
Article:
Mukuria, C., Rowen, D. orcid.org/0000-0003-3018-5109, Harnan, S. et al. (4 more authors) (2019) An updated systematic review of studies mapping (or cross walking) measures of health related-quality of life to generic preference-based measures to generate utility values. Applied Health Economics and Health Policy, 17 (3). pp. 295-313. ISSN 1175-5652

https://doi.org/10.1007/s40258-019-00467-6

An updated systematic review of studies mapping (or cross walking) measures of health related quality of life to generic preference-based measures to generate utility values

Clara Mukuria, Donna Rowen, Sue Harnan, Andrew Rawdin, Ruth Wong, Roberta Ara, John Brazier

School of Health and Related Research (ScHARR), University of Sheffield, UK, Regent Court, 30 Regent Street, S1 4DA

Correspondence to: Clara Mukuria, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, S1 4DA, UK. c.mukuria@sheffield.ac.uk

Keywords: Mapping; Cross walking; Preference-based measures; QALYs

Short Running Title: Updated systematic review of studies mapping to preference-based measures

ABSTRACT:
Background: Mapping is an increasingly common method used to predict instrument-specific preference-based health state utility values (HSUVs) from data obtained from another health-related quality of life (HRQoL) measure. There have been several methodological developments in this area since a previous review up to 2007.

Objective: To provide an updated review of all mapping studies that map from HRQoL measures to target generic preference-based measures (EQ-5D measures, SF-6D, HUI measures, QWB, AQoL measures, 15D/16D/17D, CHU-9D) published from January 2007 to October 2018.

Data Sources: A systematic review of English language articles using a variety of approaches: searching electronic and utilities databases, citation searching, targeted journal and website searches.

Study selection: Full papers of studies that mapped from one health measure to a target preference-based measure using formal statistical regression techniques.

Data extraction: Undertaken by 4 authors using predefined data fields including measures, data used, econometric models and assessment of predictive ability.

Results: There were 180 papers with 233 mapping functions in total. Mapping functions were generated to obtain EQ-5D-3L/EQ-5D-5L-EQ-5D-Y (n=147), SF-6D (n=45), AQoL-4D/AQoL-8D (n=12), HUI2/HUI3 (n=13), 15D (n=8) CHU-9D (n=4) and QWB-SA (n=4) HSUVs. A large number of different regression methods were used with ordinary least squares (OLS) still being the most common approach (used ≥75% times within each preference-based measure). The majority of studies assessed the predictive ability of the mapping functions using mean absolute or root mean squared errors (n=192, 82%), but this was lower when considering errors across different categories of severity (n=92, 39%) and plots of predictions (n=120, 52%).

Conclusions: The last 10 years has seen a substantial increase in the number of mapping studies and some evidence of advancement in methods with consideration of models beyond OLS and greater reporting of predictive ability of mapping functions.

Key points for decision-makers

- Mapping or cross-walking enables utility values to be generated from other health related quality of life measures in studies where the preferred preference-based measure has not been used
- This updated review shows that compared to 10 years ago, there are now many mapping functions (n=233 across 180 published studies) which can be used to generate utility values from a large number of measures with the majority being to EQ-5D-3L
- Mapping studies now report important information that enable decision-makers to better understand and interpret mapped utility estimates, but mapping remains a second-best alternative to using the preference-based measures directly

1 INTRODUCTION

Resource allocation decisions in healthcare are frequently informed by Health Technology Assessment (HTA) using cost-effectiveness analysis where benefits are measured using quality adjusted life years (QALYs). QALYs are a measure that combine both health-related quality (HRQoL) and quantity of life, and can be used across patient groups and interventions, reflecting both changes in HRQoL and life expectancy. QALYs are generated by multiplying a quality adjustment weight (health state utility value (HSUV)) for health by duration, where the HSUV is often measured using a preference-based measure (PBM).

PBMs are made up of a descriptive system which describes HRQoL based on dimensions such as physical functioning, pain, social functioning and emotional functioning and with severity levels within each dimension. This allows HRQoL to be classified using a combination of the dimensions and severity levels in each dimension. For example, EQ-5D-3L has 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) each with 3 levels of severity (none, moderate, severe) which results in 243 health states that can be used to describe HRQoL [1]. Descriptive systems for PBMs have a HSUV on a 0 (dead) to 1 (full health) scale, with some measures having negative values for health states considered to be worse than dead. The HSUVs are usually derived from representative members of the general public who complete a preference elicitation task such as time trade off (TTO) or standard gamble (SG) to generate a value set. Patients complete the PBM and the value set is then applied to their responses to generate HSUVs.

Generic PBMs can be used across different conditions while condition-specific PBMs are specific to a condition [2]. The focus here is on generic PBMs. There are a number of generic PBMs, with the most common used being the EQ-5D-3L [3], though there is a new 5 level version, the EQ-5D-5L [1]. The other measures include the SF-6D [4, 5], the Health Utilities Index (HUI2 [6] and HUI3 [7]), the Assessment of Quality of Life measures (AQoL-4D, AQoL-6D and AQoL-8D) [8], the 15D [9], and the Quality of Wellbeing-Self Administered (QWB-SA) [10]. The majority of PBMs are adult measures although there are child or adolescent versions: 16D and 17D [11, 12] and EQ-5D-Y [13]; while AQoL-6D has been adapted for use in adolescent populations [14]. HUI2 was originally designed for children and the original version of QWB-SA was tested in both adult and children populations [15]. The Child Health Utility 9 dimensions (CHU-9D) is another generic measure developed specifically for use in children [16].

PBMs vary with respect to the descriptive systems and the value sets. International agencies that review HTA evidence to inform resource allocation decisions often recommend the use of a particular PBM to generate QALYs [17]. For example, the National Institute for Health and Care Excellence (NICE) recommend that EQ-5D-3L is used in England [18] and the National Health Care Institute (Zorginstituut **(**ZIN)) recommends the EQ-5D-5L in the Netherlands [19]. Furthermore, many international agencies recommend that PBMs are scored using their own country value set, for example

the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia recommend that HSUVs are generated using Australian general population preference weights [20], and the same requirement is also recommended for France [21], the Netherlands [22], Spain [23], England [18] and Canada [24]. However, many clinical effectiveness studies that are used to inform HTA either do not include PBMs, or do not include the PBM that is preferred by the target reimbursement authority. This has led to a rapid increase in the number of mapping papers published in recent years that estimate functions to predict HSUVs.

A mapping function is a prediction equation that is generated using the statistical relationship between a measure, referred to as the start measure which is used in the study of interest and a target PBM (e.g. EQ-5D-3L), estimated using regression analyses [25, 26]. The mapping function is then applied in the study of interest to predict the target PBM, where the target PBM has not been administered directly. There are a number of considerations when undertaking mapping (see for example Longworth and Rowen [25], Wailoo et al [27] and Ara et al [26]).

Conceptual overlap between the start measure and the target PBM is an important consideration before mapping can be undertaken [28]. For example, if the target PBM covers generic aspects of HRQoL such as physical, mental and social functioning, then the source measure should also have questions and/ or dimensions which would be mapped onto these concepts. Regression analysis relies on the existence of a relationship between the two measures and where there is little conceptual overlap, mapping will not be a suitable solution for generating HSUVs. Beyond this, there are a number of methodological choices that can have an impact on the resultant mapping functions. The dataset used to undertake mapping is usually a convenient pre-existing dataset that has both the start measure and the target PBM such as a trial or observational study. It is important to consider whether the available dataset provides what is required to generate mapping functions that are appropriate for the population to which it will be applied, for example the range in clinical severity and age. Depending on the sample size, regression analysis can be used to obtain mapping functions to predict either HSUVs or the individual health dimensions of the PBM, typically referred to as response mapping. A larger dataset covering the full range of severity is usually required for response mapping to enable all the severity levels within each dimension of the PBM to be estimated. The benefit of response mapping is that individual country value sets can be applied to the predicted dimensions, whereas mapping to the HSUVs restricts the predictions to the specific country value set used in the regression analyses.

In addition to the choice of whether to predict HSUVs or responses to the dimensions, analysts also have a choice with regards to whether to use overall scores, dimension scores or items from the start measure as predictor variables. Age and gender may be considered important and their inclusion has been recommended where available [27]. Other variables such as clinical outcomes could also be included although the choice is partly determined by what variables are available in both datasets e.g. if

only dimension scores, age and gender are available, then the mapping function can only use these variables.

There are also options regarding which regression method to use. A previous systematic review [29] found the most common approach when mapping to HSUVs was ordinary least squares (OLS) but other approaches that reflect the distribution of the HSUVs had also been applied. These included Tobit and Censored Least Absolute Deviation (CLAD) to deal with the limited range of the HSUVs as their maximum value is 1. More recently, methods have been developed to deal with multi-modal distributions such as mixture models, and methods that combine different approaches such as the adjusted limited dependent variable mixture models (ALDVMM) [26, 27]. Techniques such as ordered logit or probit and multinomial logit or probit have been used for response mapping to PBM dimensions.

Finally, there are different statistics that can be used to assess the predictive abilities of the mapping functions. Much of the focus is on how well the mapping functions predict HSUVs compared to the observed HSUVs, since this is indicative of the accuracy of the mapping function at predicting utilities in another dataset. Errors can be generated that report the difference between observed and predicted HSUVs. However, mean errors can mask large differences where both positive and negative differences exist as well as masking any bias in errors, for example, larger errors at the extremes of HSUVs index. Therefore other measures of error such as mean absolute errors (MAE), mean squared errors (MSE) or root mean squared errors (RMSE) are preferred [25, 26]. These provide an assessment of the difference between predicted and observed values with smaller values preferred. Assessment of performance may be within the dataset used to predict the function or a separate dataset, with the latter usually based on a smaller subset of the predicted dataset that is not used in the prediction. The statistics on errors can also be presented based on severity either by HSUVs or by a different measure of severity such as a clinical measure. This is useful as there is evidence that errors may be associated with severity and the analysis can be informative for users of the mapping function. Plots of the predicted and observed HSUVs alongside corresponding errors are also informative. Studies may also assess results based on the expected sign, statistical significance and consistency of estimated coefficients for the explanatory variables.

The previous Brazier et al [29] systematic review found 30 studies reporting functions mapping from non-preference-based measures to PBMs in studies published before 2007 [29]. The most common technique was ordinary least squares (OLS), with very few studies using any other approaches. Most of the studies reported R-squared values, mean errors and overall RMSE or MSE. This previous review recommended that research should be undertaken to test the accuracy of mapping functions, and that a framework should be developed for the critical appraisal of mapping studies to enable policy makers to critically assess HSUVs generated using mapping studies. Since this time there has been a large number of mapping studies published and the methods to estimate mapping functions have increased. Several papers that make recommendations around the estimation of mapping functions including data

requirements, regression models, performance assessment and reporting standards have also been published [25-27, 30]. In addition a database developed by researchers at the Health Economics Research Centre (HERC) at Oxford University for studies mapping to EQ-5D measures is available [31]. This provides a valuable resource for researchers searching for appropriate mapping functions for EQ-5D. However, although the database includes other PBMs where they have been included in studies mapping to EQ-5D measures, it does not include studies that only map to other PBMs which may be useful for other international agencies beyond the UK. One aim of this review is to address this gap.

The overall objective of this paper was to update the 2010 mapping review [29] to identify studies that map to generic adult and child PBMs from other HRQoL measures including other preference-based measures. This would provide a resource for researchers searching for studies that estimate mapping functions that are appropriate for use in their context in order to meet particular international reimbursement agency requirements.

2 METHODS

2.1 Database and search terms

The search strategy comprised of 5 different searches: electronic databases, utilities databases, citation searching of key publications, targeted journal searching and targeted website searching. The search was limited by publication date from January 2007 to October 2018. Three electronic databases were searched: MEDLINE, Web of Science and the Cumulative Index to Nursing and Allied Health Literature (CINAHL). The search strategy used free-text synonyms for generic preference based measures of interest 'EQ-5D, SF-6D, HUI measures, QWB, AQoL measures, 15D/16D/17D and CHU-9D'. This was combined with free-text terms for map*, crosswalk*, cross walk*, cross-walk*, estimat* or predict*, transfer, transformation or deriv*. The search strategy used here differs from the original review which did not include all the generic preference-based measures in the search terms, but had EQ-5D, HUI and QWB. Searches were used across databases and were limited to English language publications. The search strategies are presented in Appendix 1.

Two utilities databases were searched: the HERC database of mapping [31] and the ScHARR Health State Utilities Database (ScHARRHUD) [32], a searchable database providing published HSUVs. Free-text terms for map*, crosswalk*, cross walk*, cross-walk*, estimat* or predict*, transfer, transformation or deriv* in the title field was applied. Citation searches were undertaken based on the previously published mapping review [29] in the Scopus (Elsevier) database. Citation searches were also undertaken in the Web of Science (Clarivate Analytics), Scopus (Elsevier) and Google Scholar for four mapping reviews [33-36] and three mapping guidelines [27, 30, 37].

Based on the previous review, four journals that have previously published the mapping papers were searched in the Web of Science (Thomson Reuter): i) Medical Decision Making ii) Value in Health iii) Medical Care and iv) Journal of Health Economics. The search strategy is presented in Appendix 1. Relevant websites included the ISPOR website via the ISPOR Scientific Presentations Database [38] and EuroQoL (developers of EQ-5D measures) website [39]. The search strategies are presented in Appendix 3.

2.1.1 Inclusion criteria

Studies that undertook statistical mapping between one HRQoL measure and one of the target generic PBMs (EQ-5D-3L/EQ-5D-5L/EQ-5D-Y, SF-6D, HUI2/HUI3, QWB-SA, AQoL-4D, AQoL-6D/AQoL-8D, 15D/16D/17D, CHU-9D) were included. This included mapping between preference-based measures. Only English language studies were included. There were no restrictions with regards to the population.

The primary interest was mapping between measures with reported mapping functions that could be applied with specific focus on generic PBMs. Conference abstracts, studies mapping to directly elicited HSUVs generated using valuation techniques such as visual analogue scale or time trade off (i.e. not to utility values generated using a generic preference-based measure), studies not using HRQoL measures as the start measure and studies mapping to condition specific PBMs were excluded. Studies that undertook factor analysis were also excluded as they are not used to predict utility values. Methodology studies were excluded where they did not provide mapping functions. Studies that only applied or tested existing functions were also excluded.

2.2 Measures

The review focuses on mapping studies where the target measure is any one of the following generic PBMs: EQ-5D-3L/EQ-5D-5L/EQ-5D-Y, SF-6D, HUI2/HUI3, AQoL-4D/AQoL-6D/AQoL-8D, 15D/16D/17D, QWB-SA and CHU-9D.

2.2.1 EQ-5D measures

EQ-5D is the most widely used generic PBM with three-level [3] and five-level [1] versions. Both versions have the same five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, but have different levels (3/5) of severity. The EQ-5D-3L has severity levels: none, moderate and severe; while the EQ-5D-5L has severity levels: none, slight, moderate, severe and unable/extreme. Preferences for the EQ-5D-3L have been derived using TTO as well as the visual analogue scale (VAS) and more recently for the EQ-5D-5L using discrete choice experiments (DCE). Value sets are available for 32 countries for the EQ-5D-3L and for 14 countries for the EQ-5D-5L [39]. In addition, there is a cross-walk algorithm that has been used to generate 3L HSUVs from the 5L descriptive system while value sets are under development [40]. This mapping algorithm can be used to generate utility values for Denmark, France, Germany, Japan, the Netherlands, Spain, Thailand, UK,

USA and Zimbabwe. The mapping algorithm has also been applied to Sri-Lankan [41] and Polish [42] EQ-5D-3L utility values as it maps to the health states not utility values. There is also a youth version EQ-5D-Y which was developed from the EQ-5D-3L for use in children and adolescents [13]. However, there are no HSUVs available for this measure and most studies apply adult HSUVs.

2.2.2 SF-6D

The SF-6D is a preference-based measure derived from the SF-36 and the SF-12. SF-36 is a non-preference-based HRQoL measure with 36 questions with eight dimensions and physical and mental health summary scores [43]. It is one of the most commonly used health measures. SF-12 is a shorter version derived from the SF-36 covering the same dimensions.. The SF-6D has six dimensions: physical functioning, role limitation, social functioning, pain, energy, mental health. Each dimension has between four and six severity levels [4]. SF-6D can describe 18,000 possible unique health states for the version obtained from the SF-36 [4] and 7500 possible unique health states for the version derived from the SF-12 [5]. Value sets are available for Australia [44], Brazil [45], China (Hong Kong) [46], Japan [47], Portugal [48], Spain [49] and the UK [4, 5], where values were elicited using SG or DCE.

2.2.3 HUI2 and HUI3

The HUI2 and HUI3 are both derived from the same 15 item self-completed questionnaire. The HUI2 was originally developed for use in paediatric oncology. HUI2 has six dimensions: sensation, mobility, emotion, cognition, self-care and pain (fertility is an extra dimension that is not commonly used), each with four or five severity levels resulting in 8000 health states [6]. HUI3 has eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain, each with five or six severity levels resulting in 972,000 health states [7]. Value sets are available for Canada [6] and UK [50] for HUI2, and for Canada [7], France [51] and Spain [52] for HUI3 using values elicited using visual analogue scale (VAS) and SG or SG on its own.

2.2.4 15D, 16D and 17D

The 15D has fifteen dimensions: mobility, vision, hearing, breathing, sleeping, eating, speech, excretion, usual activities, mental function, discomfort and symptoms, depression, distress, vitality, sexual activity [9]. Each dimension has five severity levels, resulting in 5×10^{15} health states. The value set was generated using VAS values elicited in Finland [9] with a recent value set generated using VAS for Norway [53].

The 16D (for adolescents aged 12–15 years) [11] and the 17D (for children aged 8–11 years) [12] were developed based on the 15D. The value set was generated using VAS with Finnish children aged 12-15 for the 16D while parents provided VAS values for the 17D.

2.2.5 AQoL measures

The Assessment of Quality of Life has different versions that can be used including the, AQoL-4D, AQoL-6D and the AQoL-8D with a different number of dimensions for each version [8, 54, 55]. AQoL-4D has four dimensions (independent living, senses, relationships and mental health) while AQoL-6D has six dimensions (independent living, mental health, coping, relationships, pain, senses). The latest version, the AQoL-8D, has eight dimensions: independent living, happiness, mental health, coping, relationships, self-worth, pain, senses which are based on a combination of a 35 items questionnaire resulting in $2.37*10^{23}$ health states [8]. The value set is generated using VAS and TTO values elicited in Australia.

AQoL-6D has also been adapted for use with adolescents. Value sets for the adapted AQoL-6D were generated using TTO in a sample of adolescents from Australia, Fiji, New Zealand and Tonga [14].

2.2.6 QWB-SA

The QWB-SA has 68 items in total with dichotomous symptom lists for 19 chronic symptoms, 25 acute symptoms and 14 mental health symptoms/behaviours along with 17 items that cover mobility, physical and social activity [56, 57]. The original version was developed and tested for use in both adults and children [15]. It defines 945 health states and the value set is generated using VAS values elicited in the USA.

2.2.7 CHU-9D

The Child Health Utility Index 9D (CHU-9D) has 9 dimensions: worry, sadness, annoyance, tiredness, pain, sleep, daily routine, school, and activities [16]. Each dimension has 5 levels of severity resulting in 1,953,125 health states. The value set has been generated using SG in the UK using an adult sample [58], best worst scaling (BWS) in Australia using an adolescent and adults [59, 60] and BWS with TTO in a sample of adolescents in China [61].

2.3 Study Selection

Three authors undertook study selection. SH undertook an initial eligibility assessment across all of the identified studies, excluding studies that were identified as not mapping studies from titles. CM and DR undertook further eligibility assessment based on abstracts and the full-paper review. Queries were discussed between the reviewers.

2.4 Data Extraction and Analysis

An extraction template was created by DR and reviewed by the team. The final template included information on the start measure and the target PBM, whether measures were self-completed, the country value set used, whether mapping was done to the HSUV index or the dimensions. Details about the estimation dataset including the population e.g. patient characteristics, the sample country, and the

sample size were extracted. For methods, the regression techniques was extracted and whether additional explanatory variables (clinical, age, gender) were recorded. Whether the regression coefficients were reported was noted. Information on how studies assessed performance of the mapping functions including MAE/MSE/RMSE, including by severity, and plots of predictions was also noted. Actual values e.g. regression coefficients or MAE were not extracted. Independent extraction was undertaken by AR, CM, DR and SH. AR, SH and CM did initial extraction and CM and DR double-checked the extractions.

Analysis was based on the number of studies across the preference-based measures as well as in relation to methods used. The Brazier et al [29] review noted that OLS was the most common method but they noted the limitation of using OLS for mapping. R-squared was also reported in studies in the previous review but this does not provide useful information on whether mapping functions are valid and was not extracted for this review. Therefore we considered what type of regression methods were used and what information was used in comparison of mapping functions i.e. use of MAE/RMSE or plots. No quality assessments were made with regards to the reported mapping functions as judgements about the validity of mapping functions depend in part on the context in which they will be applied.

3 RESULTS

3.1 Studies included

A total of 2967 papers were identified from the different sources after removal of 2444 duplicates (Figure 1). The number of relevant papers was reduced to 1278 based on a review of the titles and a further 707 were excluded based on review of abstracts. Further review of the remaining full papers resulted in a total of 180 papers (Figure 1).

Figure 1 Study Selection Flowchart

<insert Figure 1>

3.2 General description of studies

The mapping functions are presented in Appendix 2 ordered by the target PBM then the source measure. Some papers reported mapping functions to more than one PBM. Where a study reports a mapping to a single PBM, this was counted as a single mapping function even where different regression methods were used or different specifications tested. Where papers reported mapping to more than one PBM e.g. to EQ-5D-3L and to SF-6D this was counted as two mapping functions. There were 233 mapping functions published across the 180 papers. Highlighted rows indicate mapping functions reported in papers that estimate mapping functions to more than one PBM.

There were 8 different 15D mapping functions, 12 for AQoL measures, 4 for CHU-9D, 147 for EQ-5D measures, 13 for HUI measures, 4 for QWB-SA and 45 for SF-6D. For AQoL, 4 were mapping functions to AQoL-4D while the rest were to AQoL-8D. EQ-5D mapping functions were mainly to EQ-5D-3L (n=124) with 22 mapping functions to EQ-5D-5L and 1 to EQ-5D-Y. There were 2 mapping functions that used HUI2 and there were 12 mapping functions using HUI3.

The majority of the mapping functions were between non-preference-based measures and the target PBM (Appendix 2). The most common condition-specific non-preference-based start measures were the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ C30) [240] used in 25 mapping functions, variations of the Functional Assessment of Cancer Therapy (FACT) [241] including the general version, breast cancer, melanoma and prostate cancer (n=12) and the Health Assessment Questionnaire (HAQ) including HAQ-Disability index [242] for rheumatoid arthritis (n=13). The rest of the start measures included condition-specific HRQoL measures for a number of different conditions[1].

There were also generic non-preference-based HRQoL measures used as the start measure including the SF-12 or SF-36 [43, 243] (n=12), the Nottingham Health Profile [244] (n=2), the Patient Reported Outcomes Measure Information Systems (PROMIS) [245] (n=4) and the WHO-QoL BREF [246] (n=4). One study used a measure for older people (Older People's Quality of Life Brief Questionnaire [247]) and one used a measure for women (Women's Health Questionnaire-23 [248]) to generate 2 mapping functions. One study used a general health or self-assessed health question [122].

The generic PBMs (15D, AQoL-8D, HUI3, QWB-SA, SF-6D, EQ-5D-3L/EQ-5D-5L) were also used as start measures. Two papers provided mapping functions between these six generic PBMs [62, 63]. Two papers mapped between EQ-5D-3L and EQ-5D-5L [40, 203].

For mapping functions to CHU-9D, there were two generic measures, the KIDSCREEN [249] and the Pediatric Quality of Life Inventory (PedsQL) [250], one mental health measure (Strengths and Difficulties Questionnaire [251]) and one weight measure (Weight-Specific Adolescent Instrument for Economic Evaluation (WAItE) [252]). The PedsQL was also used to map to EQ-5D-Y.

There were 17 (7%) mapping functions that used adult general population samples, 2 used women and 1 used school children with no specified conditions. The rest of the mapping functions were based on patient samples (n=213, 91%). Clinical trials were the most common source of data. Respondents were

[1] Mental health, diabetes, fibromyalgia, heart disease, asthma, stroke, osteoarthritis, osteoporosis, vision (e.g. cataract, macular degeneration), hearing, Chronic Obstructive Pulmonary Disease (COPD), skin conditions (e.g. psoriasis), epilepsy, problems with hips or knees, neck problems, back problems, sleep problems (e.g. insomnia), Parkinson's, overactive bladder, Cushing's Syndrome, Ankylosing Spondylitis, HIV, headaches, liver disease, inflammatory bowel disease, ulcers, constipation, multiple sclerosis, obesity, growth hormone deficiency and measures for palliative care.

also recruited from populations in the community, hospital and primary care. The two most common patient populations were cancer (n=38, 16%) and rheumatoid or osteoarthritis (n=23, 10%). Some populations were mixed, for example, Chen, 2015 [62], Richardson, 2014 [63] used the Multi-instrument Comparison (MIC) dataset [253] which has all the generic PBMs and contains both patients (self-identified) and members of the general population who were healthy. A number of studies relied on the MIC dataset to undertake mapping within specified patient populations including asthma [69, 212], depression [64], diabetes [65, 202], cancer [205] and heart disease [68].

Sample sizes varied widely. The smallest sample size was 30 respondents [81] while the largest was over 130,000 respondents [161] both of which were mapping functions to the EQ-5D-3L. However, very few mapping functions used a sample size that was less than 100 (n=10, 4%).

3.3 Estimation

The majority of functions mapped only to the HSUV index for the respective measures (n=191, 82%). A few mapping functions were only to the PBM dimensions (n=4, 2%) which were all mapping to EQ-5D-3L/EQ-5D-5L. The rest were mapping functions to both the HSUV index and PBM dimensions (n=38, 16%) of which 32 were mapping to EQ-5D measures, 4 to SF-6D, 1 to HUI3 and 1 to CHU-9D. For the 15D, AQoL-4D/AQoL-8D and QWB-SA, all the mapping functions were only to the HSUV index.

The specific country value set used was restricted by available value sets for 15D (Finnish[2]), AQoL (Australian) and QWB-SA (USA). One CHU-9D mapping function was to the Australian HSUV index [60] while the rest used the UK HSUVs. One HUI3 mapping function was to the Spanish HSUV index [133] while the rest were to the Canadian HSUVs. Three SF-6D mapping functions were to the Hong-Kong HSUV index [220, 227, 229] while the rest were to UK values. There were 85 mapping functions that used the EQ-5D-3L UK country HSUV with 12 of these using these values alongside another country HSUVs. Other EQ-5D-3L country HSUVs that were used included the Canadian, USA, Chinese, Dutch, European, German, Japanese, Korean, Spanish and Swedish value sets (Appendix 2). The English value set was used in 8 mapping functions for the EQ-5D-5L while the UK crosswalk was used for 5 mapping functions. EQ-5D-5L mapping functions also used new value sets for Netherland, Spain, Canada, Uruguay, China, Japan and Korea as well as crosswalk values for USA and the Netherlands.

Most mapping functions used more than one regression method (n=142, 61%). The most common regression method used when mapping to HSUVs was OLS which was used in all the mapping functions for CHU-9D and most of the mapping functions (≥75%) for the other PBMs (Table 1). Note that this does not necessarily reflect whether or not OLS is the most appropriate method. Other methods included those that aimed to take into account the limited range of the utility values such as CLAD and Tobit in addition to beta regressions and fractional logistic models. Generalised linear

[2] Norwegian value set only recently published

models which provide flexibility in the choice of the underlying distribution were also used. Robust regressions using the MM-estimator designed to deal with the potential impact of outliers in linear regressions were also used [254]. There were also models that allowed even greater flexibility such as mixture models which enable a mixture of distributions in the regressions when estimating HSUVs. Finally, models combined different regression approaches such as OLS and logit regressions in two or three part models specified by the analyst with results combined post-analysis. The different 'parts' were determined by the distribution of the HSUVs e.g. those who had in full health (HSUV=1) vs. those who had decrements in HRQoL (HSUV<1). These two or three part models are similar to mixture models in that they attempt to address distribution of HSUVs. Methods such as the ALDVMM combined consideration of the distribution of the data and the limited rage of HSUVs [203].

EQ-5D measures and SF-6D had the largest number of regression approaches applied when mapping to the HSUVs index (n=19 and n=17 respectively, Table 1) followed by HUI measures (n=11), CHU-9D (n=8), 15D and AQoL measures (n=6 each) and QWB-SA (n=5). For response mapping to PBM dimensions, ordered/ multinomial/ multivariate/ generalised/ partial proportional logit or probit regressions were undertaken. Response mapping was undertaken for EQ-5D measures (n=32), HUI3 (n=1), SF-6D (n=3) and CHU-9D (n=1).

Table 1 Regression methods used by PBM

<insert Table 1>

3.4 Specification and performance

In addition to the HRQoL source measures, age was included as a potential predictor in 51% (n=119) of the reported mapping functions while gender was included in 55% (n=126). Clinical outcomes such as body mass index (BMI) were included 20% (n=46) of the time (Table 2).

Table 2 Number of mapping functions including additional variables, coefficients and performance indicators by preference-based measure

<Insert Table 2>

The majority of the mapping functions (n=224, 96%) reported the model coefficients either in the paper or provided a way to generate the HSUV via a separate method such as an excel sheet or program. Some mapping functions were provided in separate programs to estimate HSUVs e.g. Adams et al, 2010 [132]. Some papers using Bayesian networks did not report regression coefficients e.g. Borchani et al, 2012 [164]. Other papers did not report coefficients because the results because authors judged that results were not suitable for use as mapping functions due to poor performance. For example both Bafus et al, 2012 [81] and Dzingina et al, 2017 [163] found that there were large differences between observed and predicted values.

Most of the mapping functions (n=192, 82%) had overall MAE, MSE or RMSE as part of assessment of performance (Table 2). Performance across different categories of severity was reported in 92 (39%) mapping functions while plots were provided for 120 mapping functions (52%). Plots included scatter or linear plots of observed against predicted HSUVs with or without errors and plots of the start measure against observed and predicted HSUVs.

4 DISCUSSION

This review identified 180 papers that met the inclusion criteria. The studies undertook a range of different regression methods, with the HSUV index as the most common dependent variable. Most of the start measures were non-preference-based measures with many studies including age and gender while a more limited number also included clinical measures such as BMI. Most of the studies assessed performance using MAE, RMSE or MSE while some studies also included plots of predicted HSUVs against the observed HSUVs or start measures, rather than relying on R-squared statistics as reported in the previous review [29].

There were 180 mapping papers with 233 mapping functions identified over the review period (2007 to 2018) compared to 30 papers which were included in the previous review (1996 to early, 2007) [29]. This reflects the growth in mapping studies that has taken place in the last decade. Half (n=15, 50%) of the mapping functions included in the previous review were mapping to EQ-5D whereas functions mapping to EQ-5D measures are reported in more than half (n= 147, 63%) of the mapping functions in this current review. This may reflect the recommendation to use EQ-5D and acceptance by NICE [18] of mapped HSUVs in their methods guide for health technology assessment which may have driven demand for mapping to EQ-5D. EQ-5D-3L is also the most widely used generic PBM.

Other PBMs had less mapping functions in both the previous review (HUI measures n=8, 27%; SF-6D n=5, 17%; AQoL measures n=2, 7%; QWB-SA and 15D n=1, 3%) and this review (HUI measures n=13, 6%; SF-6D n=45, 19%; AQoL measures n=12, 5%; 15D n=8, 3% and QWB-SA n=4, 2%). There has been an increase in the number of mapping functions to the SF-6D in this review compared to the previous review. Although there are a number of mapping functions available for 15D, AQoL measures, HUI measures and QWB-SA, many of the mapping functions rely on the same dataset, the MIC dataset. In addition, this current review included mapping functions to the CHU-9D which was under development during the last review. One EQ-5D mapping function was to the EQ-5D-Y but no other mapping functions to child or adolescent generic PBMs were identified.

Patient populations were the main type of datasets used to estimate the mapping functions (n= 213, 91%), rather than members of the general population which was also the case in the previous review where 20 (67%) papers used patient populations. The use of patient populations is potentially better suited to address concerns regarding using appropriate populations to generate mapping functions in terms of demographic characteristics as well as severity [27]. However, where the source is trial data

rather than registry data, this may result in small sample size. Trial data may also reflect very specific characteristics such as severity levels due to trial inclusion criteria. This can be useful if a mapping function is to be used in a similar population but may be restrictive if mapping functions are being applied in trials with different inclusion criteria or real world populations. The growing use of HSUVs and generic PBMs has led to growth in the inclusion of these types of measures in routine use e.g. as part of routine measures following hip and knee replacements in England [255] or registry data [256] which provide larger datasets of individuals who have received care. This offers alternative sources of data that can be used to undertake mapping as well as to test performance of mapping functions in the future.

The most common regression method to estimate mapping functions mapping to the HSUV index was OLS in both reviews. Brazier et al, 2010 [29] noted that OLS shows a systematic pattern in over-predicting at the lower end and under-predicting at the upper end of HSUVs and that alternative approaches should be tested. The previous review had one study that used Tobit and two used CLAD to address the bounded nature of HSUVs while less restrictive linear models such as the generalised linear model (GLM) were also used [29]. In the current review, most mapping functions were estimated using more than one regression method (n=142, 61%). In addition to CLAD, Tobit and GLM which were used in the earlier review, other methods commonly included in this review were two/three part models, beta regressions, fractional logistic regressions and mixture models. EQ-5D measures and SF-6D had the largest number of regression approaches applied when mapping to the HSUVs index which reflects where mapping was used. For EQ-5D-3L, methods to address the distribution of the UK values which have a multi-modal distribution with a large proportion of values at 1 and a gap between 1 and the next value (i.e. 0.883) has also increased the number of approaches used. For example, methods such as ALDVMM have been developed to address the specific nature of the EQ-5D-3L UK value set [128]. These are aimed at replicating not only mean predicted HSUVs but also the distribution of the HSUVs index.. These regression methods can be extended to mapping to EQ-5D HSUVs from other country value sets as well as other measures e.g. Gray et al 2018 [212] uses these approaches for HUI3. The appropriateness of methods will depend on the measure and distribution as well as standard tests applied to assess these methods. For example, SF-6D UK value set has a limited range but does not have the multimodal distribution that is seen in the UK EQ-5D-3L value set, which can impact on model selection.

Response mapping was undertaken using ordered or multinomial probit or logit regressions with 5 (17%) studies reported using this approach in the previous review which was a similar proportion to the current review (16%, n=37). All the studies in the previous review used response mapping to map to EQ-5D-3L. Response mapping has also mainly been undertaken for the EQ-5D measures (n=32) in the current review. An additional 3 mapping functions to SF-6D, one to HUI3 and CHU-9D used response mapping in the current review. Application of response mapping across all measures remains relatively

low with the exception of EQ-5D measures. This reflects both the number of dimensions in the measure and the availability of large samples that cover the range of severity within each dimension.

Most mapping functions (n=192, 82%) in this review reported MAEs or RMSEs to assess the performance of mapping functions. These statistics were supported by examining errors by severity to assess whether there were systematic patterns in the errors (n=92, 39%) or plots of observed HSUVs or predicted HSUVs and errors (n=120, 120%). This is an improvement on the previous review where the pattern of errors was only reported in 2 studies (7%) [29].

Regression coefficient values, which are important if mapping functions are to be applied in external datasets, were reported for most of the mapping functions (96%). Some mapping functions were provided in separate programs to estimate HSUVs which was useful particularly where standard OLS techniques were not used therefore coefficient values could not simply be applied directly to the external dataset. Some papers did not report coefficients results because the authors judged that mapping functions performed poorly based on predictive ability. Although most papers reported coefficients alongside assessment of performance of models, validity of mapping functions will depend on other information such as the mapping dataset and how it compares to the dataset where mapping functions will be applied. For example, Woodcock et al [110] found that there were differences in which mapping function was preferred when they were applied to different d.

Although this review provides a useful resource for analysts, there are some limitations. There was no quality assessment of the included studies, and no judgement of whether regression methods, model specifications or predictive ability were appropriate. Any judgements on quality cannot be generalised as the appropriateness of mapping functions relies on assessment of applicability for the context while appropriateness of methods relies on the target PBM. Though the aim of this review was not to examine whether published studies are in accordance with published recommendations including recent ISPOR Taskforce guidelines [27], our extracted data enables a general assessment of whether important information is reported. Our findings suggest that since the 2010 review [29], authors have increased the number of regression methods that they use as well as their reporting of predictive ability of mapping functions across different categories of severity. Despite OLS remaining the most commonly used regression method, there has been a wider use of other regression methods that are likely to be more appropriate for the distribution of HSUV data. This is a promising development since the 2010 review as it suggests that researchers are taking into account the distribution of their data.

Where mapping is required to generate HSUVs to inform HTA, the widespread availability of mapping functions estimated using patient populations rather than the general population is likely to mean that the datasets used to estimate mapping functions are most similar to the clinical trial or observational datasets to which they are applied. This may in turn lead to reduced error in mapped HSUVs. In addition, the availability of more published mapping functions may allow their performance to be tested

in separate datasets. For example, a number of mapping functions are available for the EORTC-QLQ-C30 to EQ-5D-3L which allows them to be tested against each other. For example, Woodcock et al 2018 [110] tested 5 existing EORTC QLQ-C30 mapping functions.

Acceptability to generate HSUVs using mapping in HTA submissions to international agencies has increased since the previous review which is likely to have encouraged the proliferation of mapping studies. This acceptability is also reflected in the use of the crosswalk mapping algorithm [40] from EQ-5D-5L to EQ-5D-3L as an official scoring approach by the EuroQoL group. This crosswalk algorithm can potentially be applied to any other studies which already have EQ-5D-3L HSUVs but not EQ-5D-5L utility values as has been done for Poland [42] and Sri-Lanka [41]. As there are many more EQ-5D-3L country value sets than EQ-5D-5L, this offers an interim alternative to generating EQ-5D-5L values.

New developments in models that take into account the distribution of HSUVs and greater understanding of how the use of mapping in economic evaluation impacts on incremental cost effectiveness ratios enable the science to both be better undertaken and better understood. However, there is still room for improvement and cause to apply caution. Authors are encouraged to follow published recommendations around mapping best practice [25-27, 30], including the recent ISPOR Taskforce guidelines [27]. It is recommended that error term distributions, variance and covariance are reporting in mapping studies to enable better understanding of the accuracy of mapped estimates when they are used in cost-effectiveness modelling [26, 27]. It is also recommended that a plot (with values reported in a table) is always included of predicted versus observed HSUVs conditional on the start measure, to enable researchers using the mapping function to understand the potential accuracy of the mapped estimates [27]. It is hoped that the mapping literature will take these recommendations on board over the coming decade.

Mapping is not guaranteed to generate accurate mapped HSUVs. Care must be taken to ensure overlap between the measures that are mapped from and to as if the generic preference-based measure is inappropriate in that patient population mapping to this measure is also inappropriate. Mapping should not be used to avoid including a generic preference-based measure in a trial or key observational study since mapped estimates increase uncertainty and are not preferable to direct administration of the preference-based measure.

This review provides an important resource for researchers enabling the identification of possible mapping studies for use to predict HSUVs for a range of different conditions and for a range of different preference-based measures.

Author Contributions

CM and DR reviewed studies, extracted and analysed the data from the review and wrote the manuscript. SH and AR reviewed studies, extracted data and contributed to the manuscript. RW

undertook the searches and contributed to the manuscript. RA and JB were involved in designing the analysis and contributed to the manuscript.

Compliance with Ethical Standards: JB and CM were supported in the preparation and submission of this paper by the HEOM Theme of the NIHR CLAHRC Yorkshire and Humber. www.clahrc-yh.nihr.ac.uk. The views expressed are those of the author(s), and not necessarily those of the NHS, the NIHR or the Department of Health. CM, DR, SH, AR, RW, RA and JB have no conflicts of interests to declare.

Data Availability: Full data extracted from each of the studies is available in the electronic supplements.

References

1. Herdman, M., et al., *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).* Quality of Life Research, 2011. **20**(10): p. 1727-1736.
2. Rowen, D., et al., *The Role of Condition-Specific Preference-Based Measures in Health Technology Assessment.* PharmacoEconomics, 2017. **35**(S1): p. 33-41.
3. Brooks, R., *EuroQol: the current state of play.* Health Policy, 1996. **37**(1): p. 53-72.
4. Brazier, J., J. Roberts, and M. Deverill, *The estimation of a preference-based measure of health from the SF-36.* Journal of Health Economics, 2002. **21**(2): p. 271-292.
5. Brazier, J.E. and J. Roberts, *The Estimation of a Preference-Based Measure of Health From the SF-12.* Medical Care, 2004. **42**(9): p. 851-859.
6. Torrance, G.W., et al., *Multiattribute Utility Function for a Comprehensive Health Status Classification System.* Medical Care, 1996. **34**(7): p. 702-722.
7. Feeny, D., et al., *Multiattribute and Single-Attribute Utility Functions for the Health Utilities Index Mark 3 System.* Medical Care, 2002. **40**(2): p. 113-128.
8. Richardson, J., et al., *Modelling utility weights for the Assessment of Quality of Life (AQoL)-8D.* Quality of Life Research, 2014. **23**(8): p. 2395-2404.
9. Sintonen, H., *The 15D instrument of health-related quality of life: properties and applications.* Annals of Medicine, 2001. **33**(5): p. 328-336.
10. Kaplan, R.M., *New health promotion indicators: the general health policy model.* Health Promotion International, 1988. **3**(1): p. 35-49.
11. Apajasalo, M., et al., *Quality of life in early adolescence: a sixteendimensional health-related measure (16D).* Quality of Life Research, 1996. **5**(2): p. 205-211.
12. Apajasalo, M., et al., *Quality of life in pre-adolescence: a 17-dimensional health-related measure (17D).* Quality of Life Research, 1996. **5**(6): p. 532-538.
13. Wille, N., et al., *Development of the EQ-5D-Y: a child-friendly version of the EQ-5D.* Quality of life research, 2010. **19**(6): p. 875-886.
14. Moodie, M., et al., *Predicting Time Trade-Off Health State Valuations of Adolescents in Four Pacific Countries Using the Assessment of Quality-of-Life (AQoL-6D) Instrument.* Value in Health, 2010. **13**(8): p. 1014-1027.
15. Kaplan, R.M., J.W. Bush, and C.C. Berry, *Health status: types of validity and the index of well-being.* Health services research, 1976. **11**(4): p. 478.
16. Stevens, K., *Developing a descriptive system for a new preference-based measure of health-related quality of life for children.* Quality of Life Research, 2009. **18**(8): p. 1105-1113.
17. Rowen, D., et al., *International Regulations and Recommendations for Utility Data for Health Technology Assessment.* PharmacoEconomics, 2017. **35**(S1): p. 11-19.
18. National Insitute for Health and Care Excellence, *Guide to the methods of technology appraisal*. 2013, NICE: London.
19. Busschbach, J.J.V., B. Van Hout, and G.A. De Wit, *BIJLAGE 2: QALY en kwaliteit: van leven metingen*. 2016, Zorginstituut Nederland: Diemen.
20. Pharmaceutical Benefits Advisory Committee, *Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee*. 2013, Australian Government Department of Health: Australia:.
21. Haute Autorité de Santé, *Choices in Methods for Economic Evaluation*. 2012, HAS: France.
22. College voor zorgverzekeringen, *Guidelines for Pharmacoeconomic Research: evaluation and actualisation*. 2006, CVZ.
23. CatSalut, *Guia I Recomanacions Per A La Realització I Presentació D'avaluacions Econòmiques I Anàlisis D'impacte Pressupostari De Medicaments En L'àmbit Del Catsalut*. 2014, CatSalut: Catalonia.
24. CADTH, *Guidelines for the economic evaluation of health technologies: Canada*. 2017.
25. Longworth, L. and D. Rowen, *Mapping to Obtain EQ-5D Utility Values for Use in NICE Health Technology Assessments.* Value in Health, 2013. **16**(1): p. 202-210.
26. Ara, R., D. Rowen, and C. Mukuria, *The Use of Mapping to Estimate Health State Utility Values.* PharmacoEconomics, 2017. **35**(S1): p. 57-66.

27. Wailoo, A.J., et al., *Mapping to Estimate Health-State Utility from Non–Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task Force Report.* Value in Health, 2017. **20**(1): p. 18-27.
28. Round, J. and A. Hawton, *Statistical alchemy: conceptual validity and mapping to generate health state utility values.* PharmacoEconomics-open, 2017. **1**(4): p. 233-239.
29. Brazier, J.E., et al., *A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures.* The European Journal of Health Economics, 2010. **11**(2): p. 215-225.
30. Petrou, S., et al., *The MAPS Reporting Statement for Studies Mapping onto Generic Preference-Based Outcome Measures: Explanation and Elaboration.* PharmacoEconomics, 2015. **33**(10): p. 993-1011.
31. Dakin, H., *Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database.* Health and Quality of Life Outcomes, 2013. **11**(1): p. 151.
32. Rees, A., et al., *Development of the Scharr HUD (Health Utilities Database).* Value in Health, 2013. **16**(7): p. A580.
33. Chen, G. and J. Ratcliffe, *A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations.* PharmacoEconomics, 2015. **33**(10): p. 1013-1028.
34. Kwon, J., et al., *A Systematic Review and Meta-analysis of Childhood Health Utilities.* Medical Decision Making, 2017. **38**(3): p. 277-305.
35. Brazier, J., et al., *A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures.* Health Technology Assessment, 2014. **18**(34).
36. Kearns, B., R. Ara, and A.J. Wailoo, *A review of the use of statistical regression models to inform cost effectiveness analyses within the NICE technology appraisals programme*, in *NICE Decision Support Unit*. 2012.
37. Longworth, L. and D. Rowen, *The use of mapping methods to estimate health state utility values*, in *NICE DSU technical support*. 2011: Sheffield: Decision Support Unit, ScHARR, University of Sheffield. p. b4.
38. ISPOR. *ISPOR Scientific Presentations Database*. [cited 2017 18 September]; Available from: http://www.ispor.org/RESEARCH_STUDY_DIGEST/research_index.asp.
39. Group, E. *EuroQoL Group Website*. [cited 2017 18 September]; Available from: http://www.euroqol.org/.
40. van Hout, B., et al., *Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets.* Value in Health, 2012. **15**(5): p. 708-715.
41. Kularatna, S., et al., *Mapping Sri Lankan EQ-5D-3L to EQ-5D-5L Value Sets.* Value in Health Regional Issues, 2017. **12**: p. 20-23.
42. Golicki, D., et al., *Interim EQ-5D-5L Value Set for Poland: First Crosswalk Value Set in Central and Eastern Europe.* Value in Health Regional Issues, 2014. **4**: p. 19-23.
43. Ware, J.E. and C.D. Sherbourne, *The MOS 36-ltem Short-Form Health Survey (SF-36).* Medical Care, 1992. **30**(6): p. 473-483.
44. Norman, R., et al., *Valuing SF-6D Health States Using a Discrete Choice Experiment.* Medical Decision Making, 2013. **34**(6): p. 773-786.
45. Cruz, L.N., et al., *Estimating the SF-6D Value Set for a Population-Based Sample of Brazilians.* Value in Health, 2011. **14**(5): p. S108-S114.
46. Lam, C.L.K., J. Brazier, and S.M. McGhee, *Valuation of the SF-6D Health States Is Feasible, Acceptable, Reliable, and Valid in a Chinese Population.* Value in Health, 2008. **11**(2): p. 295-303.
47. Brazier, J.E., et al., *Estimating a preference-based index from the Japanese SF-36.* Journal of Clinical Epidemiology, 2009. **62**(12): p. 1323-1331.
48. Ferreira, L.N., et al., *A Portuguese Value Set for the SF-6D.* Value in Health, 2010. **13**(5): p. 624-630.
49. Abellán Perpiñán, J.M., et al., *Lowering The 'Floor' of the SF-6D Scoring Algorithm Using A Lottery Equivalent Method.* Health Economics, 2011. **21**(11): p. 1271-1285.
50. McCabe, C., et al., *Health state values for the HUI 2 descriptive system: results from a UK survey.* Health Economics, 2005. **14**(3): p. 231-244.
51. Le Galès, C., et al., *Development of a preference-weighted health status classification system in France: the Health Utilities Index 3.* Health Care Management Science, 2002. **5**(1): p. 41-51.

52. Ruiz, M., et al., *Adaptación y validación del Health Utilities Index Mark 3 al castellano y baremos de corrección en la población española.* Medicina Clínica, 2003. **120**(3): p. 89-96.
53. Michel, Y.A., et al., *A Norwegian 15D value algorithm: proposing a new procedure to estimate 15D value algorithms.* Quality of Life Research, 2018.
54. Hawthorne, G., J. Richardson, and R. Osborne, *The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life.* Quality of Life Research, 1999. **8**(3): p. 209-224.
55. Richardson, J.R.J., et al., *Construction of the descriptive system for the assessment of quality of life AQoL-6D utility instrument.* Health and Quality of Life Outcomes, 2012. **10**(1): p. 38.
56. Kaplan, R.M., W.J. Sieber, and T.G. Ganiats, *The quality of well-being scale: comparison of the interviewer-administered version with a self-administered questionnaire.* Psychology and Health, 1997. **12**(6): p. 783-791.
57. Seiber, W.J., et al., *Quality of well being self-administered (QWB-SA) scale*, in *San Diego: Health Services Research Center, University of California*. 2008.
58. Stevens, K., *Valuation of the Child Health Utility 9D Index.* PharmacoEconomics, 2012. **30**(8): p. 729-747.
59. Ratcliffe, J., et al., *Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best–Worst Scaling.* Health Economics, 2016. **25**(4): p. 486-496.
60. Ratcliffe, J., et al., *Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm.* Social Science & Medicine, 2016. **157**: p. 48-59.
61. Chen, G., et al., *Scoring the Child Health Utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff.* Quality of Life Research, 2018: p. 1-14.
62. Chen, G., et al., *Mapping between 6 Multiattribute Utility Instruments.* Medical Decision Making, 2015. **36**(2): p. 160-175.
63. Richardson, J., et al., *Comparing and Explaining Differences in the Magnitude, Content, and Sensitivity of Utilities Predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D Multiattribute Utility Instruments.* Medical Decision Making, 2014. **35**(3): p. 276-291.
64. Mihalopoulos, C., et al., *Assessing outcomes for cost-utility analysis in depression: comparison of five multi-attribute utility instruments with two depression-specific outcome measures.* British Journal of Psychiatry, 2014. **205**(05): p. 390-397.
65. Chen, G., et al., *Diabetes and quality of life: Comparing results from utility instruments and Diabetes-39.* Diabetes Research and Clinical Practice, 2015. **109**(2): p. 326-333.
66. Bergius, S., et al., *Health-related quality of life among prostate cancer patients: real-life situation at the beginning of treatment.* Scandinavian Journal of Urology, 2016. **51**(1): p. 13-19.
67. Collado-Mateo, D., et al., *"Fibromyalgia and quality of life: mapping the revised fibromyalgia impact questionnaire to the preference-based instruments".* Health and Quality of Life Outcomes, 2017. **15**(1).
68. Chen, G., et al., *Deriving health utilities from the MacNew Heart Disease Quality of Life Questionnaire.* European Journal of Cardiovascular Nursing, 2014. **14**(5): p. 405-415.
69. Kaambwa, B., et al., *Mapping Between the Sydney Asthma Quality of Life Questionnaire (AQLQ-S) and Five Multi-Attribute Utility Instruments (MAUIs).* PharmacoEconomics, 2016. **35**(1): p. 111-124.
70. Mortimer, D., L. Segal, and J. Sturm, *Can we derive an 'exchange rate' between descriptive and preference-based outcome measures for stroke? Results from the transfer to utility (TTU) technique.* Health and Quality of Life Outcomes, 2009. **7**(1): p. 33.
71. Mortimer, D., et al., *Item-Based versus Subscale-Based Mappings from the SF-36 to a Preference-Based Quality of Life Measure.* Value in Health, 2007. **10**(5): p. 398-407.
72. Ackerman, I.N., et al., *Using WOMAC Index scores and personal characteristics to estimate Assessment of Quality of Life utility scores in people with hip and knee joint disease.* Quality of Life Research, 2014. **23**(8): p. 2365-2374.
73. Kontodimopoulos, N., et al., *Mapping the Cancer-Specific EORTC QLQ-C30 to the Preference-Based EQ-5D, SF-6D, and 15D Instruments.* Value in Health, 2009. **12**(8): p. 1151-1157.
74. Chen, G., et al., *Mapping of Incontinence Quality of Life (I-QOL) scores to Assessment of Quality of Life 8D (AQoL-8D) utilities in patients with idiopathic overactive bladder.* Health and Quality of Life Outcomes, 2014. **12**(1).
75. Chen, G., et al., *From KIDSCREEN-10 to CHU9D: creating a unique mapping algorithm for application in economic evaluation.* Health & Quality of Life Outcomes, 2014. **12**: p. 134.

76. Lambe, T., et al., *Mapping the Paediatric Quality of Life Inventory (PedsQLTM) Generic Core Scales onto the Child Health Utility Index-9 Dimension (CHU-9D) Score for Economic Evaluation in Children.* Pharmacoeconomics, 2018. **36**(4): p. 451-465.
77. Furber, G., et al., *Mapping scores from the Strengths and Difficulties Questionnaire (SDQ) to preference-based utility values.* Quality of Life Research, 2014. **23**(2): p. 403-11.
78. Robinson, T. and Y. Oluboyede, *Estimating CHU-9D utility scores from the WAItE: a mapping algorithm for economic evaluation.* Value in Health, 2018.
79. Hua, A.Y., et al., *Mapping functions in health-related quality of life: mapping from the Achilles Tendon Rupture Score to the EQ-5D.* Knee Surgery Sports Traumatology Arthroscopy, 2018. **26**(10): p. 3083-3088.
80. Badia, X., et al., *Mapping AcroQoL scores to EQ-5D to obtain utility values for patients with acromegaly.* Journal of Medical Economics, 2018. **21**(4): p. 382-389.
81. Bafus, B.T., *Evaluation of utility in shoulder pathology: Correlating the American Shoulder and Elbow Surgeons and Constant scores to the EuroQoL.* World Journal of Orthopedics, 2012. **3**(3): p. 20.
82. Mlcoch, T., et al., *Mapping the relationship between clinical and quality-of-life outcomes in patients with ankylosing spondylitis.* Expert Review of Pharmacoeconomics & Outcomes Research, 2016. **17**(2): p. 203-211.
83. Kaambwa, B., L. Billingham, and S. Bryan, *Mapping utility scores from the Barthel index.* The European Journal of Health Economics, 2011. **14**(2): p. 231-241.
84. Wailoo, A., et al., *Modeling Health State Utility Values in Ankylosing Spondylitis: Comparisons of Direct and Indirect Methods.* Value in Health, 2015. **18**(4): p. 425-431.
85. Grochtdreis, T., et al., *Mapping the Beck Depression Inventory to the Eq-5d-3l in Patients with Depressive Disorders.* Value in Health, 2015. **18**(7): p. A707.
86. Yousefi, M., et al., *Mapping Catquest Scores onto EQ-5D Utility Values in Patients With Cataract Disease.* Iranian Red Crescent Medical Journal, 2016. **19**(5).
87. Boland, M.R.S., et al., *Mapping the Clinical Chronic Obstructive Pulmonary Disease Questionnaire onto Generic Preference-Based EQ-5D Values.* Value in Health, 2015. **18**(2): p. 299-307.
88. Hoyle, C.K., M. Tabberer, and J. Brooks, *Mapping the COPD Assessment Test onto EQ-5D.* Value in Health, 2016. **19**(4): p. 469-477.
89. Badia, X., et al., *Mapping CushingQOL scores to EQ-5D utility values using data from the European Registry on Cushing's syndrome (ERCUSYN).* Quality of Life Research, 2013. **22**(10): p. 2941-2950.
90. Acaster, S., et al., *Mapping the EQ-5D index from the cystic fibrosis questionnaire-revised using multiple modelling approaches.* Health and Quality of Life Outcomes, 2015. **13**(1): p. 33.
91. Poole, C.D., et al., *Estimation of health-related utility (EQ-5D index) in subjects with seasonal allergic rhinoconjunctivitis to evaluate health gain associated with sublingual grass allergen immunotherapy.* Health and Quality of Life Outcomes, 2014. **12**(1): p. 99.
92. Ali, F.M., et al., *Mapping of the DLQI scores to EQ-5D utility values using ordinal logistic regression.* Quality of Life Research, 2017. **26**(11): p. 3025-3034.
93. Blome, C., et al., *Mapping DLQI on EQ-5D in psoriasis: transformation of skin-specific health-related quality of life into utilities.* Archives of Dermatological Research, 2012. **305**(3): p. 197-204.
94. Davison, N.J., et al., *Generating EQ-5D-3L Utility Scores from the Dermatology Life Quality Index: A Mapping Study in Patients with Psoriasis.* Value in Health, 2018. **21**(8): p. 1010-1018.
95. Herédi, E., et al., *Exploring the relationship between EQ-5D, DLQI and PASI, and mapping EQ-5D utilities: a cross-sectional study in psoriasis from Hungary.* The European Journal of Health Economics, 2014. **15**(S1): p. 111-119.
96. Norlin, J.M., et al., *Analysis of three outcome measures in moderate to severe psoriasis: a registry-based study of 2450 patients.* British Journal of Dermatology, 2012. **166**(4): p. 797-802.
97. Mlcoch, T., et al., *Mapping Quality of Life (EQ-5D) from DAPsA, Clinical DAPsA and HAQ in Psoriatic Arthritis.* Patient-Patient Centered Outcomes Research, 2018. **11**(3): p. 329-340.
98. Crott, R. and A. Briggs, *Mapping the QLQ-C30 quality of life cancer questionnaire to EQ-5D patient preferences.* The European Journal of Health Economics, 2010. **11**(4): p. 427-434.
99. Jang, R.W., et al., *Derivation of Utility Values from European Organization for Research and Treatment of Cancer Quality of Life-Core 30 Questionnaire Values in Lung Cancer.* Journal of Thoracic Oncology, 2010. **5**(12): p. 1953-1957.

100. Khan, I. and S. Morris, *A non-linear beta-binomial regression model for mapping EORTC QLQ- C30 to the EQ-5D-3L in lung cancer patients: a comparison with existing approaches.* Health and Quality of Life Outcomes, 2014. **12**(1).
101. Kharroubi, S.A., et al., *Use of Bayesian Markov Chain Monte Carlo Methods to Estimate EQ-5D Utility Scores from EORTC QLQ Data in Myeloma for Use in Cost-Effectiveness Analysis.* Medical Decision Making, 2015. **35**(3): p. 351-360.
102. Kharroubi, S.A., et al., *Bayesian statistical models to estimate EQ-5D utility scores from EORTC QLQ data in myeloma.* Pharmaceutical Statistics, 2018. **17**(4): p. 358-371.
103. Kim, E.-j., S.-K. Ko, and H.-Y. Kang, *Mapping the cancer-specific EORTC QLQ-C30 and EORTC QLQ-BR23 to the generic EQ-5D in metastatic breast cancer patients.* Quality of Life Research, 2011. **21**(7): p. 1193-1203.
104. Kim, S.H., et al., *Mapping EORTC QLQ-C30 onto EQ-5D for the assessment of cancer patients.* Health and Quality of Life Outcomes, 2012. **10**(1): p. 151.
105. Marriott, E.-R., et al., *Mapping EORTC-QLQ-C30 to EQ-5D-3L in patients with colorectal cancer.* Journal of Medical Economics, 2016. **20**(2): p. 193-199.
106. McKenzie, L. and M. van der Pol, *Mapping the EORTC QLQ C-30 onto the EQ-5D Instrument: The Potential to Estimate QALYs without Generic Preference Data.* Value in Health, 2009. **12**(1): p. 167-171.
107. Proskorovsky, I., et al., *Mapping EORTC QLQ-C30 and QLQ-MY20 to EQ-5D in patients with multiple myeloma.* Health and Quality of Life Outcomes, 2014. **12**(1): p. 35.
108. Versteegh, M.M., et al., *Mapping onto Eq-5 D for patients in poor health.* Health and Quality of Life Outcomes, 2010. **8**(1): p. 141.
109. Versteegh, M.M., et al., *Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D.* Medical Decision Making, 2011. **32**(4): p. 554-568.
110. Woodcock, F. and B. Doble, *Mapping the EORTC-QLQ-C30 to the EQ-5D-3L: An Assessment of Existing and Newly Developed Algorithms.* Med Decis Making, 2018. **38**(8): p. 954-967.
111. Burge, R., et al., *Use of health-related quality of life measures to predict health utility in postmenopausal osteoporotic women: results from the Multiple Outcomes of Raloxifene Evaluation study.* Health and Quality of Life Outcomes, 2013. **11**(1): p. 189.
112. Huamán, J.W., et al., *Cutoff values of the Inflammatory Bowel Disease Questionnaire to predict a normal health related quality of life.* Journal of Crohn's and Colitis, 2010. **4**(6): p. 637-641.
113. Cheung, Y.B., et al., *Mapping the Functional Assessment of Cancer Therapy - Breast (FACT-B) to the 5-level EuroQoL group's 5-dimension questionnaire (EQ-5D-5L) utility index in a Multi-ethnic Asian population.* Health and Quality of Life Outcomes, 2014. **12**(1).
114. Cheung, Y.-B., et al., *Mapping the English and Chinese Versions of the Functional Assessment of Cancer Therapy–General to the EQ-5D Utility Index.* Value in Health, 2009. **12**(2): p. 371-376.
115. Teckle, P., et al., *Mapping the FACT-G cancer-specific quality of life instrument to the EQ-5D and SF-6D.* Health and Quality of Life Outcomes, 2013. **11**(1): p. 203.
116. Young, T.A., et al., *Mapping Functions in Health-Related Quality of Life.* Medical Decision Making, 2015. **35**(7): p. 912-926.
117. Askew, R.L., et al., *Mapping FACT-Melanoma Quality-of-Life Scores to EQ-5D Health Utility Weights.* Value in Health, 2011. **14**(6): p. 900-906.
118. Diels, J., et al., *Mapping FACT-P to EQ-5D in a large cross-sectional study of metastatic castration-resistant prostate cancer patients.* Quality of Life Research, 2014. **24**(3): p. 591-598.
119. Skaltsa, K., et al., *Mapping the FACT-P to the Preference-Based EQ-5D Questionnaire in Metastatic Castration-Resistant Prostate Cancer.* Value in Health, 2014. **17**(2): p. 238-244.
120. Wu, E.Q., et al., *Mapping FACT-P and EORTC QLQ-C30 to Patient Health Status Measured by EQ-5D in Metastatic Hormone-Refractory Prostate Cancer Patients.* Value in Health, 2007. **10**(5): p. 408-414.
121. Monroy, M., et al., *Mapping of the Gastrointestinal Short Form Questionnaire (GSF-Q) into EQ-5D-3L and SF-6D in patients with gastroesophageal reflux disease.* Health and Quality of Life Outcomes, 2018. **16**.
122. Ara, R., et al., *Predicting preference-based utility values using partial proportional odds models.* BMC Research Notes, 2014. **7**(1): p. 438.
123. Lindkvist, M. and I. Feldman, *Assessing outcomes for cost-utility analysis in mental health interventions: mapping mental health specific outcome measure GHQ-12 onto EQ-5D-3L.* Health and Quality of Life Outcomes, 2016. **14**(1).

124. Serrano-Aguilar, P., et al., *The relationship among mental health status (GHQ-12), health related quality of life (EQ-5D) and health-state utilities in a general population.* Epidemiology and Psychiatric Sciences, 2009. **18**(3): p. 229-239.
125. Ward Fuller, G., et al., *Health State Preference Weights for the Glasgow Outcome Scale Following Traumatic Brain Injury: A Systematic Review and Mapping Study.* Value in Health, 2017. **20**(1): p. 141-151.
126. Gillard, P.J., et al., *Mapping from Disease-Specific Measures to Health-State Utility Values in Individuals with Migraine.* Value in Health, 2012. **15**(3): p. 485-494.
127. Bansback, N., et al., *Using the health assessment questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis.* Arthritis & Rheumatism, 2007. **57**(6): p. 963-971.
128. Hernández Alava, M., A.J. Wailoo, and R. Ara, *Tails from the Peak District: Adjusted Limited Dependent Variable Mixture Models of EQ-5D Questionnaire Health State Utility Values.* Value in Health, 2012. **15**(3): p. 550-561.
129. Wolfe, F., K. Michaud, and G. Wallenstein, *Scale Characteristics and Mapping Accuracy of the US EQ-5D, UK EQ-5D, and SF-6D in Patients with Rheumatoid Arthritis.* The Journal of Rheumatology, 2010. **37**(8): p. 1615-1625.
130. Nair, S.C., et al., *Does disease activity add to functional disability in estimation of utility for rheumatoid arthritis patients on biologic treatment?* Rheumatology, 2015. **55**(1): p. 94-102.
131. Hernández Alava, M., et al., *The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis.* Rheumatology, 2013. **52**(5): p. 944-950.
132. Adams, R., et al., *Understanding the Relationship between the EQ-5D, SF-6D, HAQ and Disease Activity in Inflammatory Arthritis.* PharmacoEconomics, 2010. **28**(6): p. 477-487.
133. Carreño, A., et al., *Using HAQ-DI to estimate HUI-3 and EQ-5D utility values for patients with rheumatoid arthritis in Spain.* Value in Health, 2011. **14**(1): p. 192-200.
134. Kim, H.-L., et al., *Mapping health assessment questionnaire disability index (HAQ-DI) score, pain visual analog scale (VAS), and disease activity score in 28 joints (DAS28) onto the EuroQol-5D (EQ-5D) utility score with the KORean Observational study Network for Arthritis (KORONA) registry data.* Rheumatology International, 2016. **36**(4): p. 505-513.
135. Jia, H., et al., *Predicting the EuroQol Group's EQ-5D Index from CDC's "Healthy Days" in a US Sample.* Medical Decision Making, 2010. **31**(1): p. 174-185.
136. Kay, S., et al., *Mapping EQ-5D Utility Scores from the Incontinence Quality of Life Questionnaire among Patients with Neurogenic and Idiopathic Overactive Bladder.* Value in Health, 2013. **16**(2): p. 394-402.
137. Buxton, M.J., et al., *Mapping from Disease-Specific Measures to Utility: An Analysis of the Relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's Disease and Measures of Utility.* Value in Health, 2007. **10**(3): p. 214-220.
138. Gu, N.Y., et al., *Mapping of the Insomnia Severity Index and other sleep measures to EuroQol EQ-5D health state utilities.* Health and Quality of Life Outcomes, 2011. **9**(1): p. 119.
139. Dixon, P., H. Dakin, and S. Wordsworth, *Generic and disease-specific estimates of quality of life in macular degeneration: mapping the MacDQoL onto the EQ-5D-3L.* Quality of Life Research, 2015. **25**(4): p. 935-945.
140. Vera, E., et al., *Relationship between symptom burden and health status: Analysis of the MDASI-BT and EQ-5D.* Neuro-Oncology Practice, 2018. **5**(1): p. 56-63.
141. Huang, I.C., et al., *Addressing Ceiling Effects in Health Status Measures: A Comparison of Techniques Applied to Measures for People with HIV Disease.* Health Services Research, 2007. **43**(1p1): p. 327-339.
142. Joyce, V.R., et al., *Mapping MOS-HIV to HUI3 and EQ-5D-3L in Patients With HIV.* MDM Policy & Practice, 2017. **2**(2): p. 238146831771644.
143. Ali, M., et al., *Dependency and health utilities in stroke: Data to inform cost-effectiveness analyses.* European Stroke Journal, 2016. **2**(1): p. 70-76.
144. Rivero-Arias, O., et al., *Mapping the Modified Rankin Scale (mRS) Measurement into the Generic EuroQol (EQ-5D) Health Outcome.* Medical Decision Making, 2009. **30**(3): p. 341-354.
145. Whynes, D.K., et al., *Testing for Differential Item Functioning within the EQ-5D.* Medical Decision Making, 2012. **33**(2): p. 252-260.

146. Kontodimopoulos, N., et al., *Longitudinal predictive ability of mapping models: examining post-intervention EQ-5D utilities derived from baseline MHAQ data in rheumatoid arthritis patients.* The European Journal of Health Economics, 2012. **14**(2): p. 307-314.
147. Sauerland, S., et al., *Mapping Utility Scores from a Disease-Specific Quality-of-Life Measure in Bariatric Surgery Patients.* Value in Health, 2009. **12**(2): p. 364-370.
148. Hawton, A., et al., *Using the Multiple Sclerosis Impact Scale to Estimate Health State Utility Values: Mapping from the MSIS-29, Version 2, to the EQ-5D and the SF-6D.* Value in Health, 2012. **15**(8): p. 1084-1091.
149. Hawton, A., et al., *The use of multiple sclerosis condition-specific measures to inform health policy decision-making: mapping from the MSWS-12 to the EQ-5D.* Multiple Sclerosis Journal, 2011. **18**(6): p. 853-861.
150. Sidovar, M.F., et al., *Mapping the 12-item multiple sclerosis walking scale to the EuroQol 5-dimension index measure in North American multiple sclerosis patients.* BMJ Open, 2013. **3**(5): p. e002798.
151. Kay, S. and A. Ferreira, *Mapping the 25-item National Eye Institute Visual Functioning Questionnaire (NEI VFQ-25) to EQ-5D Utility Scores.* Ophthalmic Epidemiology, 2014. **21**(2): p. 66-78.
152. Browne, C., et al., *Estimating quality-adjusted life years from patient-reported visual functioning.* Eye, 2012. **26**(10): p. 1295-1301.
153. Payakachat, N., et al., *Predicting EQ-5D utility scores from the 25-item National Eye Institute Vision Function Questionnaire (NEI-VFQ 25) in patients with age-related macular degeneration.* Quality of Life Research, 2009. **18**(7): p. 801-813.
154. Carreon, L.Y., et al., *Estimating EQ-5D values from the Neck Disability Index and numeric rating scales for neck and arm pain.* Journal of Neurosurgery: Spine, 2014. **21**(3): p. 394-399.
155. Vokó, Z., et al., *Mapping the Nottingham Health Profile onto the Preference-Based EuroQol-5D Instrument for Patients with Diabetes.* Value in Health Regional Issues, 2014. **4**: p. 31-36.
156. McDonough, C.M., et al., *Predicting EQ-5D-US and SF-6D societal health state values from the Osteoporosis Assessment Questionnaire.* Osteoporosis International, 2011. **23**(2): p. 723-732.
157. Ruiz, M.A., et al., *Mapping of the OAB-SF Questionnaire onto EQ-5D in Spanish Patients with Overactive Bladder.* Clinical Drug Investigation, 2016. **36**(4): p. 267-279.
158. Hernandez Alava, M. and A. Wailoo, *Fitting adjusted limited dependent variable mixture models to EQ-5D.* Stata Journal, 2015. **15**(3): p. 737-750.
159. Oppe, M., N. Devlin, and N. Black, *Comparison of the Underlying Constructs of the EQ-5D and Oxford Hip Score: Implications for Mapping.* Value in Health, 2011. **14**(6): p. 884-891.
160. Pinedo-Villanueva, R.A., et al., *Mapping the Oxford hip score onto the EQ-5D utility index.* Quality of Life Research, 2012. **22**(3): p. 665-675.
161. Dakin, H., A. Gray, and D. Murray, *Mapping analyses to estimate EQ-5D utilities and responses based on Oxford Knee Score.* Quality of Life Research, 2012. **22**(3): p. 683-694.
162. Cappelleri, J.C., et al., *Mapping painDETECT, a neuropathic pain screening tool, to the EuroQol (EQ-5D-3L).* Quality of Life Research, 2016. **26**(2): p. 467-477.
163. Dzingina, M.D., P. McCrone, and I.J. Higginson, *Does the EQ-5D capture the concerns measured by the Palliative care Outcome Scale? Mapping the Palliative care Outcome Scale onto the EQ-5D using statistical methods.* Palliative Medicine, 2017. **31**(8): p. 716-725.
164. Borchani, H., et al., *Markov blanket-based approach for learning multi-dimensional Bayesian network classifiers: An application to predict the European Quality of Life-5 Dimensions (EQ-5D) from the 39-item Parkinson's Disease Questionnaire (PDQ-39).* Journal of Biomedical Informatics, 2012. **45**(6): p. 1175-1184.
165. Kent, S., et al., *Mapping from the Parkinson's Disease Questionnaire PDQ-39 to the Generic EuroQol EQ-5D-3L.* Medical Decision Making, 2015. **35**(7): p. 902-911.
166. Young, M.K., et al., *Mapping of the PDQ-39 to EQ-5D scores in patients with Parkinson's disease.* Quality of Life Research, 2012. **22**(5): p. 1065-1072.
167. Cheung, Y.B., et al., *Mapping the eight-item Parkinson's Disease Questionnaire (PDQ-8) to the EQ-5D utility index.* Quality of Life Research, 2008. **17**(9): p. 1173-1181.
168. Hatswell, A.J. and S. Vegter, *Measuring quality of life in opioid-induced constipation: mapping EQ-5D-3 L and PAC-QOL.* Health Economics Review, 2016. **6**(1).

169. Hartman, J.D. and B.M. Craig, *Comparing and transforming PROMIS utility values to the EQ-5D.* Quality of Life Research, 2018. **27**(3): p. 725-733.
170. Revicki, D.A., et al., *Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample.* Quality of Life Research, 2009. **18**(6): p. 783-791.
171. Thompson, N.R., B.R. Lapin, and I.L. Katzan, *Mapping PROMIS Global Health Items to EuroQol (EQ-5D) Utility Scores Using Linear and Equipercentile Equating.* PharmacoEconomics, 2017. **35**(11): p. 1167-1176.
172. Park, S.-Y., et al., *Development of a transformation model to derive general population-based utility: Mapping the pruritus-visual analog scale (VAS) to the EQ-5D utility.* Journal of Evaluation in Clinical Practice, 2017. **23**(4): p. 755-761.
173. Geale, K., M. Henriksson, and M. Schmitt-Egenolf, *How is disease severity associated with quality of life in psoriasis patients? Evidence from a longitudinal population-based study in Sweden.* Health and Quality of Life Outcomes, 2017. **15**(1).
174. Kołtowska-Häggström, M., et al., *Using EQ-5D to Derive General Population-based Utilities for the Quality of Life Assessment of Growth Hormone Deficiency in Adults (QoL-AGHDA).* Value in Health, 2007. **10**(1): p. 73-81.
175. Khan, K.A., et al., *Mapping between the Roland Morris Questionnaire and Generic Preference-Based Measures.* Value in Health, 2014. **17**(6): p. 686-695.
176. Rundell, S.D., et al., *Mapping a Patient-Reported Functional Outcome Measure to a Utility Measure for Comparative Effectiveness and Economic Evaluations in Older Adults with Low Back Pain.* Medical Decision Making, 2014. **34**(7): p. 873-883.
177. Wong, C.K.H., et al., *Mapping the SRS-22r questionnaire onto the EQ-5D-5L utility score in patients with adolescent idiopathic scoliosis.* PLOS ONE, 2017. **12**(4): p. e0175847.
178. Goldsmith, K.A., et al., *Mapping of the EQ-5D index from clinical outcome measures and demographic variables in patients with coronary heart disease.* Health and Quality of Life Outcomes, 2010. **8**(1): p. 54.
179. Wijeysundera, H.C., et al., *Predicting EQ-5D Utility Scores from the Seattle Angina Questionnaire in Coronary Artery Disease.* Medical Decision Making, 2010. **31**(3): p. 481-493.
180. Chuang, L.-H. and P. Kind, *Converting the SF-12 into the EQ-5D.* PharmacoEconomics, 2009. **27**(6): p. 491-505.
181. Coca Perraillon, M., Y.-C.T. Shih, and R.A. Thisted, *Predicting the EQ-5D-3L Preference Index from the SF-12 Health Survey in a National US Sample.* Medical Decision Making, 2015. **35**(7): p. 888-901.
182. Conigliani, C., A. Manca, and A. Tancredi, *Prediction of patient-reported outcome measures via multivariate ordered probit models.* Journal of the Royal Statistical Society: Series A (Statistics in Society), 2014. **178**(3): p. 567-591.
183. Le, Q.A. and J.N. Doctor, *Probabilistic Mapping of Descriptive Health Status Responses Onto Health State Utilities Using Bayesian Networks.* Medical Care, 2011. **49**(5): p. 451-460.
184. Le, Q.A., *Probabilistic mapping of the health status measure SF-12 onto the health utility measure EQ-5D using the US-population-based scoring models.* Quality of Life Research, 2013. **23**(2): p. 459-466.
185. Ara, R. and J. Brazier, *Deriving an Algorithm to Convert the Eight Mean SF-36 Dimension Scores into a Mean EQ-5D Preference-Based Score from Published Studies (Where Patient Level Data Are Not Available).* Value in Health, 2008. **11**(7): p. 1131-1143.
186. Kim, S.-H., et al., *Deriving a mapping algorithm for converting SF-36 scores to EQ-5D utility score in a Korean population.* Health and Quality of Life Outcomes, 2014. **12**(1).
187. Rowen, D., J. Brazier, and J. Roberts, *Mapping SF-36 onto the EQ-5D index: how reliable is the relationship?* Health and Quality of Life Outcomes, 2009. **7**(1): p. 27.
188. Neilson, A.R., et al., *Estimating EQ-5D Weights from other Trial Outcome Measures for Use in Cost-Effectiveness Studies: An Example in Patients with Frozen Shoulder.* Shoulder & Elbow, 2013. **5**(2): p. 136-143.
189. Crump, R.T., et al., *Establishing utility values for the 22-item Sino-Nasal Outcome Test (SNOT-22) using a crosswalk to the EuroQol-five-dimensional questionnaire-three-level version (EQ-5D-3L).* International Forum of Allergy & Rhinology, 2017. **7**(5): p. 480-487.
190. Starkie, H.J., et al., *Predicting EQ-5D Values Using the SGRQ.* Value in Health, 2011. **14**(2): p. 354-360.

191. Busschbach, J.J.V., et al., *Deriving reference values and utilities for the QoL-AGHDA in adult GHD.* The European Journal of Health Economics, 2010. **12**(3): p. 243-252.
192. Poole, C.D., et al., *A comparison of physician-rated disease severity and patient reported outcomes in mild to moderately active ulcerative colitis.* Journal of Crohn's and Colitis, 2010. **4**(3): p. 275-282.
193. Dams, J., et al., *Mapping the EQ-5D index by UPDRS and PDQ-8 in patients with Parkinson's disease.* Health and Quality of Life Outcomes, 2013. **11**(1): p. 35.
194. Chan, K., et al., *Underestimation of uncertainties in health utilities dervied from mapping algorithms involving health-related quality of life measures: statistical explanations and potential remedies.* Value in Health, 2013. **16**(3): p. A49.
195. Barton, G.R., et al., *Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores?* Health and Quality of Life Outcomes, 2008. **6**(1): p. 51.
196. Wailoo, A., M. Hernandez Alava, and A. Escobar Martinez, *Modelling the relationship between the WOMAC osteoarthritis index and EQ-5D.* Health and Quality of Life Outcomes, 2014. **12**(1): p. 37.
197. Xie, F., et al., *Use of a Disease-Specific Instrument in Economic Evaluations: Mapping WOMAC onto the EQ-5D Utility Index.* Value in Health, 2010. **13**(8): p. 873-878.
198. Parker, M., A. Haycox, and J. Graves, *Estimating the Relationship between Preference-Based Generic Utility Instruments and Disease-Specific Quality-of-Life Measures in Severe Chronic Constipation.* PharmacoEconomics, 2011. **29**(8): p. 719-730.
199. Moore, A., C.A. Young, and D.A. Hughes, *Mapping ALSFRS-R and ALSUI to EQ-5D in Patients with Motor Neuron Disease.* Value Health, 2018. **21**(11): p. 1322-1329.
200. Peak, J., et al., *Predicting health-related quality of life (EQ-5D-5 L) and capability wellbeing (ICECAP-A) in the context of opiate dependence using routine clinical outcome measures: CORE-OM, LDQ and TOP.* Health & Quality of Life Outcomes, 2018. **16**(1): p. 106.
201. Gamst-Klaussen, T., et al., *Assessment of outcome measures for cost–utility analysis in depression: mapping depression scales onto the EQ-5D-5L.* 2018. **4**(4): p. 160-166.
202. Lamu, A.N., et al., *Do country-specific preference weights matter in the choice of mapping algorithms? The case of mapping the Diabetes-39 onto eight country-specific EQ-5D-5L value sets.* Quality of Life Research, 2018. **27**(7): p. 1801-1814.
203. Hernández-Alava, M. and S. Pudney, *Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis.* Journal of Health Economics, 2017. **55**: p. 139-152.
204. Ameri, H., et al., *Mapping the cancer-specific QLQ-C30 onto the generic EQ-5D-5L and SF-6D in colorectal cancer patients.* Expert Review of Pharmacoeconomics and Outcomes Research, 2018.
205. Lamu, A.N. and J.A. Olsen, *Testing alternative regression models to predict utilities: mapping the QLQ-C30 onto the EQ-5D-5L and the SF-6D.* Quality of Life Research, 2018.
206. Lee, C.F., et al., *Development of Conversion Functions Mapping the FACT-B Total Score to the EQ-5D-5L Utility Value by Three Linking Methods and Comparison with the Ordinary Least Square Method.* Applied Health Economics and Health Policy, 2018. **16**(5): p. 685-695.
207. Patton, T., et al., *Mapping between HAQ-DI and EQ-5D-5L in a Chinese patient population.* Qual Life Res, 2018. **27**(11): p. 2815-2822.
208. Coon, C., et al., *Evaluation of a crosswalk between the European Quality of Life Five Dimension Five Level and the Menopause-Specific Quality of Life questionnaire.* Climacteric, 2018: p. 1-8.
209. Kaambwa, B. and J. Ratcliffe, *Predicting EuroQoL 5 Dimensions 5 Levels (EQ-5D-5L) Utilities from Older People's Quality of Life Brief Questionnaire (OPQoL-Brief) Scores.* The Patient - Patient-Centered Outcomes Research, 2017. **11**(1): p. 39-54.
210. Abdin, E., et al., *Mapping the Positive and Negative Syndrome Scale scores to EQ-5D-5L and SF-6D utility scores in patients with schizophrenia.* Qual Life Res, 2018.
211. Wijnen, B.F.M., et al., *A comparison of the responsiveness of EQ-5D-5L and the QOLIE-31P and mapping of QOLIE-31P to EQ-5D-5L in epilepsy.* The European Journal of Health Economics, 2017. **19**(6): p. 861-870.
212. Gray, L.A., M. Hernandez Alava, and A.J. Wailoo, *Development of Methods for the Mapping of Utilities Using Mixture Models: Mapping the AQLQ-S to the EQ-5D-5L and the HUI3 in Patients with Asthma.* Value in Health, 2018. **21**(6): p. 748-757.

213. Kaambwa, B., et al., *Does Selecting Covariates Using Factor Analysis in Mapping Algorithms Improve Predictive Accuracy? A Case of Predicting EQ-5D-5L and SF-6D Utilities from the Women's Health Questionnaire.* 2018.
214. Wee, H.L., et al., *Mean Rank, Equipercentile, and Regression Mapping of World Health Organization Quality of Life Brief (WHOQOL-BREF) to EuroQoL 5 Dimensions 5 Levels (EQ-5D-5L) Utilities.* Medical Decision Making, 2018. **38**(3): p. 319-333.
215. Khan, K.A., et al., *Mapping EQ-5D utility scores from the PedsQL™ generic core scales.* Pharmacoeconomics, 2014. **32**(7): p. 693-706.
216. Dakin, H., et al., *Mapping analyses to estimate health utilities based on responses to the OM8-30 otitis media questionnaire.* Quality of Life Research, 2009. **19**(1): p. 65-80.
217. Cheung, Y.B., et al., *Mapping the Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory to the Health Utility Index Mark III.* Quality of Life Research, 2018.
218. Payakachat, N., et al., *Predicting health utilities for children with autism spectrum disorders.* Autism research : official journal of the International Society for Autism Research, 2014. **7**(6): p. 649-663.
219. Hays, R.D., et al., *Using Linear Equating to Map PROMIS® Global Health Items and the PROMIS-29 V2.0 Profile Measure to the Health Utilities Index Mark 3.* PharmacoEconomics, 2016. **34**(10): p. 1015-1022.
220. Yang, Y., et al., *Improving the mapping of condition-specific health-related quality of life onto SF-6D score.* Quality of Life Research, 2014. **23**(8): p. 2343-2353.
221. Frew, E.J., et al., *Providing an extended use of an otological-specific outcome instrument to derive cost-effectiveness estimates of treatment.* Clinical Otolaryngology, 2015. **40**(6): p. 593-599.
222. Skolasky, R.L., et al., *Predicting Health-Utility Scores From the Cervical Spine Outcomes Questionnaire in a Multicenter Nationwide Study of Anterior Cervical Spine Surgery.* Spine, 2011. **36**(25): p. 2211-2216.
223. Hollingworth, W., et al., *Exploring the impact of changes in neurogenic urinary incontinence frequency and condition-specific quality of life on preference-based outcomes.* Quality of Life Research, 2010. **19**(3): p. 323-331.
224. Kalaitzakis, E., et al., *Mapping chronic liver disease questionnaire scores onto SF-6D utility values in patients with primary sclerosing cholangitis.* Quality of Life Research, 2015. **25**(4): p. 947-957.
225. Stepanova, M., et al., *Prediction of Health Utility Scores in Patients with Chronic Hepatitis C Using the Chronic Liver Disease Questionnaire-Hepatitis C Version (CLDQ-HCV).* Value in Health, 2018. **21**(5): p. 612-621.
226. Roset, M., et al., *Mapping CushingQoL Scores onto SF-6D Utility Values in Patients with Cushing's Syndrome.* The Patient - Patient-Centered Outcomes Research, 2013. **6**(2): p. 103-111.
227. Wong, C.K.H., et al., *Predicting SF-6D from the European Organization for Treatment and Research of Cancer Quality of Life Questionnaire Scores in Patients with Colorectal Cancer.* Value in Health, 2013. **16**(2): p. 373-384.
228. Kontodimopoulos, N., *The potential for a generally applicable mapping model between QLQ-C30 and SF-6D in patients with different cancers: a comparison of regression-based methods.* Quality of Life Research, 2014. **24**(6): p. 1535-1544.
229. Wong, C.K.H., et al., *Mapping the Functional Assessment of Cancer Therapy-General or -Colorectal to SF-6D in Chinese Patients with Colorectal Neoplasm.* Value in Health, 2012. **15**(3): p. 495-503.
230. Lee, L., et al., *Mapping the Gastrointestinal Quality of Life Index to Short-Form 6D utility scores.* Journal of Surgical Research, 2014. **186**(1): p. 135-141.
231. Meacock, R., et al., *Mapping the disease-specific LupusQoL to the SF-6D.* Quality of Life Research, 2014. **24**(7): p. 1749-1758.
232. Yang, M., et al., *Mapping MOS Sleep Scale scores to SF-6D utility index.* Current Medical Research and Opinion, 2007. **23**(9): p. 2269-2282.
233. Richardson, S.S. and S. Berven, *The development of a model for translation of the Neck Disability Index to utility scores for cost-utility analysis in cervical disorders.* The Spine Journal, 2012. **12**(1): p. 55-62.
234. Zheng, Y., et al., *Mapping the neck disability index to SF-6D in patients with chronic neck pain.* Health and Quality of Life Outcomes, 2016. **14**(1).
235. Carreon, L.Y., et al., *Predicting SF-6D Utility Scores From the Neck Disability Index and Numeric Rating Scales for Neck and Arm Pain.* Spine, 2011. **36**(6): p. 490-494.
236. Carreon, L.Y., et al., *Predicting SF-6D Utility Scores From the Oswestry Disability Index and Numeric Rating Scales for Back and Leg Pain.* Spine, 2009. **34**(19): p. 2085-2089.

237. Hanmer, J., *Predicting an SF-6D Preference-Based Score Using MCS and PCS Scores from the SF-12 or SF-36.* Value in Health, 2009. **12**(6): p. 958-966.
238. Wang, P., et al., *Predicting preference-based SF-6D index scores from the SF-8 health survey.* Quality of Life Research, 2012. **22**(7): p. 1675-1683.
239. Selim, A.J., et al., *A preference-based measure of health: the VR-6D derived from the veterans RAND 12-Item Health Survey.* Quality of Life Research, 2011. **20**(8): p. 1337-1347.
240. Aaronson, N.K., et al., *The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology.* JNCI Journal of the National Cancer Institute, 1993. **85**(5): p. 365-376.
241. Cella, D.F., et al., *The Functional Assessment of Cancer Therapy scale: development and validation of the general measure.* Journal of Clinical Oncology, 1993. **11**(3): p. 570-579.
242. Fries, J.F., *The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales.* j Rheumatol, 1982. **9**: p. 789-793.
243. Ware, J.E., M. Kosinski, and S.D. Keller, *A 12-Item Short-Form Health Survey.* Medical Care, 1996. **34**(3): p. 220-233.
244. Hunt, S.M., et al., *The Nottingham health profile: Subjective health status and medical consultations.* Social Science & Medicine. Part A: Medical Psychology & Medical Sociology, 1981. **15**(3): p. 221-229.
245. Cella, D., et al., *The Patient-Reported Outcomes Measurement Information System (PROMIS).* Medical Care, 2007. **45**(Suppl 1): p. S3-S11.
246. Group, W., *Development of the World Health Organization WHOQOL-BREF quality of life assessment.* Psychological Medicine, 1998. **28**(3): p. 551-558.
247. Bowling, A., et al., *A short measure of quality of life in older age: The performance of the brief Older People's Quality of Life questionnaire (OPQOL-brief).* Archives of Gerontology and Geriatrics, 2013. **56**(1): p. 181-187.
248. Girod, I., et al., *Development of a revised version of the Women's Health Questionnaire.* Climacteric, 2006. **9**(1): p. 4-12.
249. Ravens-Sieberer, U., et al., *Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life.* Quality of Life Research, 2010. **19**(10): p. 1487-1500.
250. Varni, J.W., M. Seid, and P.S. Kurtin, *PedsQL™ 4.0: Reliability and Validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in Healthy and Patient Populations.* Medical Care, 2001. **39**(8).
251. Goodman, R., *The Strengths and Difficulties Questionnaire: A Research Note.* Journal of Child Psychology and Psychiatry, 2006. **38**(5): p. 581-586.
252. Oluboyede, Y., C. Hulme, and A. Hill, *Development and refinement of the WAItE: a new obesity-specific quality of life measure for adolescents.* Quality of Life Research, 2017. **26**(8): p. 2025-2039.
253. Richardson, J., A. Iezzi, and A. Maxwell, *Cross-national comparison of twelve quality of life instruments: MIC Paper 1 Background, questions, instruments.* 2012, Centre for Health Economics: Melbourne.
254. Yohai, V.J., *High breakdown-point and high efficiency robust estimates for regression.* The Annals of Statistics, 1987: p. 642-656.
255. Secondary Care Analysis (PROMs) and NHS Digital, *Patient Reported Outcome Measures (PROMS) in England. A guide to PROMs methodology*. 2017, NHS Digital: England.
256. Wolfe, F. and K. Michaud, *The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank.* Rheumatology, 2010. **50**(1): p. 16-24.

Figure 1: PRISMA diagram of identification, screening and inclusion of papers

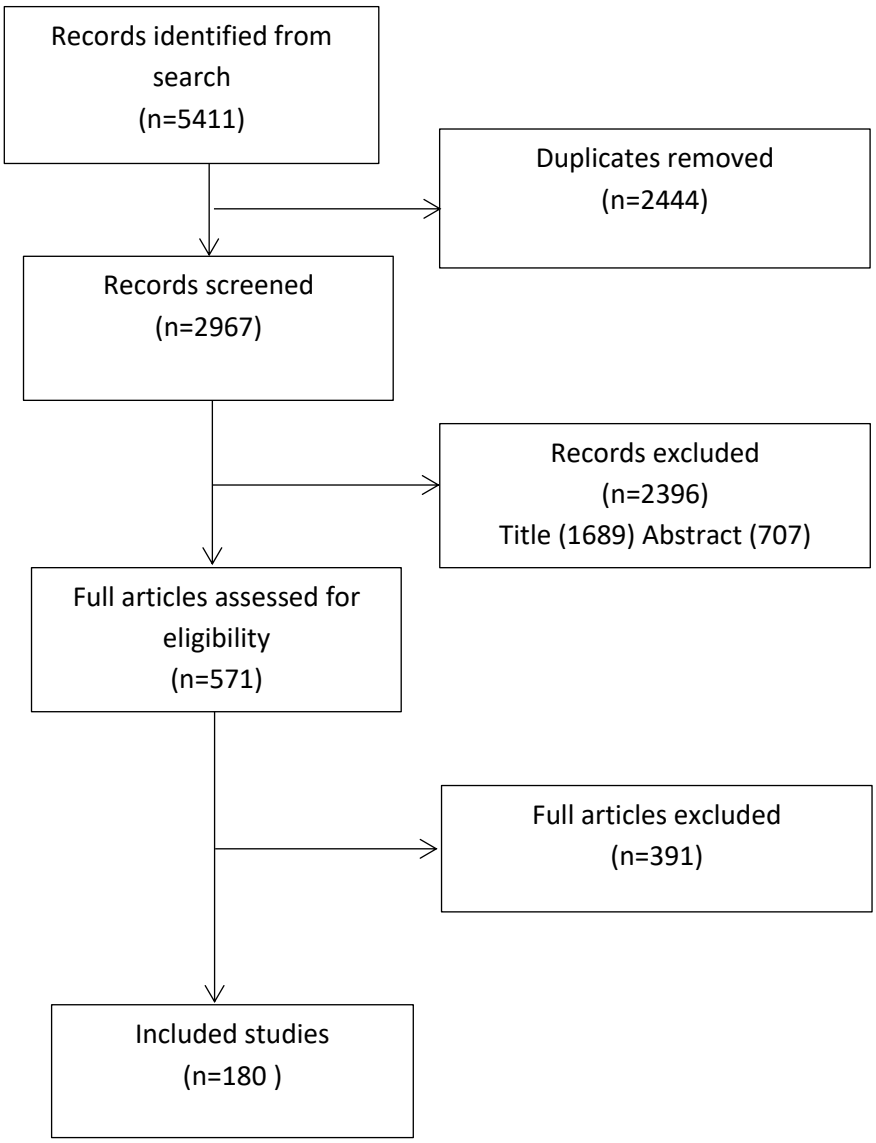

Table 1 Regression methods used by PBM

Target measure	Estimation methods	Number of mapping functions	% within each PBM
15D	Ordinary Least Squares (OLS)	7	88
	Censored Least Absolute Deviation (CLAD)	2	25
	Generalised Linear Model (GLM)	6	75
	Robust Regression MM-estimator	2	25
	Linear Geometric Mean Square Regression	1	13
	Beta regressions	1	13
AQoL-4D/ AQoL-8D	OLS	10	83
	CLAD	2	17
	GLM	8	67
	Beta regression	1	8
	Robust Regression MM-estimator	3	25
	Linear Geometric Mean Square Regression	1	8
CHU-9D	OLS	4	100
	CLAD	2	50
	GLM	3	75
	Two-part model (TPM)	1	25
	Tobit	2	50
	Beta regression	1	25
	Robust Regression MM-estimator	2	50
	Mixture models	1	25
	Response mapping	1	25
EQ-5D-3L/ EQ-5D-5L/ EQ-5D-Y	OLS	131	89
	CLAD	36	24
	GLM	29	20
	TPM	23	16
	Three-part models	2	1
	Tobit	32	22
	Beta regressions	11	7
	Robust Regression MM-estimator	6	4
	Linear Geometric Mean Square Regression	1	1
	Generalized estimating equation	3	2
	Median/quantile/quadratic / equipercentile regression	8	5
	Splining	3	2
	Mixture models	15	10
	Bayesian approaches (OLS/networks/probabilistic)	8	5
	Fractional logistic regression	5	3
	Linear equating	2	1
	Non-parametric	2	1
	Mean ranking	1	1
	Extended estimating equations	1	1
	Response mapping	32	22
HUI2/HUI3	OLS	12	100
	CLAD	2	17
	GLM	5	42

Target measure	Estimation methods	Number of mapping functions	% within each PBM
	Tobit	1	8
	TPM	2	17
	Linear Geometric Mean Square Regression	1	8
	Robust Regression MM-estimator	2	17
	Beta regression finite mixture models	2	17
	Mean Rank	1	8
	Linear equating	1	8
	Mixture models	2	17
	Response mapping	1	8
QWB-SA	OLS	3	75
	CLAD	1	25
	GLM	3	75
	Robust MM-estimator	2	50
	Linear Geometric Mean Square Regression	1	25
SF-6D (SF-12 and SF-36)	OLS	41	91
	CLAD	11	24
	GLM	16	36
	Tobit	8	18
	TPM	1	2
	Robust Regression MM-estimator	3	7
	Beta Regression	4	9
	Mixture models	1	2
	Median and kernel regression	2	4
	Generalised estimating models	1	2
	Linear Geometric Mean Square Regression	1	2
	Bayesian additive regression kernels	1	2
	Fractional logistic	1	2
	Quantile	1	2
	Extended estimating equations	1	2
	Non-parametric	1	2
	Response mapping	3	7
	Not stated and not OLS	1	2

Note: Proportion (%) add up to more than 100% within each measure as multiple estimation methods were used in most studies

OLS – Ordinary Least Squares; CLAD - Censored Least Absolute Deviation; GLM – Generalised Linear Model; TPM – Two-part model; 15D – 15 Dimensions; AQoL-4/8D – Assessment of Quality of Life 4 or 8 Dimensions; CHU-9D – Child Health Utility – 9 Dimensions; EQ-5D-3L – EQ-5D three level version; EQ-5D-5L – EQ-5D five level version; EQ-5D-Y – EQ-5D youth version; HUI2/3 – Health Utilities Index 2 or 3; QWB – Quality of Wellbeing Scale; SF-6D – Short Form – 6 Dimensions; SF-12 – Short Form 12 ; SF-36 – Short Form 36

Table 2 Number of mapping functions including additional variables, coefficients and performance indicators by preference-based measure (n (%))

Target measure	Any tested model includes clinical measures	Any tested model includes Age	Any tested model includes Gender	Are coefficients reported?	Did paper assess predictive ability? Only interested in MAE, RMSE, MSE (not ME)	Error across different categories of severity (can be subgroup means or a plot of predicted versus actual, or a plot of errors)	Plot of predictions e.g. predicted versus actual; plot of predictions alone; histogram
15D (n=8)	0	2 (25)	5 (63)	8 (100)	6 (75)	2 (25)	3 (38)
AQoL (n=12)	0	4 (33)	7 (58)	12 (100)	11 (92)	4 (33)	5 (42)
CHU-9D (n=4)	0	3 (75)	3 (75)	4 (100)	4 (100)	1 (25)	3 (75)
EQ-5D (n=147)	31 (21)	81 (55)	75 (51)	139 (95)	121 (82)	59 (40)	75 (51)
HUI2/HUI3 (n=13)	2 (15)	4 (31)	8 (62)	13 (100)	8 (62)	7 (54)	5 (38)
QWB (n=4)	0	0	3 (75)	4 (100)	3 (75)	1 (25)	1 (25)
SF-6D (n=45)	13 (29)	25 (56)	27 (60)	44 (98)	39 (87)	18 (40)	28 (62)
Total (n=233)	**46 (20)**	**119 (51)**	**128 (55)**	**224 (96)**	**192 (82)**	**92 (39)**	**120 (52)**

MAE- Mean Absolute Error; RMSE – Root Mean Squared Error; MSE – Mean Squared Error; ME – Mean Error

15D – 15 Dimensions; AQoL-4/8D – Assessment of Quality of Life 4 or 8 Dimensions; CHU-9D – Child Health Utility – 9 Dimensions; HUI2/3 – Health Utilities Index 2 or 3; QWB – Quality of Wellbeing Scale; SF-6D – Short Form – 6 Dimensions

EQ-5D – includes EQ-5D three level, five level and youth versions