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 Zhao¹, M.S.; Tania Winzenberg¹, PhD, Professor; Barbara de Graaff¹, PhD, Research fellow;
Dawn Aitken¹, PhD, Research fellow; Hasnat Ahmad¹, 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

41 **Significance and Innovations**

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

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-preference-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 $\geq 80\%$ 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 $I^2 \geq 50\%$ 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.

168 *Affected OA joint sites*

169 Fifty-nine (39%) studies focused on knee OA and 41 (27%) focused on hip OA. Thirteen (9%)
 170 studies focused on both knee and hip OA and reported HSUVs separately. Two (1%) and six
 171 (4%) studies were focused on shoulder and hand OA, respectively. Thirty studies (20%)
 172 focused on mixed OA (Figure 2C).

173 *Knee OA*

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

179 *Hip OA*

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

184 *Other types of OA*

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

188 *Mixed OA type*

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

375 extraction. TZ conducted the meta-analyses and wrote the manuscript. All authors read and
376 approved the final manuscript.

377 **Acknowledgements**

378 The authors would like to thank Mr Petr Otahal (Statistical Officer, Menzies Institute for
379 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 review 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, Allegante 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. Nilsson 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.

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.