Published in final edited form as:

Pharmacoeconomics. 2007 ; 25(2): 93–106.

Measuring Preferences for Cost-Utility Analysis: How Choice of Method May Influence Decision-Making

Christine M. McDonough, PT, MS and Anna N. A. Tosteson, ScD

Multidisciplinary Clinical Research Center in Musculoskeletal Diseases and the Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Lebanon, NH

Abstract

Preferences for health are required when the economic value of health care interventions are assessed within the framework of cost-utility analysis. The objective of this paper is to review alternative methods for preference measurement and to evaluate the extent to which method may affect health care decision making. Two broad approaches to preference measurement that provide societal health state values are considered: 1) direct measurement and 2) preference-based health state classification systems. Empirical evidence comparing systems and decision making impact suggests that preferences will have the greatest impact on economic analyses when chronic conditions or long-term sequelae are involved. At present, there is no clearly superior method and further study of cost-effectiveness ratios from alternative systems is needed to evaluate system performance. Among studies that compared alternative preference-based systems, EQ-5D tended to provide larger change scores and more favorable cost-effectiveness ratios than HUI2 and 3, while SF-6D provided smaller change scores and less favorable ratios than the other systems. However, these patterns may not hold for all applications. Although there is some evidence that incremental cost-effectiveness ratio (ICER) thresholds (e.g., \$25,000 per QALY gained) are used in decision making, they are not strictly applied and are complemented by consideration of legal, regulatory, and social equity issues. Nonetheless, as ICERs rise, the probability of acceptance of a new therapy is likely to decrease, making the differences in QALYs obtained using alternative methods potentially meaningful. It is imperative that those conducting cost-utility analyses characterize the impact that uncertainty in health state values has on the economic value of the interventions studied. Consistent reporting of such analyses would provide further insight into the policy implications of preference measurement.

Introduction

As pressures to contain costs of medical care have escalated, cost-utility analysis (CUA) has received both critical acclaim and scrutiny as a methodology to inform decision makers regarding the economic value of health care interventions.[1-6] The number of published CUAs has steadily grown over the past 20 years,[3] and use by policy-makers appears to be increasing.[7,8] In some jurisdictions, including the UK, Australia, and the Canadian provinces of Ontario and British Columbia, a formal role for CUA in pharmaceutical coverage decisions is mandated.[7,9] In other jurisdictions, such as the US, explicit use of CUA is more limited.[1,10-13]

A key source of initial resistance to CUA was concern about the validity and comparability of results between studies.[1,7,14,15] Issues of comparability have the potential to

undermine a fundamental strength of CUA, which is meant to facilitate valid economic comparisons across a wide spectrum of interventions. Studies highlighting discrepancies between methods[15-22] underscore the importance of understanding the potential impact that methodological differences may have on decision making.

Development of a Reference Case by the U.S. Panel on Cost-Effectiveness in Health and Medicine[15] provided one methodological standard for CUA. The perspective recommended for the Reference Case is societal, and the methods for valuing health outcomes include use of a generic health state classification system and community preferences, with sensitivity analysis to include patient preferences for studies of specific conditions.[15] Since publication of the Reference Case, there is evidence of improving quality in CUA methodology and reporting.[3] Support for use of reference case criteria to improve comparability is indicated by recent recommendations within the US regulatory environment.[23] Formal guidelines in other jurisdictions demonstrate that although differences exist, there may be emerging consensus on key points noted in the Reference Case[24,25] Despite these positive developments, the extent to which methodological differences in CUA may affect policy decisions remains uncertain. In this paper, we explore how alternative methods that fulfill Reference Case criteria for preference measurement, thereby yielding “societal health state values”, may affect decision making. We describe potential sources of variation in societal health state values, highlight relevant studies, and discuss research assessing decision making impact.

Methods for Estimating Societal Health State Values

The quality-adjusted life year (QALY) is the most commonly used measure for health in CUA.[26] QALYs combine the attributes of length and quality of life into a single measure. The length of time in each health state is weighted according to an associated “health state value”, on a scale anchored at 1 representing best imaginable health and 0 representing death. In our discussion, we use “preference” as a general term reflecting a health state's desirability and “health state value” (HSV) to connote the numerical strength of preference for a health state.

We consider two general approaches to HSV measurement: 1) direct preference elicitation for relevant health states and 2) preference-based health state classification systems.[27-29]

Direct Preference Measurement

The standard gamble, time tradeoff, and rating scale are commonly used preference elicitation methods with unique characteristics.[15,30-32] Differences between preferences obtained by these methods are well-documented,[21,33-35] with HSVs typically highest for standard gamble and lowest for rating scale.[36-38] A recent study comparing elicitation methods for EQ-5D and SF-6D concluded that this component alone can contribute to differences in HSV up to 0.31, and may impact on cost-effectiveness ratios.[39] One study of dialysis patients provided evidence that patient preferences from the standard gamble resulted in cost-effectiveness ratios that were higher by \$5,916 per QALY than when time tradeoff values were used. Although this is one example of how measurement approach may affect cost-effectiveness results, individual patient rather than societal HSVs were used.[21]

Direct measurement of societal HSVs is a resource intensive endeavor, requiring development of relevant health state descriptions, and access to a representative population sample. Since this is often not feasible, researchers may use condition, age and sex-specific values from a published source[40]

Preference-Based Health State Classification Systems

Preference-based health state classification systems define each respondent's health state based on a questionnaire and assign a societal HSV with a scoring algorithm that incorporates preferences from a general population sample. This approach allows researchers to use societal HSVs with minimal resources compared to direct preference measurement.

The most widely used systems[19] include the EuroQol 5D (EQ-5D),[41-43] the Health Utilities Index (HUI),[44,45] the Quality of Well-Being Scale (QWB),[46] and the SF-36-derived SF-6D.[47,48] Three basic steps in developing a system are 1) classifying health, 2) eliciting population preferences for a subset of health states, and 3) developing a scoring algorithm to assign values for the full range of health states. We address potential sources of variation and related research for each step.

Descriptive System

A generic self-report health status questionnaire is the basis for most systems. Most systems include attributes for pain, physical function, social or role function, and anxiety/depression, but differ in the number of response levels and how these are described and weighted. Some systems include other attributes, such as hearing and vision. Other differences include the perspective used in assessing health status. For example, HUI questions ask about respondents' functional capacity,[44,49,50] while EQ-5D and others ask about actual performance. The reference period also varies among systems, from “today” to “over the past four weeks”.

Each instrument characterizes a unique number of health states based on the numbers and levels of attributes included in the questionnaire. It is unknown how many health states are needed to describe health adequately, however, EQ-5D has 243 health states, HUI2 has 24,000 and HUI3 has 972,000.

Taken together, these aspects of the descriptive systems would have an impact on psychometric properties of the system, including the ability to measure health status numerically, and to detect meaningful changes in health. Inadequacy in the descriptive systems may result in ceiling and floor effects, inability to measure key attributes of health (validity), and inability to measure important change (responsiveness).

Preference Measurement

In the development of preference-based systems, a subset of the unique system-defined health states are valued by a sample of the population as the basis for estimating values for the full range of health states. Choice of direct preference elicitation method implicitly incorporates differences noted earlier into the estimation of each system's societal HSVs. EQ-5D offers value sets based on both time tradeoff and rating scale measurements.[51] The SF-6D employs the standard gamble[47,48] and the QWB uses category scaling[46,52] methods. The developers of HUI and others, for example, argued that the rating scale is not appropriate for use in CUA. They used the standard gamble and a transformation of rating scale preference measurements for HUI2 and HUI3.[36] Debate continues about the merits of various elicitation methods[53] and power transformation.[54,55]

Comparing CUA results using alternative preference measurement methods for the same system can demonstrate the isolated impact of preference elicitation method. Conner-Spady[56] compared HSVs from EQ-5D using the time tradeoff and visual analogue scale (VAS) in 436 joint replacement patients, and found a lower baseline mean, wider range, and more negative HSVs, and larger QALY gains for time tradeoff than VAS-based value sets,

indicating a more favorable cost-effectiveness ratio for EQ-5D when using TTO relative to VAS. For example, the QALY gain reported for EQ-5D with TTO-based preference weights was 5.14, compared to 3.64 for the VAS using a 10-year time horizon. If the cost of joint replacement were \$7,000, [57] the ICER would vary only slightly across methods (from \$1,362 to \$1,923 per QALY gained) and be unlikely to influence decision making.

Source of Community Preferences

The representativeness of population samples varies among systems.[58] Furthermore, eliciting and scoring population preferences for health state classification systems are resource-intensive, and HSVs are not available for every population. Therefore, preferences from a different population than that of interest are sometimes used. The extent to which differences in population preferences would contribute to variation in CUA results has been explored. In a study comparing rating scale valuations of EQ-5D health states by Finnish and US general population samples, small differences were noted that the authors concluded would not impact EQ-5D HSVs in international studies.[59] A study comparing TTO HSVs between general US and UK population samples[60] reported higher HSVs for the US sample. Differences in EQ-5D TTO HSVs between UK and Spanish populations,[61] suggest that cultural differences may influence health state valuation.

Scoring Methods

Using directly measured preferences from a sample of the general population, a statistical model is fit to estimate HSVs for the remaining health states. The HUI scoring system is based on multiattribute utility theory[62] and a multiplicative function that captures interactions among attributes and allows characterization of single attribute utility functions for levels within each attribute.[36,63] EQ-5D with York preference weights uses an additive model that includes level of severity, movement away from perfect health, and a term (the N3 term) to account for interaction between attributes when any attribute is at the worst level.[42,43] Scoring algorithms have been developed for EQ-5D for various populations. U.S. TTO-based preference weights are now available for the EQ-5D (EQ-5D-US) using a random effects model.[64] SF-6D scoring is very similar to that of EQ-5D N3 model, with a smaller decrement when any attribute is at its worst level.[65] QWB utilizes an additive function that does not allow for interaction among the attributes.[46] The range of each scale and whether they include states worse than death (e.g. negative scores) vary.

Few studies compared CUA results using different scoring algorithms for the same system. Conner-Spady[56] assessed the effect of the N3 term on EQ-5D HSVs, and reported that it resulted in lower baseline mean and effect size for values, and more QALYs gained in a sample of joint replacement patients.

Empirical Evidence

We searched the international published literature (See Appendix for search terms), and found no studies that directly investigated impact of preference measurement method on policy decisions. Though a systematic review of policy decisions was beyond the scope of this paper, we included all identified studies that provided head-to-head comparisons of the most commonly used preference-based system or addressed impact of system choice.

Comparisons of Preference-Based Health State Classification Systems

A review of 23 published cost-utility analyses conducted alongside clinical trials[19] found that 20 studies utilized a preference-based health state classification system to estimate QALYs. HUI and EQ-5D were the most commonly used systems, with 16 using EQ-5D.

The authors suggested that different systems could qualitatively impact CUA results, and called for greater reporting transparency.

Table I summarizes cross-sectional comparisons between systems. These studies found varying differences in mean HSVs across systems and offer insight into system characteristics that may contribute to variation in HSVs.[40,66-83] Generally, correlations between values from alternative systems were moderate to strong, indicating that they measure the same construct. SF-6D demonstrated floor effects, and a limited range of available scores. EQ-5D did not provide HSVs between .88 and 1, and provided lower HSVs for similar health states from other systems. HUI3 was limited in characterizing diminished mobility other than ambulation. These characteristics may be important for instrument choice relative to the condition of interest.

Evidence of differences in HSVs from cross-sectional studies is reason for concern, but longitudinal studies are necessary to understand system performance when measuring change in health. Fewer studies[56,76,84-98] report longitudinal head-to-head comparisons of preference-based systems (Table II). EQ-5D estimates were generally largest, followed by HUI3, HUI2, and finally SF-6D. Interactions have been noted between HSVs and level of improvement.[88] Pickard et al reported that change in SF-6D HSVs correlated with mental status, while change in HUI and EQ-5D related more to daily function and disability.[84] No consistent pattern of correlation was evident in the studies we reviewed.[84,92]

Comparisons of ICERs obtained using alternative systems were less common. Thomas et al[99] reported CUA results for chronic low back pain acupuncture treatment using SF-6D and EQ-5D. The ICER for SF-6D was £4,241(95% CI 191-28,026) compared to £3,598 (95% CI 189-22,035) for EQ-5D. Neumann et al[94] compared the results of a cost-effectiveness model for Alzheimer's drug treatment using HUI2 and HUI3, and found lower mean utility scores using HUI3 versus HUI2, resulting in ICERs of \$9,000/QALY for HUI3 and \$11,000/QALY for HUI2 with duration of drug effect of 18 months. The authors noted that while the difference in ICERs was slight for this analysis of drug treatment, it could be substantial for a disease prevention drug..

Overall, EQ-5D tended to have larger changes in HSVs, which generally translate to more favorable cost-effectiveness ratios when using EQ-5D compared to HUI2 or HUI3. Similarly, in studies that included EQ-5D and SF-6D, EQ-5D had larger changes. Comparing HUI2 or HUI3 and SF-6D, SF-6D tended to have smaller changes. These patterns are generally consistent with the results of cross-sectional comparisons, yet it is difficult to identify a superior system.

Potential Impact of System Choice on Decision Making

Recent commentaries illustrate the concern that both lack of data on preferences and variation in methods will impact health care decisions[100,101]. We found no empiric studies that address this question by directly investigating policy decisions relative to different measurement methods. Studies addressing the potential influence of economic evaluation on policy decisions have considered thresholds for favorable cost-effectiveness to characterize qualitative changes in study results. Chapman et al[17] assessed the affect of quality adjustment on analyses reporting both life years and QALYs published before 1998 and found qualitatively similar results for both QALYs and life years for the majority of analyses. However, quality adjustment had an important impact, causing estimates to cross thresholds for dominance of \$50,000, or \$100,000, in 18% of cases. This could be interpreted as evidence that in the majority of cases, method choice would not be likely to have a qualitative impact on the results of the studies, and therefore on downstream policy decisions. Use of QALYs had the greatest impact for analyses assessing palliative

treatments, chronic diseases, and situations where there may be negative long-term side-effects. This may indicate that for this subset of diseases, differences in estimated HSVs carried over a lifetime may have a qualitative impact on economic evaluation results and related policy decisions.

Using the same database, Schackman et al[18] examined the results of sensitivity analyses on health-related quality of life for pharmaceutical CUAs to determine the proportion of times that specific cost-effectiveness thresholds (i.e., \$20,000, \$50,000 and \$100,000 per QALY gained) were crossed. In 31% of sensitivity analyses, the ICER exceeded a threshold and in 13% of analyses, the ratio fell below a threshold. This indicates that if the magnitude of preference method-related variation reached the levels modeled in sensitivity analyses, there would be qualitative differences in results, as defined by crossing thresholds. The use of thresholds for policy decisions among a broad range of organizations is not well documented. We found some indication that thresholds of £20-30,000 are loosely used, and that overall, as ICERs increase, the likelihood of acceptance probably decreases.[102-105]

Discussion

The paucity of published evidence limits our ability to draw conclusions concerning how preference measurement method affects policy decisions. The ICERs available for comparison do not represent large enough changes to impact decision making. However, evidence from comparisons of preference-based systems support a wide range of variation in estimates. Based on the evidence reviewed here, it appears that choice of preference measure may contribute to qualitatively different ICERs under some circumstances. As ICERs rise, the probability of acceptance appears to decrease, making the differences in QALYs obtained using alternative methods potentially meaningful. This is especially important for treatments with long-term consequences and ICERs around common thresholds. In our review, EQ-5D tended to provide more favorable cost-effectiveness ratios than HUI, while SF-6D provided less favorable ratios than the other systems. Whether these patterns will hold for all applications depends on each system's ability to measure change across the full range of health.

To assess the impact of health state valuation on cost-utility analysis, our review focused on societal HSVs estimated with the most commonly used preference-based health state classification systems. However, evidence shows that in practice, the majority of CUAs do not yet meet Reference Case criteria[20,106,107]. Reviews indicate that use of preference-based systems appears to be increasing, with approximately 23% of published CUAs using community preferences, and 20% estimated from preference-based systems.[20,107] Direct valuations using standard gamble, time tradeoff, and rating scale were used 35.3% of the time, with 26.8% of valuations from patients.

Arguments for using community preferences hold that the population potentially affected by the decision should be polled. From a position of uncertainty about their own future health (i.e., 'the veil of ignorance') they would value policy decisions with the most benefit for society as a whole.[15,108] However, patients or population subgroups may be in a better position to assess the impact of condition-related health changes than a general population sample.[15,109] This is controversial in light of evidence that patients provide higher values than other groups for their health states,[15,33,67,68,75,108,110-112] and the possibility that this could cause undervaluation of preventive intervention[67] and treatment. In addition, there is evidence that differences in valuations between patients and other groups varies with other factors,[113,114] such as severity and chronicity of illness.[115,116]

As a practical matter, the choice of methods may be dictated by resource availability, making off-the-shelf tools attractive. To facilitate cost-utility analyses when primary data are not available, the Panel on Cost-Effectiveness in Health and Medicine recommended construction of a national catalogue of “off-the-shelf” community preferences for health states.[15] Work is ongoing in this area and includes population-based HSVs using multiple instruments,[40,117-120] mapping of health status data to HSVs, and catalogues or registries from the published literature.[20,121] Several methods have been explored to estimate HSVs from the SF-36,[66,122-125] the SF-12,[122,126-135] condition-specific tools,[136-138] and national survey instruments.[139-141] Some studies indicate that estimates from mapping may have important limitations[92,122,142] while others emphasize the fairly strong correlations between approaches.[23] A study investigating potential decision making impact of algorithm choice compared ICERs from various SF-36 and SF-12 algorithms and the SF-6D.[143] They reported ICERs for an asthma cohort ranging from \$30, 769 to \$63, 492 per QALY, and for a stroke cohort of \$27, 972 to \$50,000 per QALY. Many of the reported ICERs did not have overlapping confidence intervals. The growth of methods for estimating preferences using existing data is expanding the range of possible CUA applications, and warrants continued comparative study. In addition, efforts have been made to make published cost-utility analysis more accessible through registries,[144-147]some of which include information on methodology. These efforts are important steps in improving availability and transparency of cost-utility analyses.

In conclusion, it is unknown to what extent choice of preference measurement method impacts health policy decision-making. The existing evidence points toward potential impact for a subset of situations, most likely for ICERs near established thresholds when chronic diseases and/ or long-term health effects are involved. At present, there is no clearly superior method for estimating societal preferences, however, alternative systems produce a wide range in HSVs. The ability to convert health state values from one system to another would greatly improve comparability of CUAs. A recent publication of ongoing research using MEPS and National Health Interview Survey data uses linear regression to standardize estimates from various systems, and to convert estimates from one system to those of another.[148] Related research provided population-based norms.[149] Ongoing work will further our understanding of system responsiveness.[23]

Psychometric comparisons between systems were more common than comparisons of impact on ICERs and/or policy decisions. Explicit study of cost-effectiveness ratios obtained from alternative preference-based systems is needed to improve our understanding of potential policy implications. We encourage researchers who have access to comparative data to publish these estimates. Sensitivity analyses are suggested to address uncertainty associated with estimates of health effect, including variation due to measurement method. An in-depth systematic review of decisions made by policy makers relative to alternative measurement method would be a timely and important contribution to the literature. Research in these areas combined with ongoing dialogue about standardization and conversion of estimates hold promise for attaining consistency of estimates for CUA.

Acknowledgments

This study was funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (P60-AR048094) and the National Institute on Aging (R01-AG12262).

Appendix

The following key words were used in Medline to identify papers addressing the topic of preference measurement method and policy decision-making. In addition, selected author, references from papers known from previous work, and known web sites were hand-searched. Searches were limited to English language.

Cost-utility analysis
 Cost-effectiveness analysis
 Economic evaluation
 Methods
 Attitude to Health
 Health Status Indicators
 Health Status
 Quality of Life
 Quality-Adjusted Life Years
 QALYs
 Time tradeoff
 Standard gamble.
 Rating scale
 Visual analogue scale
 Utilities
 Health state values
 Values
 Preferences
 Preferences for health states
 Preference classification systems
 Preference-based
 EQ-5D.
 HUI
 SF-6D
 SF
 Quality of Well-Being.

References

1. Luce BR. What will it take to make cost-effectiveness analysis acceptable in the United States? *Medical Care* 2005;43(7 Supplement):II-44–II48.

2. Neumann PJ. Why don't Americans use cost-effectiveness analysis? *American Journal of Managed Care* 2004;10:308–312. [PubMed: 15152700]
3. Neumann PJ, Greenberg D, Olchanski NV, Stone PW, Rosen AB. Growth and quality of the cost-utility literature, 1976–2001. *Value in Health* 2005;8(1):3–9. [PubMed: 15841889]
4. Neumann PJ, Divi N, Beinfeld MT, Levine BS, Keenan PS, Halpern EF, et al. Medicare's national coverage decisions, 1999–2003: Quality of evidence and review times. *Health Affairs* 2005;24(1):243–254. [PubMed: 15647237]
5. Thorpe KE. The rise in health care spending and what to do about it. *Health Affairs* 2005;24(6):1436–1445. [PubMed: 16284014]
6. Foote SB, Neumann PJ. The impact of Medicare modernization on coverage policy: Recommendations for reform. *The American Journal of Managed Care* 2005;11(3):140–142. [PubMed: 15786852]
7. Bloom BS. Use of formal benefit/cost evaluations in health system decision making. *American Journal of Managed Care* 2004;10(5):329–35. see comment. [PubMed: 15152703]
8. Dickson M, Hurst J, Jacobzone S. Survey of pharmacoeconomic assessment activity in eleven countries. OECD. 2003
9. Siegel J. Cost-Effectiveness Analysis in US Healthcare Decision-Making: Where Is It Going? *Medical Care* 2005;43((7)(Supplement)):II-1–II-4.
10. Tunis SR. Economic analysis in healthcare decisions. *American Journal of Managed Care* 2004;10(5):301–304. [PubMed: 15152698]
11. Aspinall SL, Good C, Glassman PA, Valentino MA. The evolving use of cost-effectiveness analysis in formulary management within the Department of Veteran Affairs. *Medical Care* 2005;43(7 Supplement):II-20–II-26.
12. Neumann PJ. Evidence-based and value-based formulary guidelines. *Health Affairs* 2004;23(1):124–34. [PubMed: 15002635]
13. Guyatt GH, Baumann M, Pauker S, Halperin J, Maurer J, Owens D, et al. Addressing resource allocation issues in recommendations from clinical practice guideline panels. Suggestions from an American College of Chest Physicians Task Force. *Chest* 2006;129:182–187. [PubMed: 16424430]
14. Drummond M, Sculpher M. Common Methodological Flaws in Economic Evaluations. *Medical Care* 2005;43((7)(Supplement)):II-5–II-14.
15. Gold, MR.; Siegel, JE.; Russell, LB.; Weinstein, MC. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996.
16. Chapman RH, Stone PW, Sandberg EA, Bell C, Neumann PJ. A comprehensive league table of cost-utility ratios and a sub-table of “panel-worthy” studies. *Medical Decision Making* 2000;20:451–467. [PubMed: 11059478]
17. Chapman RC, Berger MLMD, Weinstein MC, Weeks J, Goldie SJ, Neumann PJ. When does quality-adjusting life-years matter in cost-effectiveness analysis? *Health Economics* 2004;13:429–436. [PubMed: 15127423]
18. Schackman BR, Gold HT, Stone PW, Neumann PJ. How often do sensitivity analyses for economic parameters change cost-utility analysis conclusions? *Pharmacoeconomics* 2004;22(5):293–300. [PubMed: 15061679]
19. Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health Economics* 2004;13:1203–1210. [PubMed: 15386669]
20. Bell CM, Chapman RH, Stone DA, Sandberg EA, Neumann PJ. An off-the-shelf help list: A comprehensive catalog of preference scores from published cost-utility analyses. *Medical Decision Making* 2001;21:288–294. [PubMed: 11475385]
21. Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients' well-being and consequent effect on a cost-effectiveness analysis. *Journal of Clinical Epidemiology* 1992;45(5):505–12. [PubMed: 1588356]
22. Nord E. Toward quality assurance in QALY calculations. *International Journal for Quality in Health Care* 1993;9(1):37–45.

23. Committee to Evaluate Measures of Health Benefits for Environmental H, and Safety Regulation. Valuing health for regulatory cost-effectiveness analysis: National Academy of Sciences. 2006
24. Hjelmgren J, Berggren F, Andersson F. Health economic guidelines-similarities, differences, and some implications. *Value in Health* 2001;4(3):225–250. [PubMed: 11705185]
25. National Institute of Clinical Excellence. Technology Appraisal Methods N0515; p. 20–25. <http://www.nice.org.uk/page.aspx?o=201973>
26. Greenberg D, Pliskin JS. Preference-based outcome measures in cost-utility analyses. A 20-year overview. *International Journal of Technology Assessment in Health Care* 2002;18(3):461–6. [PubMed: 12391940]
27. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. *Annual Review of Public Health* 2000;21:587–611.
28. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technology Assessment* 1999;3(9):57–81.
29. Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. *Journal of Clinical Epidemiology* 2003;56(4):317–325. [PubMed: 12767408]
30. Drummond, MF.; O'Brien, B.; Stoddart, GL.; Torrance, GW. *Methods for the Economic Evaluation of Health Care Programmes*. Second. New York: Oxford University Press; 1997.
31. Green C, Brazier J, Deverill M. Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics* 2000;17(2):151–65. [PubMed: 10947338]
32. Torrance GW. Measurement of health state utilities for economic appraisal. *Journal of Health Economics* 1986;5(1):1–30. [PubMed: 10311607]
33. Dolan P. Valuing health-related quality of life. Issues and controversies. *Pharmacoeconomics* 1998;15(2):119–127. [PubMed: 10351186]
34. Nord E. Methods for quality adjustment of life years. *Social Science & Medicine* 1992;34(5):559–69. [PubMed: 1604362]
35. Froberg DG, Kane RL. Methodology for measuring health-state preferences-II: Scaling methods. *Journal of Clinical Epidemiology* 1989;42(5):459–471. [PubMed: 2732774]
36. Torrance GW, Furlong W, Feeny D. Health utility estimation. *Expert Review of Pharmacoeconomics and Outcomes Research* 2002;2(2):99–108. [PubMed: 19807322]
37. van Osch SMC, Wakker PP, van den Hout WB, Stiggelbout AM. Correcting biases in Standard Gamble and Time Tradeoff Utilities. *Medical Decision Making* 2004;24:511–517. [PubMed: 15359000]
38. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Economics* 2002;11:447–456. [PubMed: 12112493]
39. Tsuchiya A, Brazier J, Roberts J. Comparison of valuation methods used to generate the EQ-5D and the SF-6D value sets. *Journal of Health Economics* 2006;25:334–346. [PubMed: 16271783]
40. Fryback DG, Dasbach EJ, Klein R, Klein BEK, Dorn N, Peterson K, et al. The Beaver Dam Health Outcomes Study: Initial Catalogue of Health-state Quality Factors. *Medical Decision Making* 1993;13(2):89–102. [PubMed: 8483408]
41. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37(1):53–72. [PubMed: 10158943]
42. Dolan P. Modelling valuations for EuroQol Health States. *Medical Care* 1997;35(11):1095–1108. [PubMed: 9366889]
43. Dolan P, R J. Modelling Valuations for EQ-5D Health States. An Alternative Model Using Differences in Valuations. *Medical Care* 2002;40(5):442–446. [PubMed: 11961478]
44. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics* 1995;7(6):490–502. [PubMed: 10155335]
45. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions. Health Utilities Index. *Pharmacoeconomics* 1995;7(6):503–20. [PubMed: 10155336]
46. Kaplan RM, Anderson JP. A general health policy model: update and applications. *Health Services Research* 1988;23(2):203–35. [PubMed: 3384669]

47. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 2002;21(2):271–92. [PubMed: 11939242]
48. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *Journal of Clinical Epidemiology* 1998;51(11):1115–28. [PubMed: 9817129]
49. Asada Y. Medical technologies, nonhuman aids, human assistance, and environmental factors in the assessment of health states. *Quality of Life Research* 2005;14:867–874. [PubMed: 16025586]
50. Ware JE Jr, Brook RH, Davies AR, Lohr KN. Choosing measures of health status for individuals in general populations. *American Journal of Public Health* 1981;71(6):620–5. [PubMed: 7235100]
51. EuroQol Group. EQ-5D. Jan 31. 2006 <http://www.euroqol.org/web/users/valuation.php>
52. Kaplan RM, Ganiats TG, Sieber WJ, Anderson JP. The Quality of Well-Being Scale: critical similarities and differences with SF-36. *International Journal for Quality in Health Care* 1998;10(6):509–20. see comment. [PubMed: 9928590]
53. Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis? *Health Economics* 2006;15:653–664. [PubMed: 16498700]
54. McCabe CJ, Stevens KJ, Brazier J. Utility Scores for the Health Utilities Index Mark 2. An empirical assessment of alternative mapping functions. *Medical Care* 2005;43(6):627–635. [PubMed: 15908858]
55. Stevens KJ, McCabe CJ, Brazier J. Mapping between Visual Analogue Scale and Standard Gamble data; results from the UK Health Utilities Index 2 valuation survey. *Health Economics*. 2005 in press.
56. Conner-Spady, B.; Voaklander, DC.; Suarez-Almazor, ME. The effect of different EuroQol weights on potential QALYs gained in patients with hip and knee replacement. 17th Plenary Meeting of the EuroQol Group; 2000; p. 127-137.
57. Burns AW, Bourne RB, Chesworth BM, MacDonald SJ, Rorabeck CH. Cost effectiveness of revision total knee arthroplasty. *Clinical Orthopaedics & Related Research* 2006;446:29–33. [PubMed: 16672868]
58. Fryback DG. A US valuation of the EQ-5D. *Medical Care* 2005;43(3):199–200. [PubMed: 15725975]
59. Johnson JA, Ohinmaa A, Murti B, Sintonen H, Coons SJ. Comparison of Finnish and U.S.-based visual analog scale valuations of the EQ-5D measure. *Medical Decision Making* 2000;20(3):281–9. [PubMed: 10929850]
60. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? *Medical Care* 2005;43(3):221–8. [PubMed: 15725978]
61. Badia X, Montserrat R, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Medical Decision Making* 2001;21:7–16. [PubMed: 11206949]
62. Keeney R. A group preference axiomatization with cardinal utility. *Management Science* 1976;23:140–145.
63. Torrance GW. Preferences for health outcomes and cost-utility analysis. *The American Journal of Managed Care* 1997;3(supplement):S8–S20. [PubMed: 10180342]
64. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Medical Care* 2005;43(3):203–20. [PubMed: 15725977]
65. Brazier, J.; Roberts, J.; Tsuchiya, A. A Comparison of the EQ-5D and SF-6D Across Seven Patient Groups. *Proceedings of the 18th Plenary Meeting of the Euroqol Group*; 2001; p. 9-31.
66. Hollingworth W, Deyo RA, Sullivan SD, Emerson SS, Gray DT, Jarvik JG. The practicality and validity of directly elicited and SF-36 derived health state preferences in patients with low back pain. *Health Economics* 2002;11(1):71–85. [PubMed: 11788983]
67. Gabriel SE, Kneeland TS, Melton LJ 3rd, Moncur MM, Ettinger B, Tosteson AN. Health-related quality of life in economic evaluations for osteoporosis: whose values should we use? *Medical Decision Making* 1999;19(2):141–8. [PubMed: 10231076]

68. Krahn M, Ritvo P, Irvine J, Tomlinson G, Bremner KE, Bezjak A, et al. Patient and community preferences for outcomes in prostate cancer: implications for clinical policy. *Medical Care* 2003;41(1):153–64. [PubMed: 12544552]
69. Macran S, Weatherly H, Kind P. Measuring population health: a comparison of three generic health status measures. *Medical Care* 2003;41(2):218–31. [PubMed: 12555050]
70. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D on hypothetical preferences in a general population. *Health Economics* 2005;14:1169–1189. [PubMed: 15942981]
71. Fisk JD, Brown MG, Sketris IS, Metz LM, Murray TJ, Stadnik KJ. A comparison of health utility measures for the evaluation of multiple sclerosis treatments. *Journal of Neurological and Neurosurgical Psychiatry* 2005;76:58–63.
72. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures. *International Journal of Audiology* 2005;44:157–163. [PubMed: 15916116]
73. Espallargues M, Czoski-Murray CJ, Bansback NJ, Carlton J, Lewis GM, Hughes LA, et al. The impact of age-related macular degeneration on health status utility values. *Investigative Ophthalmology & Visual Science* 2005;46(11):4016–23. [PubMed: 16249475]
74. McDonough CM, Grove MR, Tosteson TD, Lurie JD, Hilibrand AS, Tosteson AN. Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among spine patient outcomes research trial (SPORT) participants. *Quality of Life Research* 2005;14(5):1321–32. [PubMed: 16047507]
75. Feeny D, Furlong W, Saigal S, Sun J. Comparing directly measured standard gamble scores to HUI2 and HUI3 utility scores: group- and individual-level comparisons. *Social Science & Medicine* 2004;58:799–809. [PubMed: 14672594]
76. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Annals of Medicine* 2001;33(5):358–70. [PubMed: 11491195]
77. Elvik R. The validity of using health state indexes in measuring the consequences of traffic injury for public health. *Social Science & Medicine* 1995;40(10):1385–98. [PubMed: 7638647]
78. de Vries SO, Kuipers WD, Hunink MG. Intermittent claudication: symptom severity versus health values. *Journal of Vascular Surgery* 1998;27(3):422–30. [PubMed: 9546227]
79. Marra CA, Esdaile JM, Guh D, Kopec JA, Brazier JE, Koehler BE, et al. A comparison of four indirect methods of assessing utility values in rheumatoid arthritis. *Medical Care* 2004;42(11):1125–31. [PubMed: 15586840]
80. Schulz MW, Chen J, Woo HH, Keech M, Watson ME, Davey PJ. A comparison of techniques for eliciting patient preferences in patients with benign prostatic hyperplasia. *Journal of Urology* 2002;168(1):155–9. [PubMed: 12050512]
81. Stavem K, Bjornaes H, Lossius MI. Properties of the 15D and EQ-5D utility measures in a community sample of people with epilepsy. *Epilepsy Research* 2001;44(2-3):179–89. [PubMed: 11325573]
82. Luo N, Chew L, Fong K, et al. A comparison of the EuroQol-5D and the Health Utilities Index Mark 3 in patients with rheumatic disease. *The Journal of Rheumatology* 2003;30(10):2268–2274. [PubMed: 14528528]
83. Maddigan SL, Feeny D, Johnson JA. A comparison of the Health Utilities Indices Mark 2 and Mark 3 in type 2 diabetes. *Med Decis Making* 2003;23:489–501. [PubMed: 14672109]
84. Pickard AS, Johnson JA, Feeny DH. Responsiveness of generic health-related quality of life measures in stroke. *Quality of Life Research* 2005;14(1):207–19. [PubMed: 15789955]
85. Hatoum HT, Brazier JE, Akhras KS. Comparison of the HUI3 with the SF-36 preference based SF-6D in a clinical trial setting. *Value in Health* 2004;7(5):602–9. [PubMed: 15367255]
86. Feeny D, Wu L, Eng K. Comparing Short Form 6D, Standard Gamble, and Health Utilities Index Mark 2 and Mark 3 utility scores: Results from total hip arthroplasty patients. *Quality of Life Research* 2004;13:1659–1670. [PubMed: 15651537]

87. Conner-Spady, B.; Suarez-Almazor, ME. A Comparison of preference-based health status tools in patients with musculoskeletal disease. 18th Plenary Meeting of the EuroQol Group; 2001; p. 235-245.
88. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. *Medical Care* 2003;41(7):791–801. [PubMed: 12835603]
89. Bosch JL, Hunink M. Comparison of the Health Utilities Index Mark 3 (HUI3) and the EuroQol EQ-5D in patients treated for intermittent claudication. *Quality of Life Research* 2000;9:591–601. [PubMed: 11236850]
90. Suarez-Almazor M, K C, Johnson J, Skeith K, V D. Use of health status measures in patients with low back pain in clinical settings. Comparison of specific, generic and preference-based instruments. *Rheumatology* 2000;39:783–790. [PubMed: 10908699]
91. Bosch JL, Halpern EF, Gazelle GS. Comparison of preference-based utilities of the Short-Form 36 Health Survey and Health Utilities Index before and after treatment of patients with intermittent claudication. *Medical Decision Making* 2002;22(5):403–9. [PubMed: 12365482]
92. Kaplan R, Groessl EJ, Sengupta N, Sieber WJ, Ganiats TG. Comparison of measured utility scores and imputed scores from the SF-36 in patients with rheumatoid arthritis. *Medical Care* 2005;43(1): 79–87. [PubMed: 15626937]
93. Longworth L, Bryan S. An empirical comparison of EQ-5D and SF-6D in liver transplant patients. *Health Economics* 2003;12(12):1061–7. [PubMed: 14673814]
94. Neumann PJ, Sandberg EA, Araki SS, Kuntz KM, Feeny D, Weinstein MC. A comparison of HUI2 and HUI3 utility scores in Alzheimer's disease. *Medical Decision Making* 2000;20(4):413–22. [PubMed: 11059474]
95. 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 AQL. *Health Economics* 2004;13(8):793–805. [PubMed: 15322991]
96. Stavem K, Froland SS, Hellum KB. Comparison of preference-based utilities of the 15D, EQ-5D and SF-6D in patients with HIV/AIDS. *Quality of Life Research* 2005;14(4):971–80. [PubMed: 16041894]
97. Thoma A, Sprague S, Veltri K, Duku E, Furlong W. Methodology and measurement properties of health-related quality of life instruments: A prospective study of patients undergoing breast reduction surgery. *Health and Quality of Life Outcomes* 2005;3(44)
98. Lanfitt J, Vickrey B, McDermott M, Messing S, Berg A, SPencer S, et al. Validity and responsiveness of generic preference-based HRQOL instruments in chronic epilepsy. *Quality of Life Research* 2006;15:899–914. [PubMed: 16721649]
99. Thomas, KJ.; MacPherson, H.; Ratcliffe, J.; Thorpe, L.; Brazier, J.; Campbell, M., et al. *Health Technology Assessment*. Vol. 9. Winchester, England: 2005. Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain.
100. Stiggelbout AM. Health state classification systems: How comparable are our ratios? *Med Decis Making* 2006;25:223–225. [PubMed: 16751320]
101. Neumann PJ. Health utilities in Alzheimer's disease and implications for cost-effectiveness analysis. *Pharmacoeconomics* 2005;23(6):537–41. [PubMed: 15960551]
102. National Institute of Clinical Excellence. *Technology Appraisal Methods* N0515; p. 33-34. <http://www.nice.org.uk/page.aspx?o=201973>
103. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics* 2004;13:437–452. [PubMed: 15127424]
104. Pearson SD, Rawlins MD. Quality, innovation, and value for money. NICE and the British. *National Health Service JAMA* 2005;294(20):2618–2622.
105. Henry DA, Hill SR, Harris A. Drug prices and value for money. The Australian pharmaceutical benefits scheme. *JAMA* 2005;294(20):2630–2632. [PubMed: 16304078]
106. Rasanen P, Roine E, Sintonen H, Semberg-Kontinen V, Ryyanen OP, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. *International Journal of Technology Assessment in Health Care* 2006;22(2):235–41. [PubMed: 16571199]

107. Brauer C, Rosen AB, Greenberg D, Neumann PJ. Trends in the measurement of health utilities in published cost-utility analyses. *Value in Health* 2006;9(4):213–218. [PubMed: 16903990]
108. De Wit GA, Busschbach JJV, De Charro FT. Sensitivity and Perspective in the valuation of health status: whose values count? *Health Economics* 2000;9:109–126. [PubMed: 10721013]
109. Greenberg D, Pliskin JS. Using health state classification systems for utility elicitation in the elderly. *Med Decis Making* 2006;25:220–222. [PubMed: 16751319]
110. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* 1999;8(1):25–39. [see comment] [erratum appears in *Health Econ* 1999 Sep;8(6):559]. [PubMed: 10082141]
111. Polsky D, Willke RJ, Scott K, Schulman KA, Glick HA. A comparison of scoring weights for the Euroqol derived from patients and the general public. *Health Economics* 2001;10:27–37. [PubMed: 11180567]
112. Feeny D, Blanchard C, Mahon JL, Bourne R, Rorabeck C, Stitt L, et al. Comparing community-preference-based and direct standard gamble utility scores: evidence from elective total hip arthroplasty. *International Journal of Technology Assessment in Health Care* 2003;19(2):362–72. [PubMed: 12862193]
113. Sullivan PW, Gushchyan V. Systematic differences in subjective vs. societal preferences. *Quality of Life Research* 2005;14(9):2011.
114. Soucek J, Byrne MM, Kelly PA, O'Malley KO, Richardson M, Pak C, et al. Valuation of arthritis health states across ethnic groups and between patients and community members. *Medical Care* 2005;43(9):921–928. [PubMed: 16116357]
115. Insinga RP, Fryback DG. Understanding differences between self-ratings and population ratings for health in the EuroQOL. *Quality of Life Research* 2003;12(6):611–619. [PubMed: 14516171]
116. McPherson K, Myers J, Taylor WJ, McNaughton HK, Weatherall M. Self-valuation and societal valuations of health state differ with disease severity in chronic and disabling conditions. *Medical Care* 2004;42(11):1143–1151. [PubMed: 15586842]
117. Sullivan PW, Lawrence WF, Gushchyan V. A national catalogue of Preference-based scores for chronic conditions in the United States. *Medical Care* 2005;43(7):736–749. [PubMed: 15970790]
118. Sullivan PW, Gushchyan V. Preference-based EQ-5D Index scores for chronic conditions in the United States. *Med Decis Making* 2006 July-August;26
119. Luo N, Johnson JA, Shaw JW, Feeny D, Coons SJ. Self-reported health status of the general adult U.S population as assessed by the EQ-5D and Health Utilities Index. *Medical Care* 2005;43(11):1078–1086. [PubMed: 16224300]
120. Belanger, A.; Berthelot, JM.; Guimond, E.; Houle, C. A Head-to-Head Comparison of Two Generic Health Status Measures in the Household Population: McMaster Health Utilities Index (Mark 3) and the EQ-5D. 2000 Final Revision April. Ottawa: Statistics Canada, Health Analysis and Modelling Group; 2000.
121. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Medical Care* 2000;38(6):583–637. [PubMed: 10843310]
122. Sherbourne C, Unutzer J, Schoenbaum M, Duan N, Lenert L, Sturm R, et al. Can utility-weighted health-related quality-of-life estimates capture health effects of quality improvement for depression? *Medical Care* 2001;39(11):1246–1259. [PubMed: 11606878]
123. Lee TA, Hollingworth W, Sullivan SD. Comparison of directly elicited preferences to preferences derived from the SF-36 in adults with asthma. *Medical Decision Making* 2003;23:323–334. [PubMed: 12926582]
124. Fryback DG, Lawrence WF, Martin PA, Klein R, Klein BE. Predicting Quality of Well-being scores from the SF-36: results from the Beaver Dam Health Outcomes Study. *Medical Decision Making* 1997;17(1):1–9. [PubMed: 8994146]
125. Shmueli A. Subjective health status and health values in the general population. *Medical Decision Making* 1999;19(2):122–7. [PubMed: 10231074]
126. Lundberg L, Johannesson M, Isacson DG, Borgquist L. The relationship between health-state utilities and the SF-12 in a general population. *Medical Decision Making* 1999;19(2):128–40. [PubMed: 10231075]

127. Franks P, Lubetkin EI, Gold MR, Tancredi DJ. Mapping the SF-12 to preference-based instruments. *Medical Care* 2003;41(11):1277–1283. [PubMed: 14583690]
128. Franks P, Lubetkin EI, Gold MR, Tancredi DJ, Jia H. Mapping the SF-12 to the Euroqol EQ-5D Index in a national US sample. *Medical Decision Making* 2004;24:247–254. [PubMed: 15185716]
129. Sengupta N, Nichol MB, Wu J, Globe D. Mapping the SF-12 to the HUI3 and VAS in a managed care population. *Medical Care* 2004;42(9):927–937. [PubMed: 15319619]
130. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Medical Care* 2004;42(9):851–859. [PubMed: 15319610]
131. Sanderson K, Andrews G, Corry J, Lapsley H. Using the effect size to model change in preference values from descriptive health status. *Quality of Life Research* 2004;13(7):1255–64. [PubMed: 15473504]
132. Lenert LA, Sherbourne CD, Sugar C, Wells KB. Estimation of utilities for the effects of depression from the SF-12. *Medical Care* 2000;38(7):763–70. [PubMed: 10901359]
133. Lawrence WF, Fleishman JA. Predicting EuroQoL EQ-5D preference scores from the SF-12 Health Survey in a nationally representative sample. *Medical Decision Making* 2004;24(2):160–9. [PubMed: 15090102]
134. Gray AM, Rivero-Arias O, Clarke PM. Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping. *Med Decis Making* 2006;26(1):18–29. [PubMed: 16495197]
135. Sullivan PW, Ghushchyan V. Mapping the EQ-5D Index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making* 2006;26:401–409. [PubMed: 16855128]
136. Lenert LA, Sturley AP, Rapaport MH, Chavez S, Mohr PE, Rupnow M. Public preferences for health states with schizophrenia and a mapping function to estimate utilities from positive and negative symptom scale scores. *Schizophrenia Research* 2004;71(1):155–65. [PubMed: 15374583]
137. Kind P, Macran S. Eliciting social preference weights for functional assessment of cancer therapy-lung health states. *Pharmacoeconomics* 2005;23(11):1143–1153. [PubMed: 16277549]
138. Brazier J, Kolotkin RL, Crosby RD, Williams GR. Estimating a preference-based single index for the Impact of Weight on Quality of Life-Lite (IWQOL-Lite) instrument from the SF-6D. *Value in Health* 2004;7(4):490–498. [PubMed: 15449641]
139. Gold MR, Franks P, McCoy KI, Fryback DG. Toward consistency in cost-utility analyses: using national measures to create condition-specific values. *Medical Care* 1998;36(6):778–92. see comment. [PubMed: 9630120]
140. Gold M, Franks P, Erickson P. Assessing the health of the nation. The predictive validity of a preference-based measure and self-rated health. *Medical Care* 1996;34(2):163–77. [PubMed: 8632690]
141. Erickson P. Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Quality of Life Research* 1998;7:101–114. [PubMed: 9523491]
142. Feeny D. Aas godd as it gets but good enough for which applications? *Med Decis Making* 2006;26:307–309. [PubMed: 16855119]
143. Pickard AS, Wang Z, Walton SM, Lee TA. Are decisions using cost-utility analyses robust to choice of SF-36/SF-12 preference-based algorithm? *Health & Quality of Life Outcomes* 2005;3(1):11. [PubMed: 15748287]
144. Harvard Center for Risk Analysis, Harvard School of Public Health. Cost-Effectiveness Analysis Registry. Feb 2. 2006 <http://www.hsph.harvard.edu/cearegistry/>
145. The Office of Health Economics, The International Federation of Pharmaceutical Manufacturers' Associations. Health Economic Evaluation Database (HEED). Feb 2. 2006 <http://www.ohe-heed.com/>
146. The United Kingdom National Health Service. Economic Evaluation Database. Feb 2. 2006 <http://www.york.ac.uk/inst/crd/nhsdhp.htm>
147. The United Kingdom National Health Service. European Network on Health Economics Evaluation Databases. Feb 2. 2006

http://www.cordis.lu/data/PROJ_FP5/ACTIONeqDndSESSIONeq112482005919ndDOCe622ndTBLeqEN_PROJ.htm

148. Franks P, Hanmer J, Fryback DG. Relative disutilities of 47 risk factors and conditions assessed with seven preference-based health status measures in a national U.S sample. Toward consistency in cost-effectiveness analysis. *Medical Care* 2006;44(5):478–485. [PubMed: 16641667]
149. Hanmer J, Lawrence WF, Anderson JP, Kaplan R, Fryback DG. Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related Quality-of-life scores. *Med Decis Making* 2006;26:391–400. [PubMed: 16855127]

Table 1
Comparison of mean health state values in cross-sectional comparisons of systems

Study (first author, year, n)	Systems Studied	Baseline Mean (SD)	Correlation Coefficient				
			EQ5D	HUI2	HUI3	SF6D	15D ^e
Luo, 2005, <i>n</i> = 4048 General adult US population	EQ-5D	0.87(0.13)		0.67 ^d	0.66 ^d		
	HUI2	0.86(0.32)			0.87 ^a		
	HUI3	0.81(0.38)					
Petrou, 2005, <i>n</i> = 14,736 General adult UK population	EQ-5D	0.84(0.23)				0.70 ^a	
	SF-6D	0.80(0.15)					
	EQ-5D	0.79(0.23)			0.44 ^b	0.50 ^b	
Barton, 2005, <i>n</i> = 915 Hearing impairment	HUI3	0.56(0.15)				0.41 ^b	
	SF-6D	0.77(0.08)					
McDonough, 2005, <i>n</i> = 2,097 Spine disorders	EQ-5D	0.39(0.33)		0.68 ^c	0.67 ^c	0.69 ^c	
	HUI2	0.59(0.22)			0.78 ^c	0.67 ^c	
	HUI3	0.45(0.27)				0.72 ^c	
Espallargues, 2005, <i>n</i> = 209 Macular degeneration	SF-6D	0.57(0.12)					
	EQ-5D	0.72(0.22)	-	-	-	-	-
	HUI3	0.34(0.28)					
Marra, 2004, <i>n</i> = 313 Rheumatoid arthritis	SF-6D	0.66(0.14)					
	EQ-5D	0.66(0.13)	-	-	-	-	-
	HUI2	0.71(0.19)					
Feeny, 2004, <i>n</i> = 264 Teen survivors low birthweight	HUI3	0.53(0.29)					
	SF-6D	0.63(0.24)					
	HUI2	0.92(0.13)	-	-	-	-	-
Maddigan, 2003, <i>n</i> = 372 Type II diabetes	HUI3	0.84(0.16)					
	HUI2	0.78(0.17)	-	-	-	-	-
	HUI3	0.64(0.29)					
Luo, 2003, <i>n</i> = 114 Rheumatic disease	EQ-5D	0.75(0.21)			0.45 ^c		
	HUI3	0.76(0.17)					

Study (first author, year, n)	Systems Studied	Baseline Mean (SD)	Correlation Coefficient				
			EQ5D	HUI2	HUI3	SF6D	15D ^e
O'Brien, 2003, <i>n</i> = 246	HUI3	0.61(0.12)				0.58 ^a	
Cardiac disease	SF-6D	0.58(0.16)					
Schulz, 2002, <i>n</i> = 29	EQ-5D	0.81(0.12)			0.32 ^c		
Benign prostatic hyperplasia	HUI3	0.79(0.12)					
Brazier, 2001 ^d , <i>n</i> = 2605	EQ-5D	0.59(-)				0.66 ^a	
Combined patient groups	SF-6D	0.63(-)					
Stavem, 2001, <i>n</i> = 348	EQ-5D	0.81(0.23)					0.78 ^c
Epilepsy	15D ^e	0.88(0.12)					
Belanger, 2000, <i>n</i> = 1,477	EQ-5D	0.83(-)			0.69 ^a		
Canadian general population	HUI3	0.85(-)					

^a = Pearson's correlation coefficient

^b = Kendall's tau for *n* = 863

^c = Spearman's rank correlation coefficient

^d = This study reported values for patient groups: COPD, osteoarthritis, irritable bowel syndrome, low back pain, leg ulcers, menopausal symptoms, and age greater than 65 years.

^e = 15D is a preference-based health state classification system for which data were reported compared to more commonly used systems [81]

- = not reported

Table II
Studies addressing longitudinal comparisons in mean health state values provided by the most commonly used preference-based health state classification systems

Study (first author, year, n)	Systems Studied	Baseline Mean (SD)	Change in Health State Value Mean (SD)
Langfitt, 2006, <i>n</i> =64 Chronic Epilepsy	EQ-5D	0.76(0.26)	0.11(0.28)
	EQ-5D-US	0.82(0.18)	0.07(0.21)
	HUI2	0.78(0.18)	0.03(0.15)
	HUI3	0.61(0.30)	0.08(0.27)
	SF-6D	0.70(0.14)	0.08(0.16)
Kaplan, 2005, <i>n</i> = 628 Rheumatoid Arthritis	EQ-5D	0.56(0.25)	0.10(0.27)
	HUI2	0.64(0.20)	0.06(0.18)
	HUI3	0.43(0.27)	0.09(0.24)
	SF-6D(VAS)	0.43(0.16)	0.06(0.16)
Pickard, 2005, <i>n</i> = 98 Stroke	SF-6D(SG)	0.81(0.10)	0.03(0.09)
	EQ-5D	0.31(0.38)	0.32(0.38)
	HUI2	0.51(0.20)	0.12(0.23)
	HUI3	0.19(0.30)	0.25(0.32)
Stavem, 2005, <i>n</i> = 60 HIV/AIDS	SF-6D	0.55(0.09)	0.13(0.15)
	EQ-5D	0.77(0.26)	0.02(-)
	SF-6D	0.73(0.17)	0.03(-)
	15D ^a	0.86(0.14)	0.02(-)
Thoma, 2005, <i>n</i> = 41 Breast Reduction Surgery	HUI2	(-)	0.06(0.14)
	HUI3	(-)	0.12(0.19)
Hatoum, 2004, <i>n</i> = 184 Coronary Artery Disease	HUI3	0.63(0.29)	0.15(0.28)
	SF-6D	0.67(0.12)	0.08(0.13)
Feeny, 2004, <i>n</i> = 63 Total hip arthroplasty	HUI2	0.62(0.19)	0.22(0.07)
	HUI3	0.62(0.32)	0.23(0.04)
	SF-6D	0.61(0.10)	0.10(0.02)
Holland, 2004, <i>n</i> = 123 Elderly patients	EQ-5D	0.61(0.29)	-0.16(0.34)
	AQoL ^b	0.45(0.27)	-0.12(0.24)
Longworth, 2003, <i>n</i> =183 Liver transplant	EQ-5D	0.52(0.33)	0.09(0.37)
	SF-6D	0.61(0.12)	0.01(0.28)
Bosch, 2002, <i>n</i> = 87 Intermittent claudication	HUI2	0.70(0.20)	0.10(-)
	HUI3	0.66(0.20)	0.11(-)
	SF-6D	0.66(0.09)	0.08(-)
Bosch, 2000, <i>n</i> = 88 Intermittent claudication	EQ-5D	0.57(0.25)	0.22(-)
	HUI3	0.66(0.20)	0.11(-)
Suarez-Almazor, 2000, <i>n</i> = 37 Low back pain	EQ-5D	0.38(0.33)	-4.8 (17.4) ^c
	HUI2	0.49(0.19)	-1.8 (16.1) ^c

^a 15D is a preference-based health state classification system for which data were reported compared to most commonly used systems [81]

^b AQoL is a preference-based health state classification system for which data were compared to most commonly used systems [76,95]

^c = 3 month results, rescaled to 1-100, with 100 corresponding to best health

^r = Not reported