Can The EQ-5D Detect Meaningful Change? A Systematic Review

Nalin Payakachat[1], **Mir M. Ali**[2], and **J. Mick Tilford**[1,2]

[1]Division of Pharmaceutical Evaluation and Policy, Department of Pharmacy Practice, University of Arkansas for Medical Sciences, Little Rock, AR

[2]Department of Health Policy and Management, University of Arkansas for Medical Sciences, Little Rock, AR

Abstract

Background—The EQ-5D is one of the most frequently used generic, preference-based instruments for measuring the health utilities of patients in economic evaluations. It is recommended for health technology assessment by the National Institute for Health and Clinical Excellence. Because the EQ-5D plays such an important role in economic evaluations, useful information on its responsiveness to detect meaningful change in health status is required.

Objective—This study systematically reviewed and synthesized evidence on the responsiveness of the EQ-5D to detect meaningful change in health status for clinical research and economic evaluations.

Methods—We searched the EuroQol website, PubMed, PsychINFO, and EconLit databases to identify studies published in English from the inception of the EQ-5D until August 15, 2014 using keywords that were related to responsiveness. Studies that used only the EQ-VAS were excluded from the final analysis. Narrative synthesis was conducted to summarize evidence on the responsiveness of the EQ-5D by conditions or physiological functions.

Results—Of 1,401 studies, 145 were included in the narrative synthesis and categorized into 19 categories for 56 conditions. The EQ-5D was found to be responsive in 25 conditions (45%) with the magnitude of responsiveness varying from small to large depending on the condition. There was mixed evidence of responsiveness in 27 conditions (48%). Only four conditions (7%) (i.e., alcohol dependency, schizophrenia, limb reconstruction, and hearing impairment) were identified where the EQ-5D was not responsive.

Corresponding author: Nalin Payakachat, PhD, Division of Pharmaceutical Evaluation and Policy, Department of Pharmacy Practice, University of Arkansas for Medical Sciences, Little Rock, AR 72205, npayakachat@uams.edu, Tel: 1-501-686-7701, Fax: 1-501-686-5156.

Conclusion—The EQ-5D is an appropriate measure for economic evaluation and health technology assessment in conditions where it has demonstrated evidence of responsiveness. In conditions with mixed evidence of responsiveness, researchers should consider using the EQ-5D with other condition-specific measures to ensure appropriate estimates of effectiveness. These conditions should be a main focus for future research using the new EQ-5D version with 5 response levels.

1 Introduction

With health care spending constituting an ever-increasing component of national spending worldwide, economic evaluations of health care technologies have become an important tool in informing health policy and making resource allocation decisions [1;2]. Economic evaluation methods such as cost-effectiveness analysis provide a means of assessing interventions in terms of their costs in relation to their benefits measured in quality adjusted life years (QALY) [2]. The QALY is a standard metric that is recommended for cost-effectiveness analysis by the National Institute for Health and Clinical Excellence (NICE) and the U.S. Panel on Cost-Effectiveness in Health and Medicine [2;3]. It combines the value of health-related quality of life (HRQL) or preference-based health utility and life years gained into a single metric. QALY information can be measured using generic, preference-based measures of HRQL such as the EQ-5D (EuroQol – 5 dimension descriptive system), SF-6D (Short-Form – 6 dimension), and HUI (Health Utilities Index) [4–7] among others.

The EQ-5D is the preferred preference-based measure as per the NICE methods guide for health technology assessment [2]. The EQ-5D comprises five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [8]. The original measure had three response categories including no problems, some/moderate problems or severe/extreme problems on each domain. This creates a total of 243 possible combinations of unique health states (3^5) that are linked to predetermined preference-weighted scores yielded from direct utility elicitation such as a time trade-off or visual analog scale (VAS) approaches. Health utility values generated from the EQ-5D generally range from 0 (death) to 1 (perfect health). But health utility values less than 0 are possible, and represent health states considered worse than death. The EQ-5D is often administered with the EQ-VAS where respondents report their self-rated valuation of their health state on a scale of 0 – 100. However, the EQ-VAS is not generally used to calculate QALY scores for CEA due to a concern that VAS is inferior to choice-based methods [9].

The most attractive features of the EQ-5D instrument include its brevity (5 questions with 3 response categories), and the fact that it is cognitively simple. In addition, it is available in more than 150 official languages and offers several population weights (e.g. different value sets for the UK, France, Germany, Netherlands, Denmark, Spain, Japan, USA, etc.) [10]. Because the EQ-5D is so brief with only 3 levels, the major concern is its responsiveness. A new version of the EQ-5D with 5 levels (EQ-5D 5L) was developed to address this concern [11]. The EQ-5D 5L creates a total of 3,125 unique health states (5^5). The development of preference-weighted scores associated with different possible health states for the EQ-5D 5L

is still in progress and in the interim, "cross-walks" between the EQ-5D 3L value sets and the EQ-5D 5L have been developed to facilitate use [12].

Apart from the reliability and validity of an instrument, its responsiveness to detect clinical change is a critical property. If an instrument is not sensitive to clinical changes, it will fail to detect benefits or harms of interventions or treatments. From an economic evaluation perspective, use of such an instrument would result in gross misstatement about the value of different technologies and pharmaceutical products. To address this issue, condition-specific measures (CSM) are often used to measure effectiveness in clinical studies. However, CSMs cannot produce preference-based HRQL weights for economic evaluation. Thus, information on the responsiveness of generic preference-based measures such as the EQ-5D need to be evaluated across conditions and specific treatments and technologies in health care.

Previous systematic reviews of the responsiveness of the EQ-5D have been condition-specific such as mental health, skin conditions, breast cancer, asthma & chronic obstructive pulmonary disease (COPD) [13–16]. The most recent systematic review of the EQ-5D that assessed responsiveness across various conditions reported that the instrument has poor to moderate responsiveness to clinical changes [17]. However, the authors confined their literature search to just one database and had a very narrow scope for study identification, using just one keyword. The question of whether the EQ-5D can detect clinically meaningful change in different conditions thus remains.

2 Methods

2.1 Search Strategy

We conducted a systematic search of published studies that reported the psychometric properties of the EQ-5D, in particular, the instrument's responsiveness. The search strategy focused on keywords including 'EQ5D' or 'EQ-5D' or 'euroqol' or 'EuroQol-5D' with the following additional terms: sensitive to change, sensitivity, responsiveness, responsive to change, minimally clinically important difference, minimally important difference, clinical significant change, or clinically significant change. We limited our search criteria to articles that were published in English, for a time period beginning with the inception of the EQ-5D until August 15, 2014. Since the EQ-5D was developed by the EuroQol group [18], we believed that the EQ-5D references available on their database would be inclusive of all studies that involved the use of the EQ-5D instrument. Thus, we used the EuroQol group database as one of the sources to identify relevant studies. Since the EQ-5D may be used in studies of clinical conditions, interventions and treatments, we supplemented our data sources with the PubMed and PsychINFO databases. Finally, since the EQ-5D is also used in the economic evaluation of health care technologies, we believed the EconLit database to be an important data source to identify relevant studies. Moreover, including the PubMed, PsychINFO and EconLit databases would also be useful in identifying systematic reviews that were conducted regarding the EQ-5D, which were possibly not included in the EuroQol database.

2.2 Study Identification

The titles and abstracts of all the studies that were identified using the keyword search from each of the four databases were cataloged in an excel spreadsheet. We excluded duplicates across databases, based on the study titles. Articles that had full-texts in languages other than English were also excluded. We also eliminated articles that did not have one of the keywords within the abstracts, nor deemed to be relevant.

Studies that used only the EQ-VAS were omitted due to concerns that the VAS does not explicitly involve choice, nor provide a cardinal measure that is needed for QALYs. As described earlier, the EQ-5D index comprises five dimensions and three response categories per dimension creating a total of 243 possible unique health states. Each of the health states is converted to a single summary index by applying a formula that attaches weights to each of the levels in each dimension. The formula is based on the valuation of EQ-5D health states obtained from general population samples using the standardized extended version of the EQ-5D that collects health state values using the time trade-off (TTO) method. The EuroQol group suggests that the EQ-VAS be used only in instances where valuation of health states is being sought and not as part of routine clinical and economic studies. As a result, we chose to assess the responsiveness of the EQ-5D based solely on the index score and did not focus on data collected using the EQ-VAS in any of the studies. Articles that were part of a recent condition-specific systematic review (within the past 5 years) of the responsiveness of the EQ-5D for specific diseases/conditions were reviewed along with the systematic review articles. We reported responsiveness evidence from the systematic reviews along with additional evidence from recent literature. Of all the articles that were identified for full-text review, NP and MMA randomly and independently reviewed the full-text articles to extract responsiveness information for the EQ-5D.

2.3 Data extraction

We extracted data from the included studies in a tabulated form, which covered general characteristics of the study and participants. We classified the studies based on the condition/disease that was being reviewed within the study. We identified the study design, the sample characteristics, the instruments that were used in the study, and the methods and relevant results provided in the study for an assessment of responsiveness. We did not require the studies to be primarily designed to assess responsiveness, but rather provided sufficient information within the study to allow us to make an assessment.

2.4 Assessment of responsiveness

We applied the definition of responsiveness as "the extent to which an instrument can detect a clinically significant or practically important change over time" [19]. Based on this definition, we developed a list of three relevant measures of responsiveness that were extracted from published studies. These measures included 1) differences in the EQ-5D health utility scores between responders and non-responders by clinical or self-reported measures and 2) change in EQ-5D health utility values over a period of time in which health status is expected to change (e.g., before and after an intervention) with the change demonstrated by another measure of health. We reported responsiveness in terms of the standardized response means (SRM), effect size (ES), area under the receiver-operating

characteristic (AUROC) curve, or responsiveness statistic, when applicable [20]. Cohen's categories for magnitude of ES and SRM were used to identify small (<0.5), moderate (0.5 – 0.79), and large (0.8) effect sizes [21]. The magnitude of the AUROC curve was classified into five categories to represent the performance of the EQ-5D including irresponsive (< 0.5), poor (0.5 – 0.59), sufficient (0.6 – 0.69), good (0.7 – 0.79), very good (0.8 – 0.89), and excellent (0.9–1.0) [22].

2.5 Evidence Synthesis

We identified a large degree of heterogeneity between studies in terms of the study designs, population characteristics, instruments used to assess HRQL, outcome measures, and methods used for assessing responsiveness. Therefore, we did not attempt to conduct a meta-analysis for the EQ-5D responsiveness in this review. Instead, we conducted a narrative synthesis and tabulated data by conditions or physiological functions, when possible.

3 Results

3.1 Study Selection Process

A total of 1,401 studies were identified from the four databases. A preliminary screening led to the exclusion of 559 duplicate articles and 21 articles that were published in other languages. Abstracts of the remaining 821 studies were further screened for relevance. Through this process, we excluded 377 studies that did not have our specified key words within the abstract and 281 studies lacked relevance. 163 studies were reviewed in full-text. 18 studies were excluded after the review due to a) irrelevance (n=12) or b) the study used the EQ-5D VAS for determining responsiveness (n=4) or c) the study presented clinical validity or known-group validity as responsiveness (n=2). Therefore, 145 studies were included in the final synthesis. The study identification is presented in Figure 1. Among these 145 studies, there were 60 studies published on diseases or conditions previously reviewed in condition-specific systematic review articles.

3.2 Responsiveness Evidence

Table 1 provides responsiveness evidence of the EQ-5D extracted from the literature. We categorized 145 studies into 56 diseases/conditions within 19 categories of physiological functions or conditions. There were 14 systematic or narrative reviews of specific conditions regarding responsiveness of the EQ-5D. Although asthma & COPD were described in a previous systematic review [16], the study was outdated. We extracted the responsiveness information from the systematic review and added additional evidence from other studies that were published in recent years. Reported responsiveness of the EQ-5D depended not only on the magnitude of change in health status, but also varied by the types of external anchors and methods used to calculate responsiveness. Approximately half of the studies calculated responsiveness of the EQ-5D using external anchors based on either clinical measures or self-reported measures or both (external responsiveness or anchor-based method). The other half used changes in the EQ-5D index from baseline to a follow-up period to determine responsiveness (internal responsiveness or distribution-based method). A variety of statistical approaches were used to calculate or determine responsiveness including regression, *t*-test, nonparametric *t*-test, ES, SRM and the area under the ROC

curve. Most studies were conducted in Europe (UK, Germany, Netherlands, Sweden), Canada, and the U.S. Very few articles were included from Southeast Asia (Japan, Singapore).

Within 56 specific conditions, the EQ-5D was found to be responsive in 25 conditions (45%). There were only four conditions (7%) (i.e., alcohol dependency [23;24], schizophrenia [25], limb reconstruction [26], and hearing impairment [27]) where the EQ-5D lacked responsiveness. The other 27 conditions (48%) have limited or mixed evidence for the responsiveness of the EQ-5D. All studies were based on adult populations with one exception: a study of Attention Deficit Hyperactivity Disorder (ADHD) in children and adolescents [28].

Table 2 provides a summary of the conditions that the EQ-5D was found to be responsive. The magnitude of the EQ-5D responsiveness varied from small to large depending on the magnitude of the changes in the health condition. The EQ-5D was found to be responsive to health improvement in most studied conditions except for liver metastases. For health deterioration, the EQ-5D was found to be responsive to 7 out of 16 studied conditions including inflammatory arthritis, breast cancer, liver metastases, multiple myeloma, dementia, surgery, and adverse effects of HIV treatments. However, 9 conditions did not have information if the EQ-5D is responsive to worse health.

4 Discussion

The EQ-5D is one of the most frequently used generic, preference-based measures in clinical studies in Europe and North America to measure quality of life and preference-based HRQL scores for economic evaluations. Psychometric properties of the EQ-5D on validity and reliability have been more or less confirmed in published literature while evidence of responsiveness is somewhat varied [105–108]. Since responsiveness is the ability of an instrument to detect health status changes, it is an essential property that is important for health technology assessments. While NICE recommends the EQ-5D for economic evaluation, they also allow researchers to use other generic, preference-based measures such as the SF-6D or the HUI3 if there is evidence that the EQ-5D is not appropriate for that condition [2]. This study provides an extensive systematic review of the responsiveness of the EQ-5D in order to provide more succinct evidence to either support its use in cost-effectiveness analysis or to recommend alternative approaches.

We included multiple data sources and an exhaustive list of search terms to identify relevant studies beyond the previous systematic review study [17]. We found that the EQ-5D was responsive to half of the conditions reviewed. The other half had mixed evidence of responsiveness, suggesting that researchers need to incorporate CSMs along with the EQ-5D to ensure that appropriate measures of effectiveness are reported. Disease-specific measures or CSMs are generally developed based on symptoms or clinical characteristics of the conditions/diseases. Thus, they are more sensitive for the detection and quantification of small clinical changes [109]. If researchers find changes on the EQ-5D after treatments and significant clinical changes on the CSMs, it means that the EQ-5D is able to detect changes resulting from the treatments suggesting a level of responsiveness can be calculated and

confirmed. On the contrary, if researchers do not find any change on the EQ-5D after treatment but do find significant clinical changes on the CSMs, it means that the EQ-5D is not able to detect changes that resulted from the treatments. The EQ-5D was not responsive in four conditions (alcohol dependency [23;24], schizophrenia [25], limb reconstruction [26], and hearing impairment [27]). In addition, a condition where the EQ-5D may not be responsive is hemophilia, based on the results of a study that was published after our search period [110]. Ceiling effects were observed in a few studies conducted in non-acute conditions (e.g., asthma & COPD [16;50], proximal humeral fracture [74], hearing impairment [27], urinary incontinence [59]). There was no report of a flooring effect in any study.

We found large heterogeneity among the studies across several methodological issues related to measuring responsiveness that need to be addressed. We address each of these issues in turn.

Selection of external anchors—There were a variety of clinical measures and CSMs used in studies for the same condition which affected the calculations of responsiveness for the EQ-5D [35]. Depending on the relevant change in an anchor/clinical measure, the magnitudes of ES, SRM, and AUROC curve may differ across different anchors/clinical measures [111]. There is no gold standard measure for many conditions or specific guidance on what external anchors should be used in order to calculate responsiveness. Researchers normally select measures based on their familiarity with them rather than their psychometric properties. We suggest that researchers should consider selecting anchors based on relevance for the disease indication and clinical acceptability for the specific disease condition [111]. Chosen anchors should also have some relationship with the EQ-5D, otherwise it may provide misleading information in determining whether significant change has occurred. For example, a study on coronary artery disease used chest pain as an external anchor for determining EQ-5D responsiveness [29]. Although chest pain is one of the important symptoms of coronary artery disease, it may not be a good indicator of health status change for this condition at the 2-year follow-up period. Another example is from a study conducted in patients with heart failure [31]. The magnitude of responsiveness of the EQ-5D was larger when patients reported moderate improvement on a physician global rating of change (an external anchor), when compared to +1 class improvement on the New York Heart Association class.

A popular external anchor used in the literature is the global rating scale, which asks patients to report whether they got better, stayed the same, or got worse after some period of time. Several problems are apparent with this anchor as some patients will get better or worse, just by chance, and some patients may not remember how they actually felt at baseline. As a result, the variability in individual responses makes it difficult to detect treatment effects [112].

Severity of conditions—Psychometric properties of an instrument are specific to the population and condition being studied. Thus, it is important to provide information regarding both population and condition being tested. Furthermore, researchers must describe the severity of the condition when reporting psychometric properties of the

instrument such as its responsiveness. The EQ-5D, like other instruments, is more responsive to large treatment effects than to small ones [112]. Similarly, the EQ-5D is more responsive to a large change in health found in a moderate-to-severe condition than to a small change in a mild condition. In our review, we consistently found that the EQ-5D is more likely to be responsive if conditions are more severe or if a large change is observed, which is consistent with findings from the previous systematic review of the EQ-5D by Tordrup et al. [17]. For example, in patients with mild asthma, the EQ-5D could not detect any change in health states, but for patients with moderate-to-severe asthma, the EQ-5D is responsive to health improvement [16;50]. We also observed that the ceiling effect is less likely to be a problem for the EQ-5D if it is used in moderate-to-severe conditions. For this reason, the severity of studied conditions should be specified in order to help other researchers evaluate if the EQ-5D is appropriate to measure changes in the health status.

Timing of follow-up—Defining an appropriate timing for follow-up periods in studying different conditions is crucial to measuring responsiveness. Timing of follow-up should also correspond appropriately to the natural history and progression of the disease and the likelihood of change in health status within that period [113]. For example, a one-week recall period may be used to monitor the severity of pain. For measuring the benefits of total hip replacement, a short-term follow-up period could be 1 – 6 months, while 1 – 2 weeks would be appropriate for heart failure management. If researchers do not fully understand the course of the condition or do not monitor patients in a timely manner, they may be misled by unchanged outcomes and ultimately miss the opportunity to demonstrate effectiveness of the interventions.

Statistical methods—Several responsiveness statistics are used in the literature and there is no consensus on alternative measures [112]. In addition, there is no information on how to relate different responsiveness statistics to one another, which makes it difficult to summarize the magnitude of responsiveness. Clear guidance on measures and the development of a common metric are needed.

Definition of known-group validity or clinical validity vs. responsiveness—
Some studies tested known-group or clinical validity of the EQ-5D and reported that the EQ-5D is sensitive to the condition. While known-group or clinical validity is one of the important psychometric properties of measures, it is not the same as responsiveness. Known-group or clinical validity does not require evidence of change in health over time, but rather compares health status between healthy and ill individuals [112]. For this reason, we excluded articles that reported known-group validity or clinical validity instead of responsiveness after full-text review. Incorrectly applying a definition of known-group validity for reporting responsiveness could be misleading.

Different population weights of the EQ-5D—Different population weights can affect the magnitude of the responsiveness statistics such as the ES and SRM. Using the UK population weights may result in a slightly lower magnitude of responsiveness compared to the same study using the US population weights. However, using German population weights may results in a much lower magnitude than using the UK or US population weights

and could lead to nil responsiveness of the EQ-5D for the same patient populations [23;38]. A similar phenomenon was observed in another study that applied both UK and Dutch population weights to detect changes among patients with depression [114]. The EQ-5D UK index showed a moderate magnitude of responsiveness whereas the EQ-5D Dutch index showed a large magnitude. The scoring algorithm of the EQ-5D index reflects societal preference-based valuations attached to each domain and this valuation may vary between different populations. This observation is consistent with the most recent systematic review of the EQ-5D [17].

Coping—Coping is a potential mediating mechanism that has been shown to have an association with improved quality of life, especially psychological functioning across chronic conditions [115;116]. It is defined as a dynamic process of an individual's cognitive and/or behavioral attempt to manage stress-related situations. Cancers are the most studied condition for this particular phenomenon [117;118]. Hearing and vision impairments are other conditions where coping mechanisms may have influence on detecting health change using the EQ-5D. While clinical change of hearing and vision can be detected by audiometer tests and vision tests (e.g., Snellen, visual field tests), patients might report themselves to be the same (due to coping) on all five dimensions of the EQ-5D. On the contrary, patients might report changes on the HUI3 which contains hearing and vision dimensions [7].

4.1 Study Limitations

There are limitations in our study that we need to acknowledge. First, we did not report responsiveness information from original studies for those conditions that have had systematic reviews published previously, to avoid redundancy. Instead, we reported the results from the systematic reviews for those conditions and provided updated information if the systematic review was outdated. However, most systematic review studies evaluated all psychometric properties of different preference-based measures including the EQ-5D with limited information on the measures' responsiveness. Also, the information presented in previous condition-specific systematic reviews was dependent on authors' research scope and styles of reporting. For example, Davis and Wailoo (2013) reported responsiveness with a focus on significant changes in patients with urinary incontinence over time and did not report regarding the magnitude of responsiveness [59]. Mulhern et al. (2014) reported responsiveness on both significant changes and the magnitude of responsiveness for patients with mental health problems [36]. As a result, although we saved time in reviewing studies, the prior systematic reviews might not provide enough depth of information on responsiveness. Secondly, although we utilized several key words to identify studies for evaluating the responsiveness of the EQ-5D, it is possible that we missed some studies that had responsiveness information in the text, but not in the title or abstract. However, we believe that it is unlikely for a study to assess the responsiveness of the EQ-5D and not mention any of our broad keywords in the abstract. Another limitation is the sample size in the original studies. Most studies that we reviewed were conducted in relatively small samples (<200) and attrition became a prominent problem when calculating responsiveness of the EQ-5D based on responders and non-responders. In order to minimize the issues that come with calculating responsiveness with small sample sizes, we excluded studies that had sample sizes less than 30 in each group at the follow-up period. In terms of reporting

responsiveness, there were some discrepancies between the two reviewers (NP and MMA), but all discrepancies were deliberated and resolved by re-review and discussion. Finally, we recognize that there is discrepancy between patient and population weights and there is debate over whose preference weights should be used in economic evaluation. Patient's perception may not be the same as that of the general population which could result in smaller or larger magnitude of responsiveness [17].

5 Conclusion

In summary, we systematically reviewed the evidence describing the responsiveness of the EQ-5D for a large number of conditions. Researchers who seek evidence on QALYs for economic evaluation of interventions should review the conditions and population described to determine if the EQ-5D is a responsive measure. In the case of conditions with mixed evidence of responsiveness, other CSMs are recommended to use alongside the EQ-5D to determine effectiveness. Conditions lacking responsiveness might be a prime area of focus for future research using the EQ-5D 5L.

Acknowledgments

Abbreviation List

AUC	Area under the curve
AUROC	Area under the receiver-operating characteristic curve
COPD	Chronic obstructive pulmonary disease
CSM	Condition-specific measures
EQ-5D 5L	EuroQol – 5 dimension 5 level descriptive system
EQ-5D	EuroQol – 5 dimension 3 level descriptive system
EQ-VAS	EuroQol – visual analogue scale
ES	Effect size
EULAR	European League against Rheumatism
HRQL	Health-related quality of life
HUI	Health Utilities Index
HUI3	Health Utilities Index Mark 3
NICE	National Institute for Health and Clinical Excellence
NYHA	New York Heart Association class
QALY	Quality-adjusted life year
SF-6D	Short-Form – 6 dimension
SRM	Standardized response means

| **TTO** | Time trade-off |
| **WOMAC** | Western Ontario and McMaster Universities Osteoarthritis Index |

References

1. Weinstein MC, Skinner JA. Comparative effectiveness and health care spending--implications for reform. N Engl J Med. 2010; 362(5):460–5. [PubMed: 20054039]

2. NICE. [Accessed 15 Nov 2014] Guide to the methods of technology appraisal 2013. National Institute for Health and Care Excellence. 2013. http://www.nice.org.uk/article/pmg9/resources/non-guidance-guide-to-the-methods-of-technology-appraisal-2013-pdf

3. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. JAMA. 1996; 276(15):1253–8. [PubMed: 8849754]

4. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001; 33(5):337–43. [PubMed: 11491192]

5. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002; 21(2):271–92. [PubMed: 11939242]

6. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care. 2004; 42(9):851–9. [PubMed: 15319610]

7. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, Denton M, Boyle M. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002; 40(2):113–28. [PubMed: 11802084]

8. Oemar, M.; Oppe, M. EQ-5D-3L User Guide. EuroQol Group; 2013. http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf [Accessed 15 Nov 2014]

9. Gray, AM.; Clarke, P.; Wolstenholme, J.; Wordsworth, S. Measuring, valuing, and analysing health outcomes. In: Gray, AM.; Clarke, P.; Wolstenholme, J.; Wordsworth, S., editors. Applied methods of cost-effectiveness analysis in healthcare. New York, NY: Oxford University Press; 2011. p. 83-118.

10. Szende, A.; Oppe, M.; Devlin, N. EuroQol Group Monographs. Vol. 2. Dordrecht, The Netherlands: Springer; 2007. EQ-5D value sets: Inventory, comparative review and user guide.

11. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011; 20(10):1727–36. [PubMed: 21479777]

12. EuroQol Group. EQ-5D-5L Value Sets. EuroQol Research Foundation; 2014. http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-5l-value-sets.html [Accessed 15 Nov 2014]

13. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, Jones ML, Paisley S, O'Cathain A, Barkham M, Knapp M, Byford S, Gilbody S, Parry G. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess. 2014; 18(34):vii–viii. xiii–xxv, 1–188.

14. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. Eur J Health Econ. 2014 Epub ahead of print.

15. Davies, N.; Gibbons, E.; Mackintosh, A.; Fitzpatrick, R. A structured review of patient-reported outcome measures for women with breast cancer. Patient-reported Outcome Measurement Group, Department of Public Health, University of Oxford; 2009. http://phi.uhce.ox.ac.uk/pdf/CancerReviews/PROMs_Oxford_BreastCancer_012011.pdf [Accessed 15 Nov 2014]

16. Pickard AS, Wilke C, Jung E, Patel S, Stavem K, Lee TA. Use of a preference-based measure of health (EQ-5D) in COPD and asthma. Respir Med. 2008; 102(4):519–36. [PubMed: 18180151]

17. Tordrup D, Mossman J, Kanavos P. Responsiveness of the EQ-5D to clinical change: is the patient experience adequately represented? Int J Technol Assess Health Care. 2014; 30(1):10–19. [PubMed: 24499622]

18. EuroQol Group. EQ-5D References - Reference Search. EuroQol Research Foundation; 2014. http://www.euroqol.org/eq-5d-references/reference-search.html [Accessed 15 Nov 2014]

19. Walters, SJ. Quality of Life Outcomes in Clinical Trials and Health-Care Evaluation: a Practical Guide to Analyusis and Interpretation. Chichester, UK: John Wiley & Sons; 2009.

20. Hays, RD.; Revicki, DA. Reliability and validity (including responsiveness). In: Fayers, PM.; Hays, RD., editors. Assessing quality of life in clinical trials. 2. New York, NY: Oxford University Press; 2005. p. 25-39.

21. Cohen J. A power primer. Psychol Bull. 1992; 112(1):155–9. [PubMed: 19565683]

22. Simundic, AM. [Accessed 15 Nov 2014] Measures of diagnostic accuracy: basic definition. 2011. http://www.ifcc.org/ifccfiles/docs/190404200805.pdf

23. Günther OH, Roick C, Angermeyer MC, König HH. Responsiveness of EQ-5D utility indices in alcohol-dependent patients. Drug Alcohol Depend. 2008; 92(1–3):291–5. [PubMed: 17888587]

24. Essex HN, White IR, Khadjesari Z, Linke S, McCambridge J, Murray E, Parrott S, Godfrey C. Quality of life among hazardous and harmful drinkers: EQ-5D over a 1-year follow-up period. Qual Life Res. 2014; 23(2):733–43. [PubMed: 24026632]

25. Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. Value Health. 2011; 14(6):907–20. [PubMed: 21914513]

26. Burton M, Walters SJ, Saleh M, Brazier JE. An evaluation of patient-reported outcome measures in lower limb reconstruction surgery. Qual Life Res. 2012; 21(10):1731–43. [PubMed: 22179990]

27. Yang Y, Longworth L, Brazier J. An assessment of validity and responsiveness of generic measures of health-related quality of life in hearing impairment. Qual Life Res. 2013; 22(10):2813–28. [PubMed: 23709096]

28. Bouwmans C, van der Kolk A, Oppe M, Schawo S, Stolk E, van Agthoven M, Buitelaar J, van Roijen L. Validity and responsiveness of the EQ-5D and the KIDSCREEN-10 in children with ADHD. Eur J Health Econ. 2014; 15(9):967–77. [PubMed: 24233919]

29. Kiessling A, Henriksson P. Time trends of chest pain symptoms and health related quality of life in coronary artery disease. Health Qual Life Outcomes. 2007; 5:13. [PubMed: 17341296]

30. Schweikert B, Hahmann H, Leidl R. Validation of the EuroQol questionnaire in cardiac rehabilitation. Heart. 2006; 92(1):62–7. [PubMed: 15797936]

31. Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. Health Qual Life Outcomes. 2006; 4:89. [PubMed: 17125512]

32. Kaplan RM, Tally S, Hays RD, Feeny D, Ganiats TG, Palta M, Fryback DG. Five preference-based indexes in cataract and heart failure patients were not equally responsive to change. J Clin Epidemiol. 2011; 64(5):497–506. [PubMed: 20685077]

33. Pickard AS, Johnson JA, Feeny DH. Responsiveness of generic health-related quality of life measures in stroke. Qual Life Res. 2005; 14(1):207–19. [PubMed: 15789955]

34. Hunger M, Sabariego C, Stollenwerk B, Cieza A, Leidl R. Validity, reliability and responsiveness of the EQ-5D in German stroke patients undergoing rehabilitation. Qual Life Res. 2012; 21(7):1205–16. [PubMed: 21971874]

35. Nosyk B, Sun H, Guh DP, Oviedo-Joekes E, Marsh DC, Brissette S, Schechter MT, Anis AH. The quality of eight health status measures were compared for chronic opioid dependence. J Clin Epidemiol. 2010; 63(10):1132–44. [PubMed: 20236799]

36. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Soeteman D, Brazier J. Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. Br J Psychiatry. 2014; 205(3):236–43. [PubMed: 24855127]

37. Peasgood, T.; Brazier, J.; Papaioannou, D. A systematic review of the validity and responsiveness of EQ-5D and SF-6D for depression and anxiety. The University of Sheffield; UK: White Rose Research Online. http://eprints.whiterose.ac.uk/74659/1/12.15.pdf [Accessed 15 Nov 2014]

38. Sonntag M, Konnopka A, Leichsenring F, Salzer S, Beutel ME, Herpertz S, Hiller W, Hoyer J, Joraschky P, Nolting B, Pöhlmann K, Stangier U, Strauss B, Willutzki U, Wiltink J, Leibing E, König HH. Reliability, validity and responsiveness of the EQ-5D in assessing and valuing health

status in patients with social phobia. Health Qual Life Outcomes. 2013; 11:215.10.1186/1477-7525-11-215.:215-11 [PubMed: 24365384]

39. Brettschneider C, König HH, Herzog W, Kaufmann C, Schaefert R, Konnopka A. Validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with somatoform disorders. Health Qual Life Outcomes. 2013; 11:3. [PubMed: 23305192]

40. Harrison MJ, Davies LM, Bansback NJ, McCoy MJ, Verstappen SM, Watson K, Symmons DP. British Society for Rheumatology Biologics Register Control Centre Consortium. . The comparative responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. Qual Life Res. 2009; 18(9):1195–205. [PubMed: 19777373]

41. Hagel S, Lindqvist E, Petersson IF, Nilsson JÅ, Bremander A. Validation of outcome measurement instruments used in a multidisciplinary rehabilitation intervention for patients with chronic inflammatory arthritis: linking of the International Classification of Functioning, Disability and Health, construct validity and responsiveness to change. J Rehabil Med. 2011; 43(5):411–9. [PubMed: 21448555]

42. Gaujoux-Viala C, Rat AC, Guillemin F, Flipo RM, Fardellone P, Bourgeois P, Fautrel B. Responsiveness of EQ-5D and SF-6D in patients with early arthritis: results from the ESPOIR cohort. Ann Rheum Dis. 2012; 71(9):1478–83. [PubMed: 22419775]

43. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol. 1997; 36(5):551–9. [PubMed: 9189057]

44. Eversden L, Maggs F, Nightingale P, Jobanputra P. A pragmatic randomized controlled trial of hydrotherapy and land exercises on overall well being and quality of life in rheumatoid arthritis. BMC Musculoskelet Disord. 2007; 8:23. [PubMed: 17331241]

45. Linde L, Sørensen J, Ostergaard M, Hørslev-Petersen K, Hetland ML. Health-related quality of life: validity, reliability, and responsiveness of SF-36, 15D, EQ-5D [corrected] RAQoL, and HAQ in patients with rheumatoid arthritis. J Rheumatol. 2008; 35(8):1528–37. [PubMed: 18484697]

46. Buitinga L, Braakman-Jansen LM, Taal E, Kievit W, Visser H, van Riel PL, van de Laar MA. Comparative responsiveness of the EuroQol-5D and Short Form 6D to improvement in patients with rheumatoid arthritis treated with tumor necrosis factor blockers: results of the Dutch Rheumatoid Arthritis Monitoring registry. Arthritis Care Res (Hoboken). 2012; 64(6):826–32. [PubMed: 22262513]

47. Haywood KL, Garratt AM, Dziedzic K, Dawes PT. Generic measures of health-related quality of life in ankylosing spondylitis: reliability, validity and responsiveness. Rheumatology (Oxford). 2002; 41(12):1380–7. [PubMed: 12468817]

48. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. Med Care. 2003; 41(7):791–801. [PubMed: 12835603]

49. Oga T, Nishimura K, Tsukino M, Sato S, Hajiro T, Mishima M. A comparison of the responsiveness of different generic health status measures in patients with asthma. Qual Life Res. 2003; 12(5):555–63. [PubMed: 13677500]

50. Petrillo J, van Nooten F, Jones P, Rutten-van Mölken M. Utility estimation in chronic obstructive pulmonary disease: a preference for change? Pharmacoeconomics. 2011; 29(11):917–32. [PubMed: 21988291]

51. Paterson C, Langan CE, McKaig GA, Anderson PM, Maclaine GD, Rose LB, Walker SJ, Campbell MJ. Assessing patient outcomes in acute exacerbations of chronic bronchitis: the measure your medical outcome profile (MYMOP), medical outcomes study 6-item general health survey (MOS-6A) and EuroQol (EQ-5D). Qual Life Res. 2000; 9(5):521–7. [PubMed: 11190007]

52. Jenkinson C, Gray A, Doll H, Lawrence K, Keoghane S, Layte R. Evaluation of index and profile measures of health status in a randomized controlled trial. Comparison of the Medical Outcomes Study 36-Item Short Form Health Survey, EuroQol, and disease specific measures. Med Care. 1997; 35(11):1109–18. [PubMed: 9366890]

53. MacDonagh RP, Cliff AM, Speakman MJ, O'Boyle PJ, Ewings P, Gudex C. The use of generic measures of health-related quality of life in the assessment of outcome from transurethral resection of the prostate. Br J Urol. 1997; 79(3):401–8. [PubMed: 9117222]

54. Krahn M, Bremner KE, Tomlinson G, Ritvo P, Irvine J, Naglie G. Responsiveness of disease-specific and generic utility instruments in prostate cancer patients. Qual Life Res. 2007; 16(3): 509–22. [PubMed: 17091359]

55. Kimman ML, Dirksen CD, Lambin P, Boersma LJ. Responsiveness of the EQ-5D in breast cancer patients in their first year after treatment. Health Qual Life Outcomes. 2009; 7:11.10.1186/1477-7525-7-11.:11-17 [PubMed: 19200391]

56. Lee CF, Luo N, Ng R, Wong NS, Yap YS, Lo SK, Chia WK, Yee A, Krishna L, Wong C, Goh C, Cheung YB. Comparison of the measurement properties between a short and generic instrument, the 5-level EuroQoL Group's 5-dimension (EQ-5D-5L) questionnaire, and a longer and disease-specific instrument, the Functional Assessment of Cancer Therapy-Breast (FACT-B), in Asian breast cancer patients. Qual Life Res. 2013; 22(7):1745–51. [PubMed: 23054499]

57. Krabbe PF, Peerenboom L, Langenhoff BS, Ruers TJ. Responsiveness of the generic EQ-5D summary measure compared to the disease-specific EORTC QLQ C-30. Qual Life Res. 2004; 13(7):1247–53. [PubMed: 15473503]

58. Kvam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. Eur J Haematol. 2011; 87(4):330–7. [PubMed: 21668504]

59. Davis S, Wailoo A. A review of the psychometric performance of the EQ-5D in people with urinary incontinence. Health Qual Life Outcomes. 2013; 11:20. [PubMed: 23418844]

60. Kunz S. Psychometric properties of the EQ-5D in a study of people with mild to moderate dementia. Qual Life Res. 2010; 19(3):425–34. [PubMed: 20146008]

61. Gray R, Ives N, Rick C, Patel S, Gray A, Jenkinson C, McIntosh E, Wheatley K, Williams A, Clarke CE. PD Med Collaborative Group. Long-term effectiveness of dopamine agonists and monoamine oxidase B inhibitors compared with levodopa as initial treatment for Parkinson's disease (PD MED): a large, open-label, pragmatic randomised trial. Lancet. 2014; 384(9949): 1196–205. [PubMed: 24928805]

62. Luo N, Ng WY, Lau PN, Au WL, Tan LC. Responsiveness of the EQ-5D and 8-item Parkinson's Disease Questionnaire (PDQ-8) in a 4-year follow-up study. Qual Life Res. 2010; 19(4):565–9. [PubMed: 20204708]

63. Schrag A, Spottke A, Quinn NP, Dodel R. Comparative responsiveness of Parkinson's disease scales to change over time. Mov Disord. 2009; 24(6):813–8. [PubMed: 19199355]

64. Langfitt JT, Vickrey BG, McDermott MP, Messing S, Berg AT, Spencer SS, Sperling MR, Bazil CW, Shinnar S. Validity and responsiveness of generic preference-based HRQOL instruments in chronic epilepsy. Qual Life Res. 2006; 15(5):899–914. [PubMed: 16721649]

65. Janssen MF, Lubetkin EI, Sekhobo JP, Pickard AS. The use of the EQ-5D preference-based health status measure in adults with Type 2 diabetes mellitus. Diabet Med. 2011; 28(4):395–413. [PubMed: 21392061]

66. Riazi A, Cano SJ, Cooper JM, Bradley JL, Schapira AH, Hobart JC. Coordinating outcomes measurement in ataxia research: do some widely used generic rating scales tick the boxes? Mov Disord. 2006; 21(9):1396–403. [PubMed: 16755585]

67. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, McGirt MJ. Accurately measuring outcomes after surgery for adult Chiari I malformation: determining the most valid and responsive instruments. Neurosurgery. 2013; 72(5):820–7. discussion 827. [PubMed: 23381490]

68. Sagberg LM, Jakola AS, Solheim O. Quality of life assessed with EQ-5D in patients undergoing glioma surgery: what is the responsiveness and minimal clinically important difference? Qual Life Res. 2014; 23(5):1427–34. [PubMed: 24318084]

69. Pink J, Petrou S, Williamson E, Williams M, Lamb SE. Properties of patient-reported outcome measures in individuals following acute whiplash injury. Health Qual Life Outcomes. 2014; 12:38.10.1186/1477-7525-12-38.:38-12 [PubMed: 24625124]

70. Tidermark J, Bergström G, Svensson O, Törnkvist H, Ponzer S. Responsiveness of the EuroQol (EQ 5-D) and the SF-36 in elderly patients with displaced femoral neck fractures. Qual Life Res. 2003; 12(8):1069–79. [PubMed: 14651424]

71. Tidermark J, Bergström G. Responsiveness of the EuroQol (EQ-5D) and the Nottingham Health Profile (NHP) in elderly patients with femoral neck fractures. Qual Life Res. 2007; 16(2):321–30. [PubMed: 17033908]

72. Frihagen F, Grotle M, Madsen JE, Wyller TB, Mowinckel P, Nordsletten L. Outcome after femoral neck fractures: a comparison of Harris Hip Score, Eq-5d and Barthel Index. Injury. 2008; 39(10): 1147–56. [PubMed: 18656868]

73. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, McGirt MJ. Accurately measuring the quality and effectiveness of cervical spine surgery in registry efforts: determining the most valid and responsive instruments. Spine J. 201310.1016/j.spinee.2013.07.444

74. Olerud P, Tidermark J, Ponzer S, Ahrengart L, Bergström G. Responsiveness of the EQ-5D in patients with proximal humeral fractures. J Shoulder Elbow Surg. 2011; 20(8):1200–6. [PubMed: 22014617]

75. Parsons N, Griffin XL, Achten J, Costa ML. Outcome assessment after hip fracture: is EQ-5D the answer? Bone Joint Res. 2014; 3(3):69–75. [PubMed: 24648420]

76. Zampelis V, Ornstein E, Franzén H, Atroshi I. A simple visual analog scale for pain is as responsive as the WOMAC, the SF-36, and the EQ-5D in measuring outcomes of revision hip arthroplasty. Acta Orthop. 2014; 85(2):128–32. [PubMed: 24479622]

77. Ostendorf M, van Stel HF, Buskens E, Schrijvers AJ, Marting LN, Verbout AJ, Dhert WJ. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. J Bone Joint Surg Br. 2004; 86(6):801–8. [PubMed: 15330018]

78. Goodwin PC, Ratcliffe J, Morris J, Morrissey MC. Using the knee-specific Hughston Clinic Questionnaire, EQ-5D and SF-6D following arthroscopic partial meniscectomy surgery: a comparison of psychometric properties. Qual Life Res. 2011; 20(9):1437–46. [PubMed: 21547357]

79. Lin FJ, Samp J, Munoz A, Wong PS, Pickard AS. Evaluating change using patient-reported outcome measures in knee replacement: the complementary nature of the EQ-5D index and VAS scores. Eur J Health Econ. 2014; 15(5):489–96. [PubMed: 23765331]

80. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford). 1999; 38(9):870–7. [PubMed: 10515649]

81. Giesinger K, Hamilton DF, Jost B, Holzner B, Giesinger JM. Comparative responsiveness of outcome measures for total knee arthroplasty. Osteoarthritis Cartilage. 2014; 22(2):184–9. [PubMed: 24262431]

82. Dawson J, Boller I, Doll H, Lavis G, Sharp R, Cooke P, Jenkinson C. Responsiveness of the Manchester-Oxford Foot Questionnaire (MOXFQ) compared with AOFAS, SF-36 and EQ-5D assessments following foot or ankle surgery. J Bone Joint Surg Br. 2012; 94(2):215–21. [PubMed: 22323689]

83. Maher AJ, Kilmartin TE. An analysis of Euroqol EQ-5D and Manchester Oxford Foot Questionnaire scores six months following podiatric surgery. J Foot Ankle Res. 201210.1186/1757-1146-5-17

84. DeVine J, Norvell DC, Ecker E, Fourney DR, Vaccaro A, Wang J, Andersson G. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. Spine (Phila Pa 1976). 2011; 36(21 Suppl):S69–74. [PubMed: 21897347]

85. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, Glassman SD, McGirt MJ. Accurately measuring the quality and effectiveness of lumbar surgery in registry efforts: determining the most valid and responsive instruments. Spine J. 2014; 14(12):2885–91. [PubMed: 24768731]

86. Solberg T, Johnsen LG, Nygaard ØP, Grotle M. Can we define success criteria for lumbar disc surgery?: estimates for a substantial amount of improvement in core outcome measures. Acta Orthop. 2013; 84(2):196–201. [PubMed: 23506164]

87. Staerkle RF, Villiger P. Simple questionnaire for assessing core outcomes in inguinal hernia repair. Br J Surg. 2011; 98(1):148–55. [PubMed: 20814965]

88. Lee L, Mata J, Augustin BR, Carli F, Morin N, Latimer E, Feldman LS. A comparison of the validity of two indirect utility instruments as measures of postoperative recovery. J Surg Res. 2014; 190(1):79–86. [PubMed: 24629417]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

89. Tosh J, Brazier J, Evans P, Longworth L. A review of generic preference-based measures of health-related quality of life in visual disorders. Value Health. 2012; 15(1):118–27. [PubMed: 22264979]

90. Ang M, Fenwick E, Wong TY, Lamoureux E, Luo N. Utility of EQ-5D to assess patients undergoing cataract surgery. Optom Vis Sci. 2013; 90(8):861–6. [PubMed: 23851304]

91. Fenwick EK, Xie J, Ratcliffe J, Pesudovs K, Finger RP, Wong TY, Lamoureux EL. The impact of diabetic retinopathy and diabetic macular edema on health-related quality of life in type 1 and type 2 diabetes. Invest Ophthalmol Vis Sci. 2012; 53(2):677–84.10.1167/iovs.11-8992 [PubMed: 22205611]

92. Maes IH, Joore MA, Cima RF, Vlaeyen JW, Anteunis LJ. Assessment of health state in patients with tinnitus: a comparison of the EQ-5D and HUI mark III. Ear Hear. 2011; 32(4):428–35. [PubMed: 21221004]

93. Wu AW, Hanson KA, Harding G, Haider S, Tawadrous M, Khachatryan A, Pashos CL, Simpson KN. Responsiveness of the MOS-HIV and EQ-5D in HIV-infected adults receiving antiretroviral therapies. Health Qual Life Outcomes. 201310.1186/1477-7525-11-42

94. König HH, Ulshöfer A, Gregor M, von Tirpitz C, Reinshagen M, Adler G, Leidl R. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. Eur J Gastroenterol Hepatol. 2002; 14(11):1205–15. [PubMed: 12439115]

95. Stark RG, Reitmeir P, Leidl R, König HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflamm Bowel Dis. 2010; 16(1):42–51. [PubMed: 19475674]

96. Bushnell DM, Martin ML, Ricci JF, Bracco A. Performance of the EQ-5D in patients with irritable bowel syndrome. Value Health. 2006; 9(2):90–7. [PubMed: 16626412]

97. Obradovic M, Lal A, Liedgens H. Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. Health Qual Life Outcomes. 201310.1186/1477-7525-11-110

98. Barton GR, Sach TH, Avery AJ, Doherty M, Jenkinson C, Muir KR. Comparing the performance of the EQ-5D and SF-6D when measuring the benefits of alleviating knee pain. Cost Eff Resour Alloc. 200910.1186/1478-7547-7-12

99. Chapman JR, Norvell DC, Hermsmeyer JT, Bransford RJ, DeVine J, McGirt MJ, Lee MJ. Evaluating common outcomes for measuring treatment success for chronic low back pain. Spine (Phila Pa 1976). 2011; 36(21 Suppl):S54–68. [PubMed: 21952190]

100. Campbell H, Rivero-Arias O, Johnston K, Gray A, Fairbank J, Frost H. UK MRC Spine Stabilization Trial. Responsiveness of objective, disease-specific, and generic outcome measures in patients with chronic low back pain: an assessment for improving, stable, and deteriorating patients. Spine (Phila Pa 1976). 2006; 31(7):815–22. [PubMed: 16582856]

101. Soer R, Reneman MF, Speijer BL, Coppes MH, Vroomen PC. Clinimetric properties of the EuroQol-5D in patients with chronic low back pain. Spine J. 2012; 12(11):1035–9. [PubMed: 23199409]

102. Aggarwal R, Wilke CT, Pickard AS, Vats V, Mikolaitis R, Fogg L, Block JA, Jolly M. Psychometric properties of the EuroQol-5D and Short Form-6D in patients with systemic lupus erythematosus. J Rheumatol. 2009; 36(6):1209–16. [PubMed: 19369452]

103. Coast J, Peters TJ, Richards SH, Gunnell DJ. Use of the EuroQoL among elderly acute care patients. Qual Life Res. 1998; 7(1):1–10. [PubMed: 9481146]

104. Holland R, Smith RD, Harvey I, Swift L, Lenaghan E. Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQoL. Health Econ. 2004; 13(8):793–805. [PubMed: 15322991]

105. The measurement and valuation of health status using EQ-5D: A European perspective. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2003.

106. Feeny, D. Preference-based meausres: utility and quality-adjusted life years. In: Fayers, P.; Hays, R., editors. Assessing quality of life in clinical trials. 2. New York, NY: Oxford University Press; 2005. p. 405-430.

107. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. Pharmacoeconomics. 2000; 17(1):13–35. [PubMed: 10747763]

108. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ. 2004; 13(9):873–884. [PubMed: 15362179]

109. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care. 1989; 27(3 Suppl):S217–32. Review. [PubMed: 2646490]

110. Grosse SD, Chaugule SS, Hay JW. Estimates of utility weights in hemophilia: implications for cost-utility analysis of clotting factor prophylaxis. Expert Rev Pharmacoecon Outcomes Res. 2015:1–17. Epub ahead of print.

111. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006; 4:70. [PubMed: 17005038]

112. Streiner, DL.; Norman, GR. Measuring change. In: Streiner, DL.; Norman, GR., editors. Health measurement scales: A practical guide to their development and use. 4. New York, NY: Oxford University Press; 2008. p. 277-297.

113. Fairclough, DL. Study design and protocol development. In: Fairclough, DL., editor. Design and analysis of quality of life studies in clinical trials. 2. Boca Raton, FL: Chapman & Hall/CRC; 2010. p. 29-52.

114. Gerhards SA, Huibers MJ, Theunissen KA, de Graaf LE, Widdershoven GA, Evers SM. The responsiveness of quality of life utilities to change in depression: a comparison of instruments (SF-6D, EQ-5D, and DFD). Value Health. 2011; 14(5):732–9. [PubMed: 21839412]

115. Lazarus RS. Coping theory and research: past, present, and future. Psychosom Med. 1993; 55(3):234–47. [PubMed: 8346332]

116. Lazarus, RS.; Folkman, S. Stress, appraisal and coping. New York, NY: Springer; 1984.

117. Roesch SC, Adams L, Hines A, Palmores A, Vyas P, Tran C, Pekin S, Vaughn AA. Coping with prostate cancer: a meta-analytic review. J Behav Med. 2005; 28(3):281–93. [PubMed: 16015462]

118. Wonghongkul T, Dechaprom N, Phumivichuvate L, Losawatkul S. Uncertainty appraisal coping and quality of life in breast cancer survivors. Cancer Nurs. 2006; 29(3):250–7. [PubMed: 16783127]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Key Points for Decision Makers

- A quality-adjusted life year (QALY) is a standard metric for cost-effectiveness analysis recommended by the National Institute for Health and Clinical Excellence (NICE) and the U.S. Panel on Cost-Effectiveness in Health and Medicine. The EQ-5D is one of the most frequently used preference-based measures that can produce QALY information and is preferred by NICE.

- It is vitally important to know the responsiveness of the EQ-5D, i.e., the ability to detect health status change. This study provides an extensive systematic review of the responsiveness of the EQ-5D to provide evidence in support or against its use in cost-effectiveness analyses.

- This study found that the EQ-5D is responsive in almost half of the conditions reviewed and not responsive in a small number of conditions (alcohol dependency, schizophrenia, limb reconstruction, and hearing impairment). The other half of the conditions reviewed show mixed evidence of responsiveness, suggesting that researchers need to incorporate condition-specific measures along with the EQ-5D to ensure that appropriate measures of effectiveness are reported.
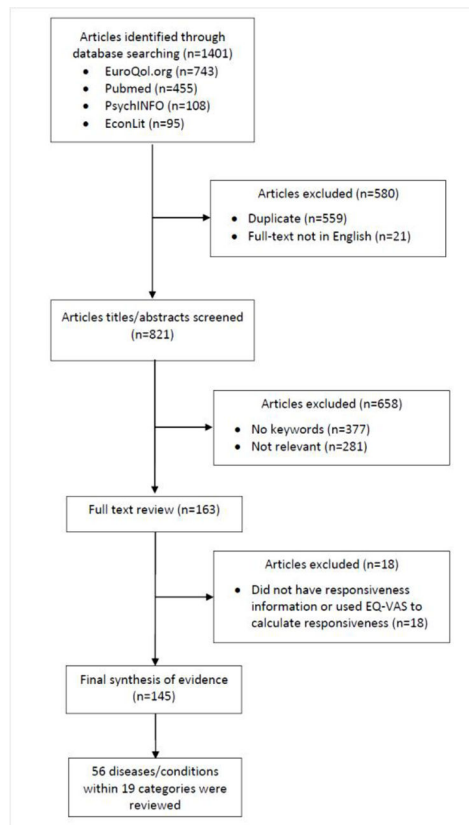
Figure 1.
Study Identification process

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Characteristics of identified studies on responsiveness of the EQ-5D.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
Cardio-vascular Disease	Coronary artery disease	1	Sweden [29]	Chest pain was used as the main outcome. No change was reported on EQ-5D index after 2-year follow-up. Responsiveness statistics was not reported.	EQ-5D is not responsive to change when chest pain is used as an anchor.
	Cardiac rehabilitation	1	Germany [30]	ES was large (0.74) from admission to discharge but marginal (ES=0.09) between discharge and three months follow-up.	EQ-5D is responsive to only a large clinical improvement after the intervention for this condition.
	Heart failure	2	USA and Canada [31;32]	The study used a Global Rating Scale (GRS) and a clinical measure (NYHA) as anchors. They found that the EQ-5D index (US norm) was responsive to change only in the group that reported moderate increase (+2 to +4) on the GRS (ES=0.13; Guyatt's ES=0.15; SRM=0.17) but did not find significant change related to the NYHA (ES=0.06, SRM=0.08), [31]. Another study that used only a distribution method (SRM) did not find any change on EQ-5D index after self-report at baseline (SRM=0.03) [32].	EQ-5D is responsive only in patients that have moderate clinical improvement, based on self-reported GRS.
	Stroke	2	Canada [33] and Germany [34]	Both studies used external anchors to determine responsiveness. EQ-5D index was responsive to improvement post-stroke. ES varied from small to large, depending on the degree of improvement and external anchor used. ES for some improvement or better health state ranged from 0.34 to 1.28, depending on anchor used. ES for large improvement ranged from 1.26 to 1.94. For worse health states, EQ-5D was not responsive, ES of −0.05. Similarly, SRM ranged from 0.73 to 1.24 for some improvement and 1.19 to 1.76 for large improvements. The *t*-statistic ranged from 3.61 to 7.62 for some improvement and 8.44 to 9.21 for large improvement.	EQ-5D is responsive to improvement but not deterioration for stroke patients.
Substance Abuse	Chronic opioid dependence	1	Canada [35]	The study used 20% clinical change on Addiction Severity Index as an external anchor. Mean difference between 12-month end point and baseline was not significant in the response group. However, area under the ROC curve was statistical significant with area under the ROC curve of 0.61.	The EQ-5D is responsive to clinical change in this sample, based on the AUROC.
	Alcohol-dependency	2	Germany [23]; UK [24]	Alcohol consumption was used as an anchor to test responsiveness of the EQ-5D [23]. The group that drastically decreased or increased alcohol consumption had a small but significant changed on the EQ-5D index (ES=0.42, t- statistic= 2.15). However, the change on the EQ-5D index was found	EQ-5D may not be a proper measure for detecting health status change in relation to alcohol dependency.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
				(ES=0.30) in the group who did not have any change on alcohol consumption. (ES=0.30) in the group who did not have any change on alcohol consumption. Another study [24] studied harmful alcohol consumption using web-based survey. Linear regression model was used to predict changes in EQ-5D score at 12 months adjusted for EQ-5D baseline score. However, no significant change in EQ-5D scores was found.	
Mental Disorders	Mental health	1	Systematic review [36]	Mental health conditions included in the review were anxiety, depression, personality disorder, and schizophrenia, based on 7 large clinical trials. The EQ-5D was responsive to common mental health problems such as anxiety, depression and personality disorder (SRM: 0.45 – 0.58) but was less responsive to schizophrenia (SRM=0.02–0.12). Ceiling effects were a prominent problem that impaired ability of the EQ-5D to detect change over time.	EQ-5D is responsive to anxiety, depression and personality disorder but not responsive to schizophrenia.
	Attention Deficit Hyperactivity Disorder (ADHD)	1	Netherlands [28]	The sample in this study was children and adolescents (6 to 18 years old). ES was 0.52 (responder vs. non-responder).	Parent-reported EQ-5D is responsive to ADHD in this sample.
	Anxiety & Depression	1	Systematic review [37]	The EQ-5D is responsive to clinical improvement in both depressed (ES=0.44) and anxious (ES=0.39– 0.99; SRM=0.46–0.54) patients.	The EQ-5D is well established on its responsiveness in this population.
	Schizophrenia	1	Systematic review [25]	There were mixed evidence on responsiveness of the EQ-5D. SRM of the EQ-5D was small (0.39) for an improvement on BPRS of at least 25%. SRMs were smaller in the group with deterioration or improvement on BPRS <25% (SRM=0.17 and 0.05, respectively). Another study found that EQ-5D index did not respond to changes in most symptom or functioning measures. However, one study reported a large ES of 1.13 for olanzapine treated patient before and after treatment and a moderate to large effect size of 0.78 to 0.96 for other antipsychotics.	Using the EQ-5D in patients with schizophrenia is not encourage due to mixed evidence on its responsiveness.
	Social phobia	1	Germany [38]	The study used external anchor to assess EQ-5D responsiveness. ES for the improvement group at 1-year and 1.5-year follow-up were 0.50 and 0.64 (using UK weights) with the ES being reduced to half these values when using German weights. However, there was no change on the EQ-5D index among the deterioration group.	The EQ-5D is responsive to improvement but not deterioration in this sample.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
Musculo-skeletal Disease	Somatoform disorders	1	Germany [39]	The EQ-5D was responsive only in the group of patients who had clinical improvement after 1- year follow up. The ES was small but significant (ES=39 from EQ-5D UK weight; ES=26 from EQ-5D German weight). The ES for subjects that got worse at follow up was not significant (ES= −0.18 from EQ-5D UK weight; ES= −0.22 from EQ-5D index German).	The EQ-5D is responsive to improvement but not deterioration in this sample.
	Inflammatory arthritis	3	UK [40], Sweden [41;42]	In the first study, the EQ-5D VAS was used to categorize patients into improved and deteriorated groups, compared to baseline. ES was much smaller (0.22) in the improved group, compared to the deteriorated group (0.62) [40]. Another study showed that the EQ-5D was responsive to improved health to a similar degree (SRM=0.2) after a 1-year follow up [41]. The last study categorized patients into improved and deteriorated between baseline and follow-up at 6 months, 12 months and 24 months based on disease activity. They found that the EQ-5D was responsive to clinical improvement with a moderate to large ES and SRM ranging from 0.50 to 0.61 after a follow- up of 6 months to 2 years. However, the ES and SRM was small in the deteriorated group ranging from 0.07 to −0.20 at follow-up.	The EQ-5D is responsive to clinical change in patients with arthritis, in both improved and deteriorated groups. It is more responsive in patients that had a great improvement.
	Rheumatoid arthritis	4	UK [43;44], Denmark [45], Netherlands [46]	In the first study, patients reporting improvement over 3 months had a SRM of 0.70 [43]. In the second study of rehabilitation, ES or SRM is not reported due to no change in EQ-5D index [44]. The study from Denmark did not report responsiveness statistics but concluded that the EQ-5D was able to detect changes in health status [45]. In the final study, among subjects that perceived improvement over 1 year, SRM was 0.76 while it was 0.18 among non- improvers (AUC=0.72).	EQ-5D is responsive to clinical improvement but not deterioration in patients with rheumatoid arthritis.
	Ankylosing spondylitis	1	UK [47]	The SRM was relatively small (0.25) for the group that transitioned to better health at 6 months while the group that transitioned to worse health had an even smaller SRM (−0.11).	The evidence for the EQ-5D in this condition is limited.
	Musculoskeletal disease	1	Canada [48]	The study included patients with rheumatoid arthritis, low back pain, fibromyalgia, and psoriatic arthritis. ES was reported at 0.53 for the improved group and −0.58 for the group with worse health.	The EQ-5D is responsive to patients with arthritis, low back pain and fibromyalgia. However, researchers should consider using the EQ-5D with other condition-specific measures (CSM) along with the EQ-5D since its responsiveness may depend on the severity of the conditions.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
Respiratory Conditions	Asthma & COPD	2	Systematic review [16], Japan [49]	A systematic review of 16 asthma studies and 12 COPD studies identified few studies with responsiveness information [16]. For asthma, the EQ-5D was less sensitive in patients with mild condition. The ES was small in magnitude (<0.5) for improvement in asthma over a 3–6 month follow-up period. For COPD, the ES was reported in a moderate magnitude (ES=−0.55) for patients whose global rating improved. For both patient populations, the EQ-5D was not responsive for patients who reported a worse global change rating (ES=−0.07). A small study in Japan reported a small ES (0.37–0.41) for a 3–6 month follow-up period [49] for asthma patients.	The EQ-5D can be used in moderate-to-severe asthma and COPD patients. Patients with mild asthma, researchers may encounter low responsiveness of the EQ-5D to detect changes and also ceiling effects.
	COPD	1	Systematic review [50]	The EQ-5D has limited responsiveness for patients with stable COPD states who have moderate to no change with treatment. In patients with COPD exacerbation states, the EQ-5D is responsive to large changes over 2–10 week period (changes on EQ-5D index from 0.68 to 0.80, p<0.001 at 2 weeks and 0.79, p<0.001 at 10 weeks). No responsiveness statistics were reported.	The EQ-5D is responsive to change in COPD patients who have moderate-to-severe exacerbation. Researchers should consider using CSMs along with EQ-5D.
	Chronic bronchitis	1	UK [51]	The ES was small in magnitude (0.38).	The evidence for EQ-5D responsiveness is limited due to small sample size and small ES detected for this sample. Researchers should consider using CSMs along with EQ-5D.
Cancer	Benign prostatic hypertrophy	2	UK [52;53]	The study compared two different surgical interventions and found no significant change on EQ-5D [52]. However, this study did not have an external outcome measure apart from the EQ-5D. The second study found significant changes on the EQ-5D after a 6-month follow up period. However, they did not calculate ES.	Responsiveness of the EQ-5D is limited. More evidence is needed.
	Prostate cancer	1	Canada [54]	The EQ-5D is fairly responsive to changes in health status, both worse and better, to a similar degree. The area under the ROC curve ranged from 0.62–0.66 for the group that got worse and 0.65–0.67 for the group that got better at 2 months and 12 months after treatment.	Based on the area under the ROC curve, the EQ-5D is responsive to change in patients with prostate cancer. However, there is no information on effect size. Thus, researchers should consider using CSMs along with EQ-5D.
	Breast cancer	3	Netherlands [55], Singapore [56], structured review [15]	The EQ-5D was responsive in detecting moderate to large change in health status (both deterioration and improvement) at 12 month follow-up with a SRM of −0.52 and 0.62 respectively. The groups that reported smaller changes (deterioration and improvement) at	The EQ-5D is responsive to moderate-to-large change in patients with breast cancer.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
				follow-up had an SRM of −0.05 and 0.16 respectively [... ES of 0.69 and 0.52 ... were reported for subjects that had a change in performance status and quality of life respectively. [56]. The ES of the EQ-5D for patients with health status changes were in a moderate magnitude [15].	
	Liver metastases	1	Netherlands [57]	The EQ-5D was responsive to a moderate to large change of health status after surgery with an ES of −1.33 at 2-week and an ES of −0.58 at 3-month follow up periods.	The EQ-5D is responsive to moderate-to-large change in health status. It can be used to measure changes in a short-term follow-up period up to 3 months after surgery in this sample.
	Multiple myeloma	1	Norway [58]	The EQ-5D was responsive to change in health status, both improvement (SRM=0.43) and deterioration (SRM=0.45), after a 3-month follow-up period.	The EQ-5D is responsive to change in health status in this condition.
Urinary Disorders	Urinary incontinence	1	Systematic review [59]	There was mixed evidence reported in the systematic review on EQ-5D responsiveness. Only one out of 8 studies reported SRM for the EQ-5D. The magnitude was small (SRM=0.26) for patients reporting improvement at 5-month follow up. There were several CSMs used across different studies.	Using the EQ-5D in clinical trials for this sample should be accompanied with validated CSMs. Responsiveness evidence of the EQ-5D is limited.
Neurological Disorders	Dementia	1	Germany [60]	The EQ-5D was more responsive to patients with a worse health status at 1 year follow-up (ES=0.41) than patients reporting better health status at follow-up (ES=0.12) in people with mild-to-moderate dementia.	The EQ-5D is responsive to health status change in this condition.
	Parkinson's disease	3	Multi-country [61], Singapore [62], Germany & UK [63]	No anchor was used to evaluate responsiveness of the EQ-5D. All three studies evaluated responsiveness using changes on the EQ-5D index from baseline to the follow-up period. The randomized trial reported no change on the EQ-5D scores after 3-year follow up period while there was slight improvement on CSM [61]. However, the improvement on CSM did not reach statistical significance. The other two studies were secondary analyses of existing data. Both studies reported that the EQ-5D was responsive to change over a 4-year follow-up period. However, there was discrepancy between the ES reported. The study from Singapore reported a much greater ES (ES=1.06) and a SRM of 0.63, compared to another study which reported an ES of −0.19 and a SRM of −0.24.	The evidence of responsiveness of the EQ-5D for this condition is mixed and inconsistent. Although one study reported ES of the EQ-5D in a large magnitude, the study sample size was small (n=31) and they did not use any external anchor. Different population weights used such as German vs. UK could result in different ES. More research is needed in this condition.
	Epilepsy	1	US [64]	The EQ-5D is responsive to change in health status (2-year follow-up). The reported ES using UK and US weights was 0.35 while the SRM was 0.30 and 0.29 respectively.	The EQ-5D is responsive to change in this condition.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
Endocrine System	Diabetes (Type 2)	1	Systematic review [65]	Responsiveness of the EQ-5D was reported in 7 studies and evaluated using various statistical analyses. Six studies showed that the EQ-5D was responsive to treatment and health status change. Only one study investigating the long-term management of diabetic peripheral neuropathic pain did not see any change on the EQ-5D index. The systematic review did not report ES or SRM but concluded that the EQ-5D is responsive for patients with type 2 diabetes based on significant statistical changes on EQ-5D health index overtime after treatments.	The EQ-5D is responsive to change in this condition.
Genetic Disorders	Friedreich's ataxia	1	UK [66]	The ES reported in this study for the EQ-5D was small (ES=0.13) and insignificant after 1-year follow up.	Evidence is insufficient for responsiveness of the EQ-5D in this rare condition.
Surgery	Brain	2	US [67], Norway [68]	The study was conducted in patients with Chiari I (suboccipital area), the ES of the EQ-5D was large for responders (improvement of health after the surgery) with SRM of 1.78 but small for non-responder group with SRM of 0.15 [67]. On the contrary, the study conducted for patients with gliomas, the ES of the EQ-5D was large for deterioration of health (SRM=0.72) but was not significant for the improved group (SRM=0.04) [68].	The EQ-5D responsiveness is different for different type of brain conditions and type of surgery. Other CSM should be used along with the EQ-5D for detecting change in health status.
	Acute whiplash associated disorder	1	UK [69]	The EQ-5D was responsive for patients reporting better health status change with a small to moderate ES. From baseline to follow-up at 12 months, the group of patients got worse had SRM range of −0.17 to −0.97 and ES range of −0.18 to −0.84. The ranges of SRM and ES for the group of patients that got better were 0.39–0.56 and 0.42–0.58, respectively.	The EQ-5D can be used to monitor health status change for this sample.
	Neck	4	Sweden [70;71], Norway [72], UK [73]	The SRM of the EQ-5D was reported as moderate-to-large magnitude (SRM of 0.46 to 1.14). The area under the ROC curve ranged from 0.68 to 0.76.	All studies consistently reported responsiveness of the EQ-5D in this sample. The EQ-5D is responsive to monitor health improvement after surgery for neck and cervical spine injury.
	Arm (Proximal humeral)	1	Sweden [74]	Despite the small sample size, the study showed that the EQ-5D was responsive to both improvement and deterioration between the 4 month and 12 month follow-up. For subjects that clearly improved, SRM was 0.67 while those that clearly deteriorated, SRM was −0.72. Those who had marginally improved or marginally deteriorated, SRM was small (0.04). The area under the ROC curve ranged from 0.71 to 0.81 when comparing the four subgroups of patients	The EQ-5D is responsive to changes of health after surgery for proximal humeral fractures.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
				(clearly improved, marginally improved, marginally improved, marginally deteriorated, clearly deteriorated). (clearly improved, marginally improved, marginally improved, marginally deteriorated, clearly deteriorated).	
	Hip	3	UK [75], Sweden [76], Netherlands [77]	The ES of the EQ-5D was reported at 0.67 after 1-month follow-up, 0.30 after 4 month follow-up, and 0.27 after 12 month follow-up [75]. The ES was reported at 1.2–1.3 for 1–2 year follow up period [76;77].	All studies consistently reported responsiveness of the EQ-5D in this sample. The ES size seems to be varied, depending on follow-up period. The EQ-5D is sensitive to monitor health improvement after surgery for hip replacement.
	Knee	4	UK [78–80], Switzerland [81]	The ES of the EQ-5D was reported at 0.9 for 6- week follow-up period. The ES ranged from 1.2 to 1.4 for 6-to-24-month follow-up periods [78–80]. The SRM ranged from 0.56 to 1.2 after 6- to-24-month follow-up [81].	All studies consistently reported responsiveness of the EQ-5D in this sample. The EQ-5D is sensitive to monitor health improvement after surgery for knee replacement.
	Foot/ankle	2	UK [82;83]	The ES of the EQ-5D were reported at 0.83 and 0.90 after 6-month and 9-month follow up periods [82;83].	Both studies reported that the EQ-5D was responsive among patients who received surgeries for foot/ankle problems.
	Limb reconstruction	1	UK [26]	The study reported that the EQ-5D did not detect health differences correctly among three groups of patients between baseline and the 1-year follow up (SRM=0.40 for the better health group, SRM=0.74 for the same health group, and SRM=−0.07 for the worse health group).	The EQ-5D is not responsive to health status change after limb reconstruction surgery.
	Low back	1	Systematic review [84]	Only two studies included in the systematic review reported EQ-5D responsiveness. The pooled mean ES was 0.78 but it did not reach statistical significance due to a small sample size.	The evidence of EQ-5D responsiveness for patients receiving surgery for low back pain is limited. Researchers should consider using CSMs along with the EQ-5D.
	Lumbar	2	US [85], Norway [86]	The AUCs of the EQ-5D were 0.80–0.97. The EQ-5D was responsive to health improvement among patients who received lumbar surgery from baseline to 1-to-2 years follow-up.	The EQ-5D is sensitive to health improvement of patients undergoing lumbar surgery.
	Inguinal hernia	1	Switzerland [87]	Based on a sample size of 51, the reported SRM for the EQ-5D was 0.65.	Due to a small sample size, more research is warranted to support

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
					responsiveness of the EQ-5D in this sample. responsiveness of the EQ-5D in this sample.
	Colorectal	1	Canada [88]	The EQ-5D index from patients with complication was significantly lower than patients without complication (SRM=−0.21). However, SRM calculated from baseline to 4 week and 8 week follow-up were small and insignificant (−0.10 and −0.01, respectively).	There is limited evidence on responsiveness of the EQ-5D in patients receiving colorectal surgery. More study is warranted.
Vision	Visual disorders	1	Systematic review [89]	Only a few studies on post cataract surgery and age-macular degeneration reported EQ-5D responsiveness. However, no responsiveness statistics was provided. There was no responsiveness available for glaucoma, diabetic retinopathy, conjunctivitis and other visual disorders.	The evidence of responsiveness for the EQ-5D is limited for patients with post cataract surgery and macular degeneration. Since there is no information on a magnitude of the responsiveness, researchers should consider using CSMs along with the EQ-5D for the above conditions.
	Post cataract surgery	1	Singapore [90]	The prospective cohort study reported that the EQ-5D was responsive to health improvement after cataract surgery with a calculated ES of 0.35.	The ES of the EQ-5D is small. Again, CSM is recommended to use along with the EQ-5D.
	Diabetic retinopathy and diabetic macular edema	1	Australia [91]	The EQ-5D index was not different between patients who had and did not have diabetic retinopathy or diabetic macular edema. No responsiveness statistics was reported.	There is no evidence on responsiveness in this population.
Hearing	Hearing impairment	1	Review study [27]	The EQ-5D was not sensitive to detect changes after different hearing interventions (ES=0.05–0.06).	The EQ-5D is not sensitive to detect changes among patients with hearing impairment.
	Tinnitus	1	Netherlands [92]	The EQ-5D was responsive to health improvement after treating tinnitus but magnitude of response was small (ES=0.19; SRM=0.22, the area under the curve=0.58).	There is limited evidence of responsiveness of the EQ-5D for patients with tinnitus.
Infection	HIV	1	Systematic review [93]	The EQ-5D was responsive to health changes among patients with HIV. The ES ranged between 0.3–0.5 on each of its five dimensions. The ES of −0.42 and −0.05 were reported for patients with and without adverse events.	The EQ-5D is responsive to health status change due to adverse events in this sample.
Gastro-intestinal system	Inflammatory bowel disease	3	Germany [94;95], UK & Spain & Germany [96]	All three studies consistently reported that the EQ-5D was responsive to health status changes in this sample. The ES ranged from 0.29 to 0.33 and SRM of 0.67 for health improvement and SRM of −0.53 for health deterioration.	The EQ-5D is responsive to health changes in this sample.

Category	Condition/Disease	Number of Studies	Study Location	Reported Responsiveness	Conclusion[a]
Pain	Chronic pain	1	Germany [97]	Responsiveness of the EQ-5D was moderate to high in magnitude for chronic pain that increased with pain severity. SRMs were reported at of 0.65, 1.33 and 2.02 based on baseline WOMAC category of mild to moderate, moderate to severe and severe to extreme.	The EQ-5D is responsive for pain treatment.
	Knee pain	1	UK [98]	The mean EQ-5D was significant for patients that had improved <20% (p<0.01) and patients that improved >20% (p<0.001).	The EQ-5D is responsive in this population.
	Chronic low back pain	3	Systematic review [99], UK [100], Netherlands [101]	The systematic review published in 2011 did not find responsiveness evidence of the EQ-5D in this sample [99]. The other two studies reported responsiveness of the EQ-5D. Patients with improved health from baseline to 12- month follow-up had an ES of 0.97 and a SRM of 1.02 while those with deteriorated health had an ES and SRM of −0.34. However, patients with stable health had an ES of 0.40 and SRM of 0.39. The area under the ROC curve ranged from 0.59 to 0.79.	Responsiveness of the EQ-5D is inconclusive due to limited evidence.
Skin	Skin conditions	1	Systematic review [14]	Plaque psoriasis and psoriatic arthritis conditions have evidence of responsiveness for the EQ-5D with the mean score change of 0.17, p< 0.05. Acne has a moderate ES of 0.44 – 0.53. There was no responsiveness evidence on Hidradenitis suppurativa and hand eczema. Venous leg ulcers has mixed evidence on responsiveness. In the healed group, ES ranged from 0.31–0.50. In the unhealed group, ES ranged from 0.10–0.16. However, there was no significant difference on the EQ-5D index between the two groups.	The EQ-5D is responsive to change in plaque psoriasis and psoriatic arthritis.
Autoimmune System	Systematic lupus erythematosus	1	US [102]	There was no change of the EQ-5D after follow-up.	There is no evidence of responsiveness of the EQ-5D in this sample.
Elderly	Acute care	1	UK [103]	The mean difference in EQ-5D scores between baseline and 4 weeks follow-up were −0.046 for patients with stroke, 0.110 for patients with fractured neck of femur, 0.130 for patients with elective total hip replacement and 0.310 for patients with elective total knee replacement. No t-statistics were reported.	For acute conditions (fractured neck of femur, hip fracture, and knee fracture), surgery appeared to improve health status of elderly patients, based on mean EQ-5D change. However, the sample size was relatively small, more research is needed to evaluate EQ-5D responsiveness in elderly populations.
	Home-based medication review	1	UK [104]	The study reported an ES of 0.55 for the EQ-5D at 6-month follow-up compared to baseline.	The evidence of EQ-5D responsiveness is limited for this intervention.

[a]Based solely on the authors' opinion, therefore, they may or may not reflect the original study's conclusion.

Table 2

A summary of EQ-5D responsiveness in different conditions.[a]

Condition	Evidence of Responsiveness		Reference
	Health Improvement	Health Deterioration	
Cardiac rehabilitation	Yes (Large)	No	[30]
Heart failure	Yes (Small)	No	[31;32]
Stroke	Yes (Small to Large)	No	[33;34]
Chronic opioid dependence	Yes[b]	NA	[35]
Mental Health	Yes (Small to Moderate)	NA	[36]
Attention Deficit Hyperactivity Disorder	Yes (Moderate)	NA	[28]
Anxiety & Depression	Yes (Small to Large)	NA	[37]
Social phobia	Yes (Moderate)	No	[38]
Somatoform disorder	Yes (Small)	No	[39]
Inflammatory arthritis	Yes (Small to Moderate)	Yes (Small)	[40–42]
Rheumatoid arthritis	Yes (Moderate)	No	[43–46]
Asthma	Yes (Small)	No	[16;49]
COPD	Yes (Moderate)	No	[16;50]
Prostate cancer	Yes[b]	Yes[b]	[54]
Breast cancer	Yes (Moderate)	Yes (Moderate)	[15;55;56]
Liver metastases	No	Yes (Moderate to large)	[57]
Multiple myeloma	Yes (Small)	Yes (Small)	[58]
Dementia	Yes (Small)	Yes (Small)	[60]
Epilepsy	Yes (Small)	NA	[64]
Diabetes (Type 2)	Yes[c]	NA	[65]
Surgery (brain, acute whiplash, neck, arm, hip, knee, foot/ankle, lumbar)	Yes (Varies by location/type of surgery)	Yes (Varies by location/type of surgery)	[67–83;85;86]
Post cataract surgery	Yes (Small)	NA	[90]
HIV	Yes (Small)	NA	[93]
Pain	Yes (Varies by type of conditions and treatments)	No	[97–101]
Skin conditions (psoriasis, acne)	Yes (Small to Moderate)	NA	[14]

Yes – The EQ-5D is responsive to health changes; No – The EQ-5D is not responsive to health changes; "NA" – Evidence of responsiveness not available

[a]The summary is based solely on the authors' opinion.

[b] Area under the receiver operator characteristic curve was between 0.6 and 0.69 (sufficient evidence of responsiveness).

[c] No responsiveness statistics was reported.