of comorbidities on health-related quality of life of people with osteoarthritis over 10 years

Ting Zhao [1], Tania Winzenberg[1], Dawn Aitken[1], Barbara de Graaff[1], Hasnat Ahmad[1], Graeme Jones [1] and Andrew J. Palmer [1,2]

Abstract

Objective. To investigate the impact of total number and patterns of comorbidities on health-related quality of life (HRQoL) and identify the most prevalent and influential comorbidity patterns in people with OA over 10 years.

Methods. Participants from the Tasmanian Older Adult Cohort aged 50–80 years, with self-reported OA and data on comorbidities and HRQoL were included. Participants were interviewed at baseline ($n = 398$), 2.5 ($n = 304$), 5 ($n = 269$) and 10 years ($n = 191$). Data on the self-reported presence of 10 chronic comorbidities were collected at baseline. HRQoL was assessed using the Assessment of Quality of Life-4-Dimensions. The long-term impacts of the number and of the nine most prevalent combinations of cardiovascular (CVD), non-OA musculoskeletal (Ms), metabolic and respiratory comorbidities on HRQoL over 10 years were analysed using linear mixed regressions.

Results. Compared with comorbidity-free OA participants, the health state utility (HSU) of those with 2 or ≥ 3 comorbidities was respectively -0.07 and -0.13 units lower over 10 years, largely driven by reduced scores for independent living, social relationships and psychological wellness. Comorbidity patterns including 'CVD+Ms' were most influential, and associated with up to 0.13 units lower HSU, mostly through negative impacts on independent living (up to -0.12), psychological wellness (up to -0.08) and social relationship (up to -0.06).

Conclusion. Having more comorbidities negatively impacted OA patients' long-term HRQoL. OA patients with CVD and non-OA musculoskeletal conditions had the largest HSU impairment, and therefore optimal management and prevention of these conditions may yield improvements in OA patients' HRQoL.

Key words: osteoarthritis, comorbidity count, comorbidity patterns, HRQoL, HSUs, the AQoL-4D

Rheumatology Key messages

- Having more comorbidities negatively impacts OA patients' long-term HRQoL.
- The types and combinations of comorbidities vary in effect sizes and health dimensions influenced.
- 'Cardiovascular + non-OA musculoskeletal' is the most prevalent and influential comorbidity combination in the OA population.

Introduction

Osteoarthritis (OA) is a chronic and progressive condition that leads to substantial socioeconomic burden and low individual health-related quality of life (HRQoL) [1, 2]. People with OA are more likely to have comorbidities and lower HRQoL than people without OA, with a systematic review reporting a prevalence of any comorbidity of 67% in people with OA compared with 56% in those without [3]. Multi-attribute utility instruments (MAUIs) such as the Assessment of Quality of Life-4-Dimensions (AQoL-4D) can be used to measure HRQoL. Health dimension scores can be combined according to the AQoL-4D algorithm to generate a weighted index of the health state utility (HSU)—measuring the strength of an individual's or society's preference for a given health state [4].

Comorbidities have a negative impact on HRQol of people with OA, with most previous studies using the total number of comorbidities (comorbidity count) as the comorbidity measure [5–8]. Whilst these studies have improved our understanding of the potentially additive

[1]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia and [2]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia

Submitted 8 November 2020; accepted 11 April 2021

Correspondence to: Andrew J. Palmer, Menzies Research Institute Tasmania, University of Tasmania, Medical Science 1 Building, 17 Liverpool St., Hobart, Tasmania 7000, Australia,
E-mail: andrew.palmer@utas.edu.au

CLINICAL SCIENCE

impact, their reliance on an implicit and rather unrealistic assumption of a uniform HRQoL impact of all comorbidities (both individually as well as in combination with each other) is a limitation. In reality, comorbidities vary considerably between studies due to differences in data sources and study populations [5–8]. Moreover, the impact of specific conditions or combinations of conditions may be caused by each condition's specific pathophysiology, symptoms and treatments that may involve synergistic effects rather than additive effects on HRQoL [9–11]. Thus, it is also important to consider the HRQoL impacts of different patterns (i.e. types and combinations) of comorbidities in OA to improve on the existing count method.

Additionally, the long-term effects of comorbidities on HRQoL of people with OA are not well researched. To date, only a single study has reported that comorbidities are associated with increased long-term disability in people with OA [12], but none have investigated effects on HSUs and individual health dimensions. Knowing how comorbidities impact OA patients over the longer term could help guide decision-making of clinicians and people with OA with regards to prioritizing management choices for comorbidities to optimize HRQoL.

Our study aim was to fill these major evidence gaps by investigating the impact of total number and patterns of comorbidities on HRQoL (in terms of HSUs and health dimension scores) in people with OA over 10 years. It also identifies the most prevalent and influential comorbidity pattern in OA patients.

Methods

Study design

Participants came from the Tasmanian Older Adult Cohort (TASOAC), a prospective, population-based longitudinal study aimed at identifying the environmental, genetic and biochemical factors associated with the development and progression of OA at multiple joint sites (e.g. knee, hip, hand and spine). Participants aged 50–80 years in 2002 were selected from the electoral roll in Southern Tasmania, using sex-stratified simple random sampling without replacement. Participants were excluded if they resided in an aged care facility or were unable to have a knee MRI scan. The research was approved by the Southern Tasmanian Health and Medical Human Research Ethics Committee and written informed consent to participate in the study was obtained from all participants. Only those who reported having doctor-diagnosed OA were included in our study. Data were collected using interview-administered questionnaires at four periods: baseline (from February 2002 to September 2004, $n = 398$), 2.5 years later (range 1.4–4.8 years, $n = 304$), 5 years later (range 3.6–6.9 years, $n = 269$) and 10 years later (range 9–13 years, $n = 191$).

Diagnosis of osteoarthritis

At baseline, participants were asked 'Have you had been told by a doctor that you have osteoarthritis at any of these sites'. Seven sites were listed: neck, back, hands, shoulders, hips, knees and feet. Participants chose 'yes' or 'no' for each site. Participants answering 'yes' to at least one site were considered as diagnosed with OA and were included in our study. We categorized participants as having knee and/or hip OA if at least one of the sites of knee and hip (individually or in combination with other joint sites) was indicated, and as having other types of OA if at least one of the five joint sites (other than knee and hip) was indicated.

Comorbidity counts and patterns

Data on comorbidities were collected at baseline by asking 'Have you ever been diagnosed by a doctor as having any of the following'. Ten chronic conditions were listed: diabetes, heart attack (a coronary, coronary occlusion, coronary thrombosis, myocardial infarction), hypertension (high blood pressure), thrombosis, asthma, bronchitis/emphysema, osteoporosis, hyperthyroidism, hypothyroidism and RA. Participants were given the choice between answering 'yes' or 'no' to each comorbid condition. Those answering 'yes' were considered as having that comorbidity. The total number of comorbidities of each participant was calculated by summing the 'yes' answers for each comorbidity and were classified into four groups: 0, 1, 2 and ≥ 3.

Based on the type of individual comorbid conditions, patients were then categorized into four broad comorbidity groups: (i) cardiovascular disease (CVD) group, which included heart attack, hypertension and thrombosis; (ii) non-OA musculoskeletal disease (Ms) group, including osteoporosis and RA; (iii) metabolic disease (Me) group, including hyperthyroidism, hypothyroidism and diabetes; and (iv) respiratory disease (Re) group, including asthma and bronchitis/emphysema.

Sixteen potential comorbidity patterns were identified based on the absence or presence of the above-mentioned four comorbidity types individually or in combination with each other [i.e. comorbidity-free OA participants (control group); OA participants with one of the four (CVD, Ms, Me or Re) comorbidity types—four possible patterns; OA participants with any two comorbidity types—six patterns; OA participants with any three comorbidity types—four patterns; and OA participants with all four comorbidity types—one pattern].

Measurement of health-related quality of life

HRQoL was assessed at baseline, 2.5, 5 and 10 years using the well validated AQoL-4D questionnaire, with utility weights derived from an Australian population sample [13]. AQoL-4D consists of 12 items covering four dimensions. The four dimensions and their corresponding items are (i) independent living (self-care, activities of daily living and mobility), (ii) physical senses (sight, hearing and communication), (iii) social relationships (social isolation, relationship and family role) and (iv) psychological wellbeing (sleep, anxiety and pain). The scores for items in each dimension were transformed and

calculated to provide dimension scores and an overall index of HSU using the AQoL-4D algorithm (Supplement 1, available at *Rheumatology* online) [4]. HSU and dimensional scores range from 0.00 to 1.00, with higher score indicating better HRQoL.

Other characteristics

Sex and date of birth were collected at baseline, and weight and height were collected at all periods. Weight was measured to the nearest 0.1 kg (with no shoes/socks/bulky clothing/headwear) using a single pair of calibrated electronic scales (Seca, Seven Hills [NSW], Australia Delta Model 707). Height was measured to the nearest 0.1 cm (barefooted) using a stadiometer. BMI was calculated as weight (kg)/(height (m)^2, and participants were grouped into three categories of BMI, normal weight (BMI $<$25.0 kg/m^2), overweight (BMI $\geq$25.0 kg/m^2$\leq$29.9 kg/m^2) and obese (BMI $\geq$30.0 kg/m^2), following the WHO criteria [14]. Data on disability severity were collected using the Health Assessment Questionnaire Disability Index (HAQ-DI) [15] at all periods. This has eight sections (dressing, arising, eating, walking, hygiene, reach, grip and activities) with 20 items. The HAQ-DI score (ranging between 0 and 3) was calculated and categorized into two groups: no to mild disability (HAQ-DI levels 0–1) and moderate to very severe disability (HAQ-DI levels $>$1) [15].

Statistical analysis

We summarized the demographic and other features of our sample using descriptive statistics. Linear mixed regressions were conducted to estimate the impact of baseline comorbidity status (i.e. total number of comorbidities and comorbidity patterns) on HRQoL (in terms of HSUs and individual health-dimension scores) of people with OA over 10 years, with adjustment for age (50–59 = 1, 60–69 = 2, 70–79 = 3, $\geq$80 = 4), sex (male = 1, female = 2) and time of follow-up. In order to produce estimates with a meaningful level of precision, comorbidity patterns with $<$10 participants were excluded from the regression analysis so a total number of nine most prevalent comorbidity patterns that were assessed for their influence on HRQoL. Subgroup analyses were conducted to evaluate whether the HRQoL impacts of (the number and patterns of) comorbidities vary between patients with knee and/or hip OA and other types of OA. The difference in HSUs of participants with different numbers and patterns of comorbidity were evaluated based on both statistical and clinical significance. Statistical significance was set as a *P*-value $\leq$0.05 (two-tailed) and results were considered clinically significant if they met or exceeded the minimum clinically important difference (MID) in the AQoL-4D HSUs of 0.06 (95% CI: 0.03, 0.08) units for the Australian population [16]. Regression estimates and 95% CI of the effects of comorbidity status on HRQoL were reported, which represent the average difference in HRQoL (HSU and dimensional) scores of each comorbidity category of OA patients compared with the reference (no comorbidity) group over 10 years. All analyses were conducted in Stata (Stata 15.1, StataCorp, College Station, TX, USA).

Results

Three hundred and ninety-eight participants with OA were included in our analyses (Table 1). The mean (S.D.) age at baseline was 64.0 (7.6) years. From Table 1, 39.4% (n = 157) of participants were female, 69.8% (n = 278) were overweight or obese, and 51.0% (n = 203) had knee and/or hip OA. The median (interquartile range) number of comorbidities for OA participants at baseline was 1 (1–2). Comorbidity counts ranged between 0 and 7, with 24.6% (n = 98) participants having no comorbidities at baseline, and 33.4% (n = 133), 20.9% (n = 83) and 21.1% (n = 84) having one, two and three or more comorbidities, respectively. The mean

Table 1 The demographic and other characteristics of participants at baseline (n = 398)

Variable	OA participants (n = 398)
Sex, *n* (%)	
Female	157 (39.4)
Male	241 (60.6)
Age, *n* (%)	
50–60 years	139 (34.9)
60–70 years	154 (38.7)
70–80 years	105 (26.4)
BMI, *n* (%)	
Normal ($<$25.0 kg/m^2)	120 (30.2)
Overweight (25.0–29.9 kg/m^2)	156 (39.2)
Obese ($\geq$30.0 kg/m^2)	122 (30.7)
HAQ-DI, *n* (%)	
No to mild (HAQ-DI 0-1)	349 (87.7)
Moderate to very severe (HAQ-DI $>$1)	49 (12.3)
OA affected join sites[a], *n* (%)	
Knee and/or hip OA	203 (51.0)
Other types of OA	195 (49.0)
Number of comorbidities[b], *n* (%)	
0	98 (24.6)
1	133 (33.4)
2	83 (20.9)
$\geq$3	84 (21.1)
AQoL-4D, mean (S.D.)	
HSUs	0.71 (0.20)
Independent living score	0.94 (0.13)
Social relationships score	0.92 (0.13)
Physical senses score	0.93 (0.09)
Psychological wellness score	0.85 (0.10)

[a]OA affected joint sites were based on self-reported OA diagnosis. Other types of OA included OA at joint sites of hand, shoulder, feet, neck and back. [b]Number of comorbidities was based on the self-reported comorbidities. AQoL-4D: Assessment of Quality of Life-4-Dimensions; HAQ-DI: Health Assessment Questionnaire Disability Index; HSUs: health state utilities.

Table 2 Number of participants by comorbidity patterns at baseline ($n = 398$)

Comorbidities type[a]	*n* (%)
No comorbidities	**98 (24.6)**
CVD only	**83 (20.9)**
Non-OA Ms only	**38 (9.5)**
Me only	7 (1.8)
Re only	**25 (6.3)**
CVD+Ms	**42 (10.6)**
CVD+Me	8 (2.0)
CVD+Re	**23 (5.8)**
Ms+Me	3 (0.8)
Ms+Re	**12 (3.0)**
Me+Re	6 (1.5)
CVD+Ms+Me	**13 (3.3)**
CVD+Ms+Re	**23 (5.8)**
CVD+Me+Re	5 (1.3)
Ms+Me+Re	3 (0.8)
CVD+Ms+Me+Re	9 (2.3)

[a]$n = 16$ comorbidities combinations were found. Comorbidity patterns with >10 participants are given in bold. CVD included heart attack, hypertension and thrombosis; non-OA Ms included osteoporosis and RA; Me included diabetes, hyperthyroidism and hypothyroidism; Re included asthma and emphysema. CVD: cardiovascular diseases; Me: metabolic diseases; Ms: musculoskeletal diseases; Re: respiratory diseases.

(S.D.) HSU at baseline was 0.71 (0.20). The mean HSUs at each time point and by comorbidity status (i.e. total number of comorbidities and comorbidity patterns) are shown in Supplementary Figs 1–3 (available at *Rheumatology* online).

As shown in Table 2, 16 comorbidity patterns were identified at baseline, nine of which were seen in 10 or more participants. Among those with comorbidities, the most frequent comorbidity patterns were CVD (20.9%), 'CVD+Ms' (10.6%) and Ms (9.5%).

Table 3 shows the impacts of total number of comorbidities on HSU and health dimension scores over 10 years. The adjusted difference in HSU between participants with zero and one comorbidity was neither statistically significant nor clinically important over 10 years. However, participants with two comorbidities recorded clinically meaningful and statistically significant reductions in mean HSU [−0.07 (95% CI: −0.12, −0.02)] over 10 years compared with comorbidity-free participants. Participants with two comorbidities had significantly lower scores on independent living [−0.06 (95% CI: −0.09, −0.03) units], social relationships [−0.03 (95% CI: −0.06, 0.00) units], and psychological wellness [−0.03 (95% CI: −0.05, 0.00) units] compared with those without comorbidity.

Participants with three or more comorbidities recorded the largest [−0.13 (95% CI: −0.19, −0.08) units] reduction in mean HSU over 10 years compared with those without comorbidity. This reduction is both clinically important and statistically significant. All four health dimensions were also substantially impacted in this group. Results were similar for the two sub-groups of affected joint sites (i.e. knee and/or hip joints; and other joints) (Supplement 2, available at *Rheumatology* online).

Table 4 shows the impacts of comorbidity patterns on HSU and health dimension scores over 10 years. Having any single comorbidity was not associated with statistically significant reduction in HSUs compared with having none. Of the combinations of two comorbidities, 'CVD+Ms' was associated with the largest reductions in HSUs [−0.13 (95% CI: −0.20, −0.07)], driven by the impacts on independent living [−0.08 (95% CI: −0.12, −0.04)], social relationships [−0.06 (95% CI: −0.10, −0.03)] and psychological wellness [−0.04 (95% CI: −0.08, −0.01)]. The impacts on independent living and psychological wellness became greater when Me or Re was combined with 'CVD+Ms'. Results were similar for the two sub-groups of affected joint sites (i.e. knee and/or hip joints; and other joints) (Supplement 3, available at *Rheumatology* online).

Discussion

This study fills a major evidence gap by investigating the associations between the total number and patterns of comorbidities and OA-related HRQoL over 10 years using a population-based cohort and identifying the

Table 3 The impacts of total number of comorbidities on OA participants' HSUs and health dimension scores over 10 years

Number of comorbidities	HSUs	Independent living	Social relationships	Physical senses	Psychological wellness
0 ($n = 98$)	Reference	Reference	Reference	Reference	Reference
1 ($n = 133$)	−0.03 (−0.07, 0.02)	−0.02 (−0.05, 0.01)	−0.01 (−0.04, 0.01)	0.00 (−0.02, 0.02)	−0.01 (−0.04, 0.01)
2 ($n = 83$)	**−0.07 (−0.12, −0.02)**	**−0.06 (−0.09, −0.03)**	**−0.03 (−0.06, 0.00)**	0.00 (−0.02, 0.02)	**−0.03 (−0.05, 0.00)**
≥3 ($n = 84$)	**−0.13 (−0.19, −0.08)**	**−0.10 (−0.13, −0.06)**	**−0.04 (−0.07, −0.01)**	**−0.03 (−0.05, 0.00)**	**−0.06 (−0.09, −0.03)**

Data are presented as mean difference (95% CI). The reported β-coefficients of each category of comorbidity status represent average difference in HRQoL (HSU and dimensional) scores of each OA subgroup compared with the reference (no comorbidity) group over 10 years. Regressions were adjusted by age, sex and follow-up time. HRQoL: health-related quality of life; HSUs: health state utilities.

TABLE 4 The impacts of comorbidity patterns on OA participants' HSUs and health dimension scores over 10 years

Comorbidity pattern	HSUs	Independent living	Social relationships	Physical senses	Psychological wellness
No comorbidity ($n = 98$)	Reference	Reference	Reference	Reference	Reference
CVD only ($n = 83$)	−0.04 (−0.09, 0.01)	−0.03 (−0.06, 0.00)	−0.02 (−0.05, 0.01)	0.00 (−0.02, 0.02)	−0.02 (−0.04, 0.01)
Ms only ($n = 38$)	−0.05 (−0.12, 0.01)	−0.02 (−0.06, 0.02)	**−0.04 (−0.08, 0.00)**	0.01 (−0.01, 0.04)	−0.02 (−0.06, 0.01)
Re only ($n = 25$)	0.01 (−0.07, 0.08)	−0.01 (−0.06, 0.04)	0.02 (−0.02, 0.06)	0.00 (-0.03, 0.03)	0.00 (−0.04, 0.04)
CVD+Ms ($n = 42$)	**−0.13 (−0.20, −0.07)**	**−0.08 (−0.12, −0.04)**	**−0.06 (−0.10, −0.03)**	−0.02 (−0.05, 0.01)	**−0.04 (−0.08, −0.01)**
CVD+Re ($n = 23$)	−0.07 (−0.15, 0.01)	**−0.08 (−0.13, −0.03)**	−0.01 (−0.06, 0.04)	0.01 (−0.03, 0.04)	−0.03 (−0.08, 0.01)
Ms+Re ($n = 12$)	−0.04 (−0.15, 0.07)	−0.05 (−0.12, 0.02)	−0.01 (−0.07, 0.06)	−0.03 (−0.07, 0.02)	0.01 (−0.05, 0.06)
CVD+Ms+Me ($n = 13$)	**−0.13 (−0.23, −0.02)**	**−0.12 (−0.19, −0.06)**	−0.03 (−0.09, 0.04)	0.00 (−0.04, 0.04)	**−0.07 (−0.13, −0.02)**
CVD+Ms+Re ($n = 23$)	**−0.13 (−0.21, −0.05)**	**−0.10 (−0.16, −0.05)**	−0.03 (−0.07, 0.02)	0.00 (−0.04, 0.03)	**−0.08 (−0.12, −0.04)**

Data are presented as mean difference (95% CI). The reported β-coefficients of each category of comorbidity status represent average difference in HRQoL (HSU and dimensional) scores of each OA subgroup compared with the reference (no comorbidity) group over 10 years. Regression models were adjusted for age, sex and follow-up time. Combination groups with <10 respondents at baseline were excluded from modelling, leaving nine combination groups in the regressions. CVD: cardiovascular disease (including heart attack, hypertension and thrombosis); HRQoL: health-related quality of life; HSUs: health state utilities; Me: metabolic disease (including diabetes, hyperthyroidism and hypothyroidism); Ms: non-OA musculoskeletal disease (including osteoporosis and RA); Re: respiratory disease including (asthma and emphysema).

most prevalent and influential comorbidity patterns in an OA population. Comorbidities were strongly associated with lower HRQoL in people with OA—compared with comorbidity-free OA participants, people with ≥2 comorbidities experienced a clinically and statistically significant reduction of up to 0.13 HSU units over 10 years, largely driven by reduced scores on independent living (up to −0.10) and psychological wellness (up to −0.06). The commonest comorbidity combination was CVD and Ms conditions, which were associated with the largest HSU impairment, driven by negative impacts on independent living, social relationships and psychological wellness scores. These results highlight that cardiovascular and non-OA musculoskeletal conditions might be comorbidities that yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention. Our findings also emphasize the importance of adjusting for the number as well as patterns of comorbidities when assessing HRQoL in OA patients.

Comorbidities were common in our study, with CVD and MS being the most prevalent comorbid conditions. The mean age (64 years) and the proportion of participants affected by OA-related comorbidities (75%) in our sample were similar to those reported in a recently published systematic review [3]. The distributions of comorbidity count were also similar between the two studies, with 33%, 21% and 21% of our participants having 1, 2 and ≥3 comorbidities, respectively (compared with 29%, 25% and 24% in the review). The most prevalent comorbid condition in our participants was CVD (20.9%), and the most prevalent comorbidity combination was 'CVD+Ms' (10.6%). Once again, our results were consistent with previous studies that have identified the cardiovascular and non-OA musculoskeletal diseases as the two most commonly observed comorbid conditions in people with OA. This suggests that our findings can be generalized to OA populations in other geographic locations with enough degree of confidence.

Participants with two or more comorbidities recorded clinically meaningful and statistically significant reductions in mean HSUs over 10 years compared with comorbidity-free participants. Whilst the cross-sectional HSU difference between people with OA with and without comorbidities [17], and the impact of the number of comorbidities on HSU of OA patients [5] have previously been investigated, our study provides the first data of which we are aware describing impacts of OA-related comorbidities on long-term changes in HSU. As such, our findings can be instrumental in generating insights into the future health trajectories of OA patients with different comorbidity profiles.

Our study bridges another important evidence gap by assessing the nature and extent of the effects of various comorbidity patterns on the HRQoL in OA population, highlighting the importance of adjusting for the number as well as types and combinations of comorbidities. The investigation of combined effects is particularly relevant for people with OA who often have multiple comorbid

conditions [3]. We found the combination of 'CVD+Ms' was the most dominant predictor of HRQoL impairment, exhibiting additive as well as synergistic effects (as indicated by the −0.13 unit reduction in HSUs in this group, compared with −0.04 in the CVD only and −0.05 in the Ms only). The combination of 'CVD+Ms' also strongly impacted individual health dimensions with the resultant largest reduction in independent living scores, followed by social relationships and psychological wellness scores. The impacts on independent living and psychological wellness became greater when Me or Re was combined with 'CVD+Ms'. For OA with Ms only and with 'CVD+Re', the impairment was seen in social relationships and independent living, respectively. Our results show that as well as the number of comorbidities, the types and combinations of comorbidities vary in effect sizes and health dimensions influenced, and suggest that the CVD and Ms comorbid conditions might yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention. The underlying mechanisms behind the various impacts of comorbidity patterns on HSUs and health dimensions may be different [18] but was not examined in our study and this remains unclear for OA populations [3]. A more in-depth investigation of the prevalence of comorbidity patterns and their precise effects (as well as the mechanism through which these effects on various health dimensions are realized) should therefore be on the agenda for future research.

Linear mixed regression models are recommended when outcomes are measured repeatedly over time and are likely to be correlated within each patient [19]. The ability of linear mixed regression in incorporating fixed effects (the factors assumed to have the same effect across many patients) and random effects (the factors likely to vary substantially from patient to patient) and accommodating unbalanced data patterns even when the missing values are not completely at random [19] also make this method appropriate for our data.

Our study has some limitations. First, the number of participants in some comorbidity groups was quite small, and groups having <10 participants were not included in analysis. Only a small number of individuals had individual comorbidities, so we also could not identify the relative contribution of individual comorbidities within each comorbidity group. Future studies with larger sample sizes are needed to (i) validate our findings, (ii) provide more information about alternative comorbidity patterns that could not be examined in this study, and (iii) investigate how comorbidities in each category compare with each other. Second, participants were selected based on self-reports of doctor-diagnosed OA and comorbidities were also self-reported without medical record verification. While high agreement has been found between self-reported comorbidity measures and those sourced from patients' medical records [20, 21], there remains the potential for misclassification. For example, the prevalence of RA in TASOAC baseline was 12% [22], which is higher than the expected population prevalence [1]. Third, we lacked information on duration, severity and treatment histories of OA and associated comorbidities, which may have important consequences for OA patients' HRQoL. Finally, some patients may have developed comorbidities during the follow-up period, but our analysis did not consider this possibility and used only the baseline comorbidity data. As people with OA are more likely to develop comorbidities than non-OA people over time [23, 24], our results may underestimate the long-term impacts of comorbidities on HRQoL of OA participants.

Conclusion

Comorbidities were negatively associated with OA patients' HRQoL over 10 years, and the effect sizes and health dimensions influenced varied by numbers and patterns of comorbidities. The findings of the long-term impacts of comorbidities generate important insights into the future health trajectories of OA patients with different comorbidity profiles. Having three or more comorbidities was associated with 0.13, 0.10 and 0.06 unit reductions in HSU, independent living and psychological wellness scores, respectively. 'CVD+Ms' was the most prevalent and influential comorbidity combination in this OA population, and impacted HSU (−0.13 units), independent living (−0.08 units), social relationships (−0.06 units) and psychological wellness (−0.04 units). These results suggest that cardiovascular and non-OA musculoskeletal conditions might be the comorbidities that yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention.

Acknowledgements

We would like to thank all participants of TASOAC project (aimed at identifying the environmental, genetic, and biochemical factors associated with the development and progression of osteoarthritis at multiple sites) who provided the valuable data that informed the comprehensiveness of this paper. We would also like to thank all medical practitioners who were associated with TASOAC study. Finally, we thank to Mr Petr Otahal (Statistical Officer, Menzies Institute for Medical Research, University of Tasmania) for his help with the statistical analysis of this study. G.J. designed the TASOAC study that provided the data for our analyses. T.Z., A.P. and H.A. conceived and designed the initial concept and method of this manuscript, D.A., T.W. and B.G. helped in modifying and improving the original concept and methods. T.Z. accomplished the setting up of database and performed necessary data analyses and drafted the manuscript. All authors were involved in the manuscript review and final approval.

Funding: No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this article.

Disclosure statement: All authors declare no interest conflicts.

Data availability statement

The data that was generated from this study will not be deposited in a public repository due to privacy and consent restrictions. De-identified data can be made available from the corresponding author on reasonable request, subject to a data sharing agreement.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Ackerman IN, Pratt C, Gorelik A, Liew D. Projected burden of osteoarthritis and rheumatoid arthritis in Australia: a population-level analysis. Arthritis Care Res 2018;70:877–83.

2 Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clin Exp Res 2005;17:255–63.

3 Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in osteoarthritis: a systematic review and meta-analysis of observational studies. Arthritis Care Res 2020;72:991–1000.

4 Hawthorne G, Richardson J, Day N. Using the assessment of quality of life (AQoL) version 1. Melbourne: Centre for Health Program Evaluation, 2001.

5 Hosseini K, Gaujoux-Viala C, Coste J *et al.* Impact of comorbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis. Best Pract Res Clin Rheumatol 2012;26:627–35.

6 Geryk LL, Carpenter DM, Blalock SJ, DeVellis RF, Jordan JM. The impact of co-morbidity on health-related quality of life in rheumatoid arthritis and osteoarthritis patients. Clin Exp Rheumatol 2015;33:366–74.

7 Wesseling J, Welsing PM, Bierma-Zeinstra SM *et al.* Impact of self-reported comorbidity on physical and mental health status in early symptomatic osteoarthritis: the CHECK (Cohort Hip and Cohort Knee) study. Rheumatology 2013;52:180–8.

8 van Dijk GM, Veenhof C, Schellevis F *et al.* Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. BMC Musculoskelet Disord 2008;9:95.

9 Fortin M, Dubois M-F, Hudon C, Soubhi H, Almirall J. Multimorbidity and quality of life: a closer look. Health Qual Life Outcomes 2007;5:52.

10 Hunger M, Thorand B, Schunk M *et al.* Multimorbidity and health-related quality of life in the older population: results from the German KORA-Age study. Health Qual Life Outcomes 2011;9:53.

11 Mujica-Mota RE, Roberts M, Abel G *et al.* Common patterns of morbidity and multi-morbidity and their impact on health-related quality of life: evidence from a national survey. Qual Life Res 2015;24:909–18.

12 Ettinger WH, Davis MA, Neuhaus JM, Mallon KP. Long-term physical functioning in persons with knee osteoarthritis from NHANES I: effects of comorbid medical conditions. J Clin Epidemiol 1994;47:809–15.

13 Hawthorne G, Richardson J, Osborne R, McNeil H. The Australian quality of life (AQoL) instrument: initial validation. Melbourne: Centre for Health Program Evaluation, 1997.

14 James PT, Leach R, Kalamara E, Shayeghi M. The worldwide obesity epidemic. Obes Res 2001;9:228S–33S.

15 Bruce B, Fries JF. The health assessment questionnaire (HAQ). Clin Exp Rheumatol 2005;23:S14.

16 Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Aust N Z J Public Health 2005;29:136–42.

17 Tuominen U, Blom M, Hirvonen J *et al.* The effect of comorbidities on health-related quality of life in patients placed on the waiting list for total joint replacement. Health Qual Life Outcomes 2007;5:16.

18 Kriegsman DM, Deeg DJ, Stalman WA. Comorbidity of somatic chronic diseases and decline in physical functioning: the Longitudinal Aging Study Amsterdam. J Clin Epidemiol 2004;57:55–65.

19 Detry MA, Ma Y. Analyzing repeated measurements using mixed models. JAMA 2016;315:407–8.

20 Baumeister H, Kriston L, Bengel J, Härter M. High agreement of self-report and physician-diagnosed somatic conditions yields limited bias in examining mental–physical comorbidity. J Clin Epidemiol 2010;63:558–65.

21 Peeters GMEE, Alshurafa M, Schaap L, de Vet HCW. Diagnostic accuracy of self-reported arthritis in the general adult population is acceptable. J Clin Epidemiol 2015;68:452–9.

22 Laslett LL, Quinn SJ, Winzenberg TM *et al.* A prospective study of the impact of musculoskeletal pain and radiographic osteoarthritis on health related quality of life in community dwelling older people. BMC Musculoskelet Disord 2012;13:168.

23 Veronese N, Stubbs B, Solmi M *et al.* Osteoarthritis increases the risk of cardiovascular disease: data from the osteoarthritis initiative. J Nutr Health Aging 2018;22:371–6.

24 Rahman MM, Cibere J, Anis AH, Goldsmith CH, Kopec JA. Risk of type 2 diabetes among osteoarthritis patients in a prospective longitudinal study. Int J Rheumatol 2014;2014:620920.

Supplement 6.1 The Assessment of Quality of life 4-Dimensions (AQoL-4D) algorithm to calculate the dimension scores and HSUs

AQoL-4D is a multi-attribute utility instrument, it can provide individual health dimension scores that can be combined according to the AQoL-4D algorithm to obtain an overall index of the health state utility. The 'utilities' are, in effect, preference weights and final utility scores should reflect peoples' preferences more accurately than unweighted dimensional scores. The 'utilities' can be used in economic evaluations, and specifically, cost-utility analysis requiring the computation of quality-adjusted life years (QALYs).

Specifically, the AQoL-4D questionnaire consists of 12 items covering four dimensions. Each dimension has three items with four response levels. The four dimensions and their corresponding items are: 1) independent living (self-care, activities of daily living, and mobility), 2) physical senses (sight, hearing and communication), 3) social relationships (social isolation, relationship and family role) and 4) psychological wellbeing (sleep, anxiety and pain). The dis-utility value (dvQ) for 12 items is given in Table 1.

Table 1. Items dis-utility values

Dimension	Item	health level			
		1	2	3	4
Independent Living	1	0.000	0.154	0.403	1.000
	2	0.000	0.244	0.343	1.000
	3	0.000	0.326	0.415	1.000
Social Relationships	4	0.000	0.169	0.396	1.000
	5	0.000	0.095	0.191	1.000
	6	0.000	0.147	0.297	1.000
Physical Senses	7	0.000	0.145	0.288	1.000
	8	0.000	0.253	0.478	1.000
	9	0.000	0.219	0.343	1.000
Psychological Wellness	10	0.000	0.107	0.109	1.000
	11	0.000	0.141	0.199	1.000
	12	0.000	0.104	0.312	1.000

Each dimension disutility values (dvD) are estimated from the following equations:

dvD1 = (1.0989*(1-(1-0.6097*dvQ1) *(1-0.4641*dvQ2) *(1-0.5733*dvQ3)))

dvD2 = (1.0395*(1-(1-0.7023*dvQ4) *(1-0.6253*dvQ5) *(1-0.6638*dvQ6)))

dvD3 = (1.6556*(1-(1-0.2476*dvQ7) *(1-0.2054*dvQ8) *(1-0.3382*dvQ9)))

dvD4 = (1.2920*(1-(1-0.1703*dvQ10) *(1-0.2554*dvQ11) *(1-0.6347*dvQ12)))

Each dimension scores are derived from the dvD and calculated using the equation:

$$\text{Dimension score} = 1 - \text{dvD}$$

The overall index of HSU is calculated using the equation:

$$HSU = (1.04 * (1 - (0.841 * dvD1)) * (1 - (0.855 * dvD2)) * (1 - (0.931 * dvD3)) * (1 - (0.997 * dvD4))) - 0.04$$

Supplement 6.2. The impacts of total number of comorbidities on OA participants' HSUs and health dimension scores over ten years by OA affected joint sites.

Number of comorbidities	HSUs	Independent Living	Social Relationships	Physical Senses	Psychological wellness
Knee and/or hip OA					
0 (n=50)	Ref.	Ref.	Ref.	Ref.	Ref.
1 (n=59)	-0.04 (-0.12, 0.03)	-0.02 (-0.07, 0.03)	-0.03 (-0.08, 0.02)	0.01 (-0.02, 0.04)	-0.02 (-0.06, 0.01)
2 (n=45)	**-0.08 (-0.16, 0.00)**	**-0.07 (-0.12, -0.02)**	-0.05 (-0.10, 0.01)	0.00 (-0.04, 0.03)	-0.03 (-0.07, 0.01)
≥3 (n=49)	**-0.15 (-0.23, -0.07)**	**-0.12 (-0.18, -0.07)**	-0.04 (-0.09, 0.01)	-0.03 (-0.06, 0.01)	**-0.08 (-0.11, -0.04)**
Other types of OA					
0 (n=48)	Ref.	Ref.	Ref.	Ref.	Ref.
1 (n=74)	-0.02 (-0.08, 0.03)	-0.02 (-0.05, 0.01)	0.00 (-0.03, 0.03)	0.00 (-0.02, 0.02)	-0.01 (-0.04, 0.02)
2 (n=38)	-0.05 (-0.12, 0.01)	**-0.05 (-0.09, -0.01)**	-0.02 (-0.05, 0.02)	0.01 (-0.02, 0.03)	-0.02 (-0.06, 0.01)
≥3 (n=35)	**-0.11 (-0.18, -0.04)**	**-0.06 (-0.10, -0.03)**	**-0.04 (-0.07, 0.00)**	-0.02 (-0.05, 0.01)	**-0.04 (-0.08, -0.01)**

Note: Regressions were adjusted by age, sex, and the time of follow-up. Data were presented as mean difference (95% confidence interval). OA=osteoarthritis, HSUs=health state utilities. *The reported beta coefficients of each category of comorbidity status represent average difference in HRQoL (HSU and dimensional) scores of each OA subgroup compared with the reference (no comorbidity) group over ten years.*

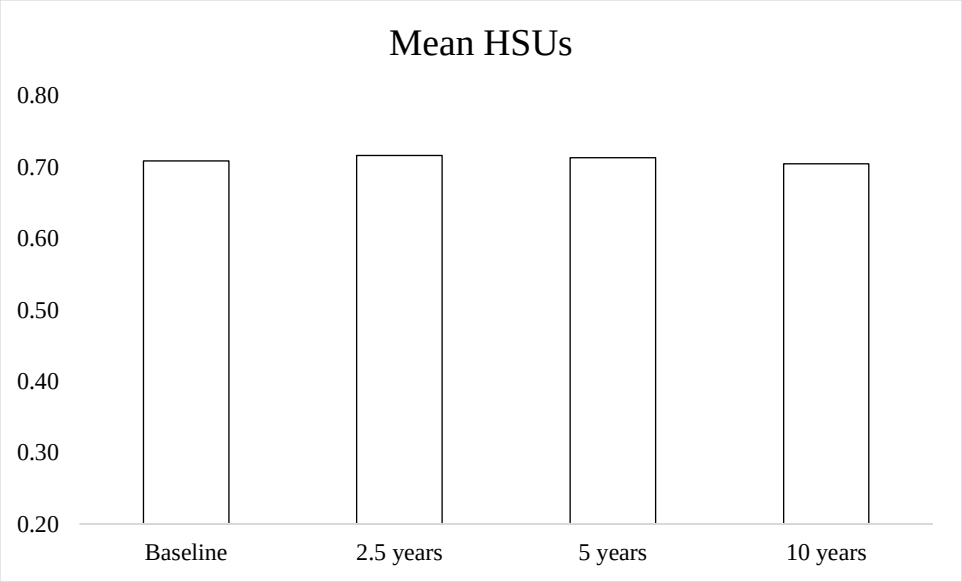
Supplement 6.3. The impacts of comorbidity patterns on OA participants' HSUs and health dimension scores over ten years by OA affected joint sites.

Comorbidity combinations	HSUs	Independent Living	Social Relationships	Physical Senses	Psychological wellness
Knee and/or hip OA					
No comorbidity (n=50)	Ref.	Ref.	Ref.	Ref.	Ref.
CVD only (n=40)	-0.07 (-0.15, 0.01)	-0.05 (-0.10, 0.01)	-0.03 (-0.09, 0.02)	0.00 (-0.04, 0.03)	-0.03 (-0.07, 0.00)
Ms only (n=13)	-0.05 (-0.17, 0.06)	0.00 (-0.08, 0.08)	**-0.09 (-0.16, -0.01)**	0.04 (-0.01, 0.08)	-0.02 (-0.08, 0.03)
Re only (n=14)	0.01 (-0.10, 0.12)	-0.01 (-0.08, 0.07)	0.03 (-0.04, 0.10)	0.01 (-0.03, 0.06)	-0.01 (-0.06, 0.05)
CVD+Ms (n=24)	**-0.15 (-0.24, -0.05)**	**-0.09 (-0.15, -0.03)**	**-0.08 (-0.14, -0.02)**	-0.03 (-0.07, 0.01)	-0.05 (-0.09, 0.00)

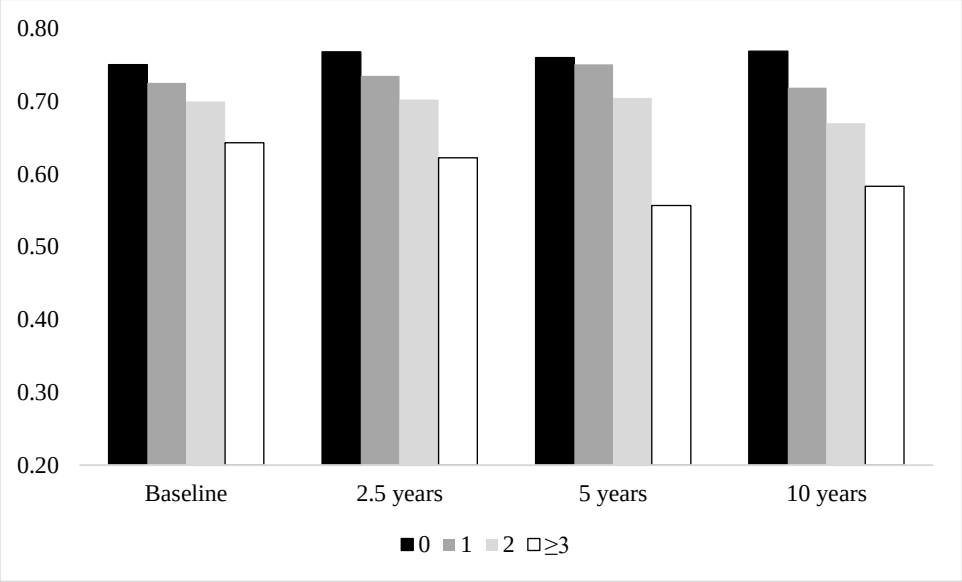
CVD+Ms+Re (n=17)	**-0.15 (-0.25, -0.04)**	**-0.12 (-0.20, -0.05)**	-0.05 (-0.12, 0.02)	0.00 (-0.04, 0.05)	**-0.08 (-0.14, -0.03)**
Other types of OA					
No comorbidity (n=48)	Ref.	Ref.	Ref.	Ref.	Ref.
CVD only (n=43)	-0.02 (-0.08, 0.04)	-0.02 (-0.05, 0.02)	0.00 (-0.03, 0.03)	0.00 (-0.03, 0.03)	-0.01 (-0.04, 0.03)
Ms only (n=25)	-0.07 (-0.14, 0.00)	-0.04 (-0.08, 0.00)	-0.02 (-0.06, 0.01)	0.01 (-0.03, 0.04)	-0.03 (-0.07, 0.01)
Re only (n=11)	0.01 (-0.08, 0.11)	0.00 (-0.05, 0.05)	0.02 (-0.03, 0.06)	-0.02 (-0.06, 0.03)	0.02 (-0.04, 0.07)
CVD+Ms (n=18)	**-0.12 (-0.19, -0.04)**	**-0.07 (-0.11, -0.02)**	**-0.05 (-0.09, -0.01)**	-0.01 (-0.05, 0.03)	-0.04 (-0.09, 0.00)
CVD+Re (n=14)	-0.03 (-0.12, 0.06)	**-0.05 (-0.10, -0.01)**	0.02 (-0.02, 0.06)	0.01 (-0.03, 0.05)	-0.01 (-0.06, 0.04)

Note: Models were adjusted by age, sex, and time. Data were presented as mean difference (95% confidence interval). Combination groups were excluded from modelling when there were less than 10 respondents at baseline, leaving 5 combination groups in each regression. OA=osteoarthritis; HSUs=health state utilities; CVD=cardiovascular disease (including Heart attack, Hypertension and Thrombosis); Ms=non-OA musculoskeletal disease (including Osteoporosis and Rheumatoid arthritis); Me=metabolic disease (including Diabetes, Hyperthyroidism and Hypothyroidism); Re=respiratory disease including (Asthma and Emphysema). *The reported beta coefficients of each category of comorbidity status represent average difference in HRQoL (HSU and dimensional) scores of each OA subgroup compared with the reference (no comorbidity) group over ten years.*

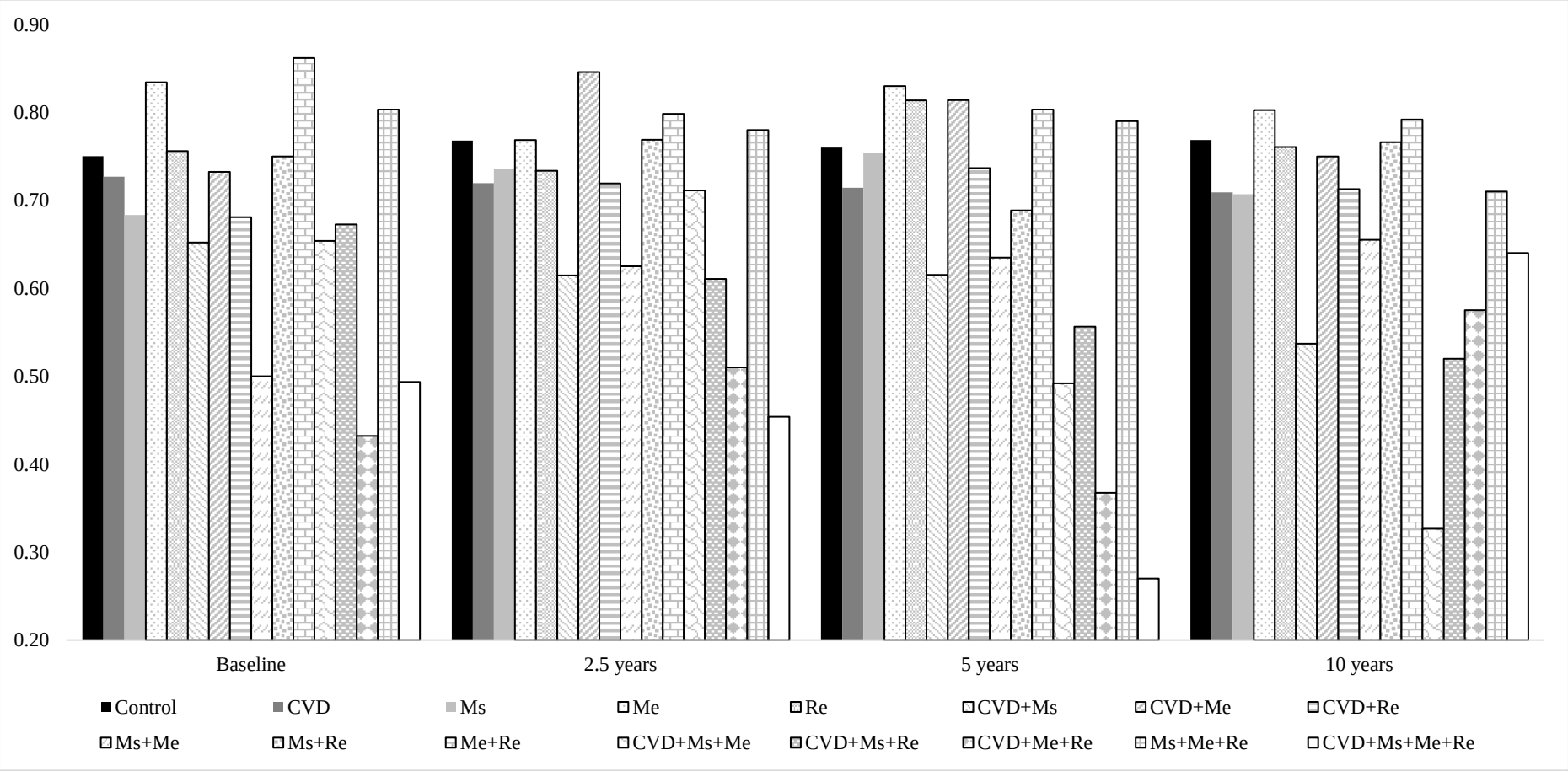
Supplement Figures 6.1-6.3



Supplement Figure 6.1. The crude mean health state utilities of participants at each timepoint



Supplement Figure 6.2. The crude health state utilities of participants at each timepoint by number of comorbidities



Supplement Figure 6.3. The crude mean health state utilities of participants at each timepoint by comorbidity patterns

Supplement 6.4 The potential impact of comorbidity incidence on results

The self-reported comorbidity data were collected at all four timepoints in TASOAC, however, the key reason that we are not able to estimate and then account for the incidence of comorbidities during follow-up periods was due to data inconsistency over time. For instance, all the covered OA-related comorbid conditions were chronic and incurable, therefore, if participants reported having a condition at a previous timepoint, it is reasonable to expect that they would have those conditions in the following timepoints as well. However, there was a lot of inconsistency in self-reported comorbidity data at four timepoints. Table 6.4.1 provides the 16 response possibilities at four timepoints for each individual condition to show the nature and extent of comorbidity related data inconsistences.

To answer the question of "Have you ever been diagnosed by a doctor as having any of the following", for each condition, participants were given two choices (1=Yes and 0, otherwise). For each single comorbidity, there could be 16 possible response combinations for four timepoints (Table 6.4.1). If we consider the response combinations related to all 16 (individual and in combination) comorbidities patterns considered in our analyses, the numbers of possible response combinations will explode. Considering the chronic nature of comorbidities, the self-reported data can be considered consistent only at 5 (i.e., Responses 1, 2, 7, 8 and 16) of the 16 possible response combinations in Table 6.4.1.

Table 6.4.1. The response possibility for individual conditions at four timepoints

Response possibility	Baseline	2.5 years	5 years	10 years	Data consistency
Response 1	**0**	**0**	**0**	**0**	**Y**
Response 2	**0**	**0**	**0**	**1**	**Y**
Response 3	0	0	1	0	N
Response 4	0	1	0	0	N
Response 5	0	1	0	1	N
Response 6	0	1	1	0	N
Response 7	**0**	**0**	**1**	**1**	**Y**
Response 8	**0**	**1**	**1**	**1**	**Y**
Response 9	1	0	0	0	N
Response 10	1	0	0	1	N
Response 11	1	0	1	0	N
Response 12	1	1	0	0	N
Response 13	1	1	0	1	N
Response 14	1	1	1	0	N
Response 15	1	0	1	1	N
Response 16	**1**	**1**	**1**	**1**	**Y**

Table 6.4.2 provides an example of response combinations for participants with asthma. As shown, there are n=30 (4+1+22+1+2) occasions at which participants reported inconsistent information about their asthma and the inconsistency rate of self-reported presence of asthma was estimated at 30/191=16%.

Table 6.4.2. Participants response for asthma at four timepoints

Response	N	Data consistency
Response 1	155	Y
Response 2	6	Y
Response 3	4	N
Response 4	1	N
Response 9	22	N
Response 10	1	N
Response 11	2	N
Total	191	

We did not evaluate the extent of inconsistent responses for all other (individual and pattern) comorbidities, but we anticipate that it will not be much different, so the follow-up information on comorbidities was not suitable to be used in our study.

In order to get some understanding of the magnitude and/or direction of bias as a result of OA-related comorbidities development after baseline, we did some corrections to participants' responses in Table 6.4.1 and calculated the total number of individuals who developed comorbidities during the study period (post-correction).Once the participants answered "Yes", all of the follow-up responses from those participants were also treated as "Yes", as due to the chronic nature of comorbidities, once they have it, they most likely have it forever. After we replaced the invalid "No" answers with the expected "Yes" answers, only the response combinations 2, 3, 4, 5, 6, 7 and 8 can be the indication of the development of comorbidities after baseline. For asthma, we found that among those who did not have self-reported asthma at baseline (response 1, 2, 3, 4, n=166), n=11 developed self-reported asthma after baseline. The 10-year cumulative asthma incidence rate among OA participants was therefore calculated at 0.07 (=11÷166). When we estimated and compared the HSUVs of n=11 participants who developed asthma after the baseline with those who did not (n=155), it was shown that the mean HSUV of the former was reduced by 0.10 units and that of later by 0.4 units from baseline to 10-years follow-up (Table 6.4.3). Therefore, it is likely we underestimated the impact of comorbidities on HRQoL of OA participants.

Table 6.4.3. The HSUVs for those who developed asthma and those who did not at four timepoints

	Baseline	2.5 years	5 years	10 years	Difference*

Developed asthma	0.58	0.62	0.54	0.48	0.10
No asthma	0.76	0.77	0.76	0.72	0.04

Note: *The difference was calculated between the baseline and 10-year follow-up.

However, as we demonstrated before, the situation will become exponentially complicated if we consider the inconsistencies facing all individual as well as combination comorbidities. Therefore, the study 4 analyses were only based on baseline self-reported comorbidities, and we have acknowledged this limitation in the main text and indicated that our results are likely to underestimate the impact of comorbidities on OA patients' HRQoL and also call for future research that based on more accurate measure of comorbidities for example medical record.

Supplement 6.5

Descriptive Statistics of TASOAC participants by study completion.

	Completed* (n=191)	Loss of follow-up^ (n=207)
Age (mean, SD)	62 (6.68)	66 (8.04)
BMI (mean, SD)	28 (4.94)	29 (5.55)
Number of comorbidity (mean, SD)	1 (1.12)	2 (1.54)
Gender		
Male	71 (37.17)	86 (41.55)
Female	120 (62.83)	121 (58.45)
OA affected join sites		
Other types of OA	109 (57.07)	86 (41.55)
Knee and/or hip	82 (42.93)	121 (58.45)
Comorbidity patterns		
No comorbidity	54 (28.27)	44 (21.26)
CVD	42 (21.99)	41 (19.81)
Ms	23 (12.04)	15 (7.25)
Me	4 (2.09)	3 (1.45)
Re	14 (7.33)	11 (5.31)
CVD+Ms	17 (8.9)	25 (12.08)
CVD+Me	3 (1.57)	5 (2.42)
CVD+Re	7 (3.66)	16 (7.73)
Ms+Me	2 (1.05)	1 (0.48)
Ms+Re	5 (2.62)	7 (3.38)
Me+Re	5 (2.62)	1 (0.48)
CVD+Ms+Me	3 (1.57)	10 (4.83)
CVD+Ms+Re	7 (3.66)	16 (7.73)
CVD+Me+Re	2 (1.05)	3 (1.45)
Ms+Me+Re	1 (0.52)	2 (0.97)
CVD+Ms+Me	2 (1.05)	7 (3.38)
AQoL-4D HSUVs (mean, SD)	0.76 (0.17)	0.67 (0.21)

Note:*Completed were those who completed all four timepoints survey. ^Loss of follow-up were those who lost at 10-year follow-up.

Chapter 7: Conclusions and future directions

7.1 Preface

In this final thesis chapter, a summary of key findings from the included studies related to health economic modelling of OA is provided. The chapter also summarises the strengths and limitations of this research thesis and provides directions for future research.

7.2 Summary of key findings from this thesis

OA is a common and costly chronic disease, affecting 1 in 5 Australians aged >45 years. Due to its chronic nature and high prevalence, OA poses a significant economic and humanistic burden to patients, their caregivers and society. Health interventions to manage OA range from core interventions, drugs, and surgical procedures, through to high technology implanted devices and digital healthcare technologies. Health economic evaluation models have been used in the OA field to compare alternative therapeutic options in terms of long-term economic costs and clinical effectiveness to identify the interventions that are effective, safe as well as being good value for money. While health economic evaluation models of OA have evolved substantially over time both because of methodological developments and better data availability, an avenue for further improvement exists both in terms of model design and the utilised model inputs.

Thus, the key focus of this research thesis is to provide directions for the improvement of health economic modelling practices of OA, internationally and in Australia. To achieve this key goal, Study 1 of this thesis comprehensively synthesised the evolution of health-economic evaluation models for all OA interventions (including treatments and preventions), with an emphasis on their strengths and weaknesses and study gaps. Findings from this part of research are important to inform the development of unbiased and openly accessible health economic models of various forms and interventions of OA to underpin decision-making in health policy and practice and to generate evidence to ensure rational use of scarce healthcare resources. As the correct implementation of multi-state health economic evaluation models not only requires the appropriate model design but also requires reliable and robust estimation of HSUVs and other key input parameters (e.g., disease management costs, and transition probabilities of changing disability levels), Study 2 of this research thesis systematically reviewed and meta-

analysed the HSUVs of people with OA to generate a HSUVs database for OA patients with different affected joint sites undergoing different treatments. Results from this part of research are instrumental to guide the HSUV choices in future health economic models of OA interventions.

Study 3 of this thesis then performed an assessment of the effects of OA on HSUVs and identified the physical and psychosocial drivers of these HSUVs effects using Australian data. Study 3 has addressed the paucity of Australian data on HRQoL impacts of OA by generating the locally driven HSUVs inputs to be used in future health economic models of OA for Australian and similar populations. Finally, as OA at its own cannot explain all the variation in HRQoL, Study 4 reported in this thesis examined the contribution of numbers and patterns of comorbidities on HRQoL of Australians with OA. These findings are important to guide the adjustment of HSUVs inputs for comorbidity patterns in the future OA health economic models of Australians and similar populations with alternative comorbidity profiles.

The key findings of this thesis are specified in the following sections.

7.2.1 The evolution of OA health economic evaluation models

Study one (reported in Chapter 3) comprehensively reviewed the evolution of health economic evaluation models for all forms of OA interventions including preventions, core treatments, adjunct non-pharmacological interventions, pharmacological and surgical treatments. It identified the weaknesses and study gaps relating to existing OA models, which provide the evidence and direction for future modelling practice to build an improved, gold-standard health economic model. The results from Chapter 3 suggested that OA modelled evaluations are of a wide variety and have evolved substantially over time, with their emphasis and complexity shifting from pharmacological-focused short-to-medium term decision-tree models to surgical-focused lifetime Markov models. Existing OA models have limitations related to the choice of model input parameters, discount rates, and model health states/events. For instance, indirect costs related to OA were mostly not considered. Discount rates were mostly consistent with local guidelines, however, most studies failed to gauge the sensitivity of the model outcomes to discount rate changes. Most studies failed to consider important model events (e.g.: cardiovascular adverse events), therapeutic adherence and treatment discontinuation. Despite clear guidelines, studies failed to pay adequate attention to lifestyle management, non-drug treatments and preventions. The reporting quality of included studies

was reasonably satisfactory; however, the title, abstract, and effectiveness measures were mostly reported inadequately.

Based on the comprehensive review of existing models, we recommend the use of Markov models for OA health economic evaluations due to their ability to incorporate repetitive and progressive (short and long-term) health events, including important medical adverse events, therapeutic adherence, and discontinuation. As the probabilities of medical adverse events, and time to events (e.g.: decision for joint replacement, revision surgery) depend on the history of previous states, we therefore recommend future studies to integrate memory into the models to avoid problems associated with the Markov assumption of memoryless-ness. We also recommend health economic evaluation studies of societal perspective to incorporate indirect costs such as costs from lost wages and the absenteeism/presenteeism costs associated with the management of OA. Finally, future models should also benefit from the recent availability of magnetic resonance imaging (MRI)-based data on OA definitions, progression and MRI-based markers, and advances in new data science.

7.2.2 OA-related HSUVs literature

Study two (reported in Chapter 4) systematically reviewed OA-related HSUVs and meta-analysed the HSUVs for people with different OA affected joint sites before and after various treatments. The systematic review identified important areas where the current evidence is lacking, namely under-represented geographical locations/ethnicities, affected OA joint sites, treatment options and HSUVs based on more sensitive MAUIs. For example, more than half (57%) of included studies were conducted in Europe, none in Africa, and only a limited number of studies (n=4) were based in Australia.

The meta-analyses of Chapter 4 generated a variety of HSUV inputs for alternative pre- and post- OA treatments to be used in future health economic evaluations models of various OA-related conditions and treatments. We found that HSUVs associated with four key treatment categories (core interventions, medication, injection, and surgery) often differed, as expected, pre- and post-treatment. Furthermore, significant inter-MAUI differences in the mean HSUVs were observed, which is as expected from alternative descriptive systems and utility algorithms backing these instruments. As such, this review provided important information that could be used by health economists and policy makers to determine the long-term disease outcomes and to identify cost-effectiveness of various OA treatments using health economic modelling techniques.

7.2.3 OA impacts on Australians' HRQoL

As informed by the findings from Study 2, there was a paucity of data on HRQoL impacts of OA in Australia and no previous study had investigated the longitudinal changes in OA-related HSUVs and their physical and psychosocial drivers. Therefore, Study 3 (reported in Chapter 5) provided a comprehensive assessment of the cross-sectional and temporal differences in HRQoL (in the forms of HSUVs and health-dimension scores) of people with and without OA using an Australian population-based cohort.

It was found that participants with OA had clinically important 0.07 (95% CI: -0.09, -0.05) units lower HSUVs over ten years compared to participants without OA. The main dimensions impacted by OA were psychological wellness, independent living, and social relationships. Whilst psychological wellness was equally impacted for both knee and/or hip OA and other types of OA, independent living and social relationships scores were substantially lower for people with knee and/or hip OA and the effect on the former increased over time. Appropriate and timely support to maintain independent living, social relationships and psychological wellness of OA patients may therefore have the potential to minimize the negative HRQoL impacts of OA.

An important finding was that OA impacts each health dimension with different intensity and timing, so interventions to improve HRQoL may need to be tailored to specific OA types and health dimensions, and the intensity of the interventions should be tailored over time. Such interventions might include those to manage psychological wellness targeting at pain, sleep, and anxiety/depression, maintain social relationships and independent living. Support to maintain psychological wellness of people with OA should be provided irrespective of the type and disease duration. However, support to maintain independent living could be more relevant to people with knee and/or hip OA, particularly those with longer disease durations.

7.2.4 Impacts of comorbidities on OA-related HRQoL

Study four (reported in Chapter 6) investigated the long-term impacts of the total number and patterns of comorbidities on OA-related HRQoL using an Australian population-based cohort and identified the most prevalent and influential comorbidity patterns in an OA population. It was found that OA participants with ≥ 2 comorbidities, especially the combination of cardiovascular and non-OA musculoskeletal conditions experienced significant reduction in HSUVs (0.13 units) over ten years compared with comorbidity-free OA participants. Lower

HSUV scores were largely driven by reduced scores on independent living, psychological wellness and social relationships. The combination of cardiovascular and non-OA musculoskeletal conditions was also the most prevalent comorbidity pattern in this OA sample.

The results from Chapter 6 highlight that cardiovascular and non-OA musculoskeletal conditions might be comorbidities that yield the greatest improvements in HRQoL of OA patients if targeted for optimal management and prevention. These results are instrumental in helping healthcare decision makers to achieve appropriate clinical judgements and management of OA patients considering the complete disease profile of patients not just the OA at its own. The findings also emphasize the importance of adjusting for the number as well as patterns of comorbidities when assessing HRQoL in OA patients. Finally, our findings of the impacts of numbers and patterns of comorbidities on HSUVs can be helpful to adjust the HSUVs model inputs for OA patients with different comorbidity profiles.

7.3 Contribution of the individual research studies to the overall aim of the thesis and their strengths and limitations

Studies reported in Chapters 3, 4, 5 and 6 of this research theses have contributed to the bridging of important evidence gaps in health economic modelling of OA (the overall aim of thesis) in two key dimensions: 1) by improving our understanding of the strengths and limitations facing existing OA health economic model types, structures and design; and 2) by providing a range of health economic model input parameters of OA-related HSUVs.

Specifically, the comprehensive systematic review of the evolution of health-economic evaluation models of all OA interventions (including preventions, core treatments, adjunct non-pharmacological interventions, pharmacological and surgical treatments) of Study 1 (reported in Chapter 3) identified the strengths, weaknesses and study gaps relating to existing OA models to provide the evidence and direction for future modelling practice in terms of model type, model structure/design/health states, model events, chosen discount rates and the choice of model input parameters. Findings from this part of the research will inform the future development of unbiased and openly accessible health economic models of OA. Findings from Study 1 suggested that a full range of important model events, available OA interventions (particularly, the lifestyle and non-drug treatments and preventions), OA management costs; and the impacts of therapeutic adherence, intervention discontinuation in OA health economic modelling practice should be considered. While Study 1 recommended the use of Markov type

state-transition models with memory integration, clear guidelines related to some area of OA health economic modelling could not be provided by our study. These included the limited capacity of study findings to set out a clear framework to guide the choice between 1) the cohort-level Markov models and individual patient-level (semi) Markov models, 2) short and long model cycles, and 3) continuous-time and discrete-time state transition models. These choices may be determined by several factors including the type of OA, data and software availability, nature of research questions to be answered, remaining life expectancy and other characteristics of the sample participants, number of follow-ups, longitudinal (balanced vs unbalanced) distribution of observations and expected computational efficiency of the model. Therefore, more research is needed to develop a clear framework to help choose between these options to further improve the development of unbiased and openly accessible health economic models of various forms and interventions for OA to underpin decision-making in health policy and practice and to generate evidence to ensure rational use of scarce healthcare resources.

Studies 2, 3, and 4 reported in Chapters 4, 5 and 6 of this thesis contributed jointly towards achieving the second key aim of this thesis (i.e., to ensure the availability of a comprehensive database of valuable HRQoL inputs to be used in future health economic models of OA, internationally and in Australia). Whilst study 2 has generated a database of useful HSU inputs that can be used in health economic models of various affected OA joint sites and interventions throughout the world (with some degree of caution), Studies 3 and 4 provided Australian derived OA-related HSUV estimates using a population-based cohort to be used in Australian OA health economic models, particularly those aimed at predicting long-term disease outcomes.

Study 2 comprehensively reviewed OA-related HSUVs to identify the areas where the current evidence is lacking and generated a database of metanalytic estimates of HSUVs to guide input choices in future health economic models of various affected OA joint sites and treatments worldwide. Whilst these results have great potential to improve the quality and consistency of OA-related HSUV parameter values for decision-analytic models, due to the inherent nature of our metanalytic results that are based on both the RCTs and observational studies, the HSUV estimates of Study 2 cannot be interpreted as direct estimates of intervention effects. However, as our meta-analytic HSUV estimates are based on a sufficiently large number of varieties of primary studies and rigorous statistical methods, these can be used (with some degree of caution) to populate health economic models of OA for people with different

affected OA joint sites undergoing different treatments. Future primary RCT studies assessing the HSUV impact of various treatments for people with different affected OA joint sites are recommended, so a database of more precise estimates of direct intervention effects can be provided.

Study 3 is the first to investigate the differences in HSUVs of Australians with and without OA and track these differences over ten years using a representative, population-based cohort of Australians with OA. Our data provided a broad indication of the nature and extent of the HRQoL impacts of OA and can be used in economic evaluation models to generate the estimates of OA's disease burden in Australia and to predict the long-term disease outcomes including life expectancy and quality adjusted life years. However, due to a limitation of our data that without giving the types of treatment patients underwent, their disease severity and duration, the utility difference between OA and no-OA population and their utility changes during follow up may be somewhat crude and, hence, less useful for cost-effectiveness analyses of OA interventions.

Study 4 is the first to investigate the long-term impacts of the total number and patterns of comorbidities on HRQoL of Australians with OA. The findings emphasized the importance of adjusting for the number as well as patterns of comorbidities when assessing HRQoL in OA patients. They have provided a road map to adjust HSUVs model inputs for OA-related comorbidities. The contribution of these findings for health economic models of OA is to guide the adjustment of HSUVs input for comorbidity numbers and patterns in the future OA health economic models of Australians and similar populations with alternative (inter-individual and within individual [longitudinal]) comorbidities numbers and patterns. An important limitation of these estimates was that HRQoL impacts of individual comorbidities as well as some comorbidity combinations were not estimated due to small sample sizes, therefore the adjustment for the HRQoL impact of individual comorbidities and some of the possible combinations of comorbidities in the health economic evaluation models cannot be achieved using these estimates.

7.4 Future directions and recommendations

While this thesis generated important insights into the current OA health economic modelling practices and provided directions for further improvements in model designs and inputs, there are some key areas where future research is urgently needed. These include the generation of other important model inputs including the costs and transition probabilities of progression of

OA, and the development and validation of a gold-standard OA economic model to identify cost-effective OA interventions. Specific areas of the future research attention are described in the following sections.

7.4.1 Transition probabilities

Transition probabilities of disease progression is one of the key input parameters for health economic evaluation models. It is well recognized that OA is one of the leading causes of disability among adults, the dynamic process of disability level in OA patients reflects the disease progression in response to disease activity and medical/personal interventions (1). The transition probabilities of disability level in OA patients can be helpful in health economic modelling evaluations to predict long-term disease outcomes. However, limited studies have investigated the transition probabilities of changing disability levels in OA for Australians and other population. Therefore, there is a need to generate those probabilities to be used in health economic evaluation models, to forecast disability trajectories of people with OA, and to investigate what sociodemographic or clinical features are associated with higher likelihood of worsening progression.

7.4.2 OA-related costs

Cost of illness is another essential input parameter for economic evaluation models. The studies investigating OA-related costs are predominantly conducted in United States and Europe (2-4). Only two studies have been conducted so far in Australia, however, even the most recent of these studies was published in 2001 (5, 6). Furthermore, these studies were either based on small sample size or did not capture all direct and indirect cost categories that are important to the detailed analysis of OA-related costs. Notably, the treatment landscape of OA in Australia has changed over time and differs substantially from that at the time of previous publications. The prevalence of OA as well as the incidence of total joint replacements (7, 8) has also recorded significant increases over time. All of this may have important consequences for OA-related economic costs. Therefore, cost of illness analyses based on high quality data sources are needed to comprehensively assess the economic burden of OA in Australia. The direct and indirect cost estimates by different categories, for example, by disability severity, affected OA joint sites and treatment types should also be investigated to generate costs inputs for future health economic models.

7.4.3 Development and validation of an improved, gold standard health economic evaluation model for Australians with OA

As stated previously, the existing OA economic models face several limitations. Therefore, benefiting from Study 1's recommendations and using the key parameters generated in this thesis as well as from the proposed future studies, the construction and validation of an improved health economic evaluation model for Australians with OA and similar population should be achieved to facilitate healthcare decision making about costs, benefits, and the value of funding OA health interventions.

7.5 Conclusions

In conclusion, this thesis presents a range of studies to fill important evidence gaps in health economics of OA. It improved our understanding of the strengths and limitations facing existing OA economic models and provided a wide range of model input parameters of OA-related HSUVs. It also provided reliable reference for optimal management of OA in Australia to improve OA people's HRQoL.

7.6 References

1. Verbrugge LM. Disability transitions for older persons with arthritis. Journal of Aging and Health. 1992;4(2):212-43.

2. Xie F, Kovic B, Jin X, He X, Wang M, Silvestre C. Economic and humanistic burden of osteoarthritis: a systematic review of large sample studies. Pharmacoeconomics. 2016;34(11):1087-100.

3. Puig-Junoy J, Zamora AR, editors. Socio-economic costs of osteoarthritis: a systematic review of cost-of-illness studies. Seminars in arthritis and rheumatism; 2015: Elsevier.

4. Salmon J, Rat A, Sellam J, Michel M, Eschard J, Guillemin F, et al. Economic impact of lower-limb osteoarthritis worldwide: a systematic review of cost-of-illness studies. Osteoarthritis and Cartilage. 2016;24(9):1500-8.

5. Mathers C, Penm R. Health system costs of injury, poisoning and musculoskeletal disorders in Australia, 1993-94: Australian Institute of Health and Welfare; 1999.

6. Lapsley HM, March LM, Tribe KL, Cross MJ, Brooks PM. Living with osteoarthritis: patient expenditures, health status, and social impact. Arthritis Care & Research: Official Journal of the American College of Rheumatology. 2001;45(3):301-6.

7. Ackerman IN, Pratt C, Gorelik A, Liew D. Projected burden of osteoarthritis and rheumatoid arthritis in Australia: a population-level analysis. Arthritis care & research. 2018;70(6):877-83.

8. Ackerman IN, Bohensky MA, Zomer E, Tacey M, Gorelik A, Brand CA, et al. The projected burden of primary total knee and hip replacement for osteoarthritis in Australia to the year 2030. BMC musculoskeletal disorders. 2019;20(1):1-10.