Title: PATIENT EXPERIENCE-BASED VALUE SETS: ARE THEY STABLE?

Authors:

A. Simon Pickard, PhD

Department of Pharmacy Systems, Outcomes and Policy, and Center for Pharmacoepidemiology Pharmacoeconomic Research, College of Pharmacy, University of Illinois at Chicago; Department of Medical Research, China Medical University Hospital, Taichung, Taiwan.

Address: 833 S. Wood Street, Rm 287 (M/C 871), Chicago, IL, USA

Telephone: 312-996-7876

Fax: 312-996-2954

E-mail: pickard1@uic.edu

Yu-Ting Hung, MS

Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago

Address: 833 S. Wood Street, Rm 287 (M/C 871), Chicago, IL, USA

Telephone: 312-996-7876

Fax: 312-996-2954

E-mail: yhung9@uic.edu

Fang-Ju Lin, PhD

Graduate Institute of Clinical Pharmacy, College of Medicine, National Taiwan University, Taipei, Taiwan; Department of Pharmacy, National Taiwan University Hospital, Taipei, Taiwan.

Address: No.33, Linsen S. Rd., Zhongzheng Dist., Taipei City 100, Taiwan (R.O.C.)

Telephone: +886-2-33668809

Fax: +886-2-33668780

E-mail: fjilin@ntu.edu.tw

Todd A. Lee, PhD

Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago

Address: 833 S. Wood Street, Rm 287 (M/C 871), Chicago, IL, USA

Telephone: 312-996-7876

Fax: 312-996-2954

E-mail: toddlee@uic.edu

Disclosure: There was no funding received for this work from National Institutes of Health (NIH), Wellcome Trust, Howard Hughes Medical Institute (HHMI), or any other organizations. There is no potential conflict of interests for all authors in this work.

Word count: 2,710 words

Number of text pages: 12 pages

Number of references: 45 references

Number of figures/tables: 3 tables (and 1 supplemental table)

PATIENT EXPERIENCE-BASED VALUE SETS: ARE THEY STABLE?

Abstract

Background. While societal preference weights are desirable to inform resource allocation decision-making, patient experienced health state-based value sets can be useful for clinical decision-making, but context may matter.

Objective. To estimate EQ-5D value sets using visual analogue scale (VAS) ratings for patients undergoing knee replacement surgery and compare the estimates before and after surgery.

Methods. We used the Patient Reported Outcome Measures (PROMs) data collected by the UK National Health Service on patients undergoing knee replacement from 2009 to 2012. Generalized least squares regression models were used to derive value sets based on the EQ-5D-3L using a development sample before and after surgery, and model performance was examined using a validation sample.

Results. A total of 90,450 pre- and post-operative valuations were included. The largest decrement in the preoperative value set was associated with the dimension of anxiety/depression, followed by self-care, mobility, usual activities, and pain/discomfort. However, pain/discomfort had a greater impact postoperatively. Compared with preoperative health problems, post-surgical health problems were associated with larger value decrements, with significant differences in several dimensions, including level 2 of mobility, level 2/3 of usual activities, level 3 of pain/discomfort, and level 3 of anxiety/depression. Similar results were observed across subgroups stratified by age and gender.

Conclusions. Findings suggest patient experience-based value sets are not stable (i.e., context such as timing matters). However, the knowledge that lower values are assigned to health states post-surgery compared to pre-surgery may be useful for the patient-doctor decision-making process.

Keywords. EQ-5D, experience-based value set, self-rated health, visual analogue scale, knee replacement

INTRODUCTION

Preferences for health as experienced by patients are recognized as distinct from preferences elicited from the general populations, which involve stated preferences for hypothetical health states.[1-4] Societal preferences derived from the general population are recommended as the primary source of values by health technology assessment (HTA) guidelines for the purpose of economic evaluations.[5,6] In contrast, value sets derived from individuals who actually "experienced" the health states provide valuable insights into self-perceived well-being that can be informative to clinical decision-making.[7-9] Although whose values should be used as a source

of information to inform decision making has been long debated,[10] relatively few investigations have been conducted on their strengths and limitations.

An experienced health state-based value set (EHS, in the terminology of Leidl et al.[9]) refers to a value set where the preference for current or experienced health is elicited, and thus only one observation per respondent can be obtained. This contrasts with the approach to estimating societal value sets where stated preferences for a set of hypothetical "given health states" (GHS) are elicited from each respondent of the general population. One commonly used generic measure of health conducive to estimating patient experience-based value sets is the EQ-5D, as it includes both a health state descriptive system and a self-rating of health using a visual analog scale (VAS). A 2013 review of EQ-5D valuation studies found that 12 of the 31 reviewed studies used VAS-based approaches to derive value sets and two of them used both time trade-off (TTO) and VAS techniques,[11] from which one experience-based value set for the German population was included.[9]

Since then, several studies have estimated value sets for EQ-5D health state descriptive system using the VAS as the dependent variable to illustrate the plausibility and usefulness of the approach.[9,12,13] When considering EHS for patients with inflammatory bowel disease as the reference gold standard, Leidl et al. found that EHS based on the German population had better predictive accuracy than the German and UK societal value sets based on stated preferences for given (hypothetical) health states (GHS).[9,14] Little et al. compared a European GHS value set to EHS value sets for patients with various conditions and demonstrated that systematic differences could dictate the conclusion of an economic evaluation.[15] Consistent with previous studies which showed that patient value sets tend to rate health states as higher (better) than societal preference-based sets,[10,16] a 2015 study found that a Swedish-based EHS value set based on a

large cohort of various patient groups tended to provide higher scores than a UK algorithm based on GHS, resulting in different conclusions and resource-allocation decisions depending on the value sets used.[17] In another Swedish study, preoperative and postoperative EHS value sets were developed for patients undergoing total hip replacement,[18] with a primary focus on contrasting the EHS based value sets with the GHS-based UK TTO value set. However, the pre and post-operative value sets, which did appear to be different, were not compared and contrasted to each other, nor were the implications of EHS-based value sets that changed with timing of the "experience" discussed.

The overall aim of this study was to estimate EQ-5D value sets using a visual analogue scale (EQ-VAS) in patients who experienced knee replacement surgery in order to test whether patient experienced health state value sets were stable by comparing them before and after surgery. Knee replacement surgery was a condition well-suited to the research question, as knee pain is a prevalent problem that affects approximately 25% of the population aged 55 years and over in England,[19] and knee replacement can potentially improve pain, mobility, and quality of life of the patients.

METHODS

Subjects

We used the Patient Reported Outcome Measures (PROMs) data collected by the UK National Health Service (NHS) on patients undergoing knee replacement between April 2009 and March 2012.[20] PROMs questionnaires were sent out by mail, and completion was voluntary. Basic patient characteristics such as gender and age group were collected, and outcome

measures, including EQ-5D and Oxford Knee Score (OKS), were self-completed a few weeks before and at least 6 months following knee replacement.

Measures

The EQ-5D is a generic, indirect preference-based measure of health that consists of a descriptive 5-dimensional health classifier and a visual analogue scale (EQ-VAS).[21] The EQ-5D descriptive system includes five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. There are three levels of response option under each dimension to indicate no problems (level 1), some/moderate problems (level 2), and extreme problems (level 3). The combination of the dimensions and levels defines a total of 243 possible health states, and each EQ-5D health state can be referred by the vector of levels across the 5 dimensions, e.g. 11223. Index-based summary scores can be obtained from the responses to the self-classifier by applying scoring algorithms that have been developed using VAS-based and choice-based techniques, i.e. the time trade-off.[11] The EQ-VAS is a direct valuation or self-rating of overall health anchored by 'best imaginable health state' at 100 and 'worst imaginable state' set at 0.[21]

Patients also completed the Oxford Knee Score (OKS), a 12-item disease-specific questionnaire designed to assess symptom severity and joint function in patients undergoing total knee replacement surgery,[22] from which a summary score can be derived ranging from 0 (worst functional outcome) to 48 (best).

Data analysis

All respondents who completed the EQ-5D classification and VAS both pre- and post-surgery were included in the analysis. In order to compare the predictive accuracy of the modeling approach, respondents were randomly selected into development (80%) and validation

(20%) samples. Regression was used to model EQ-VAS scores on the 10 main effect terms that captured the value decrement from level 1 to level 2 or 3 on each of the 5 EQ dimensions. Our intent was to examine the stability of the value weights in each level and dimension of the EQ-5D before and after knee replacement. Although several model specifications were examined, generalized least squares (GLS) was selected as the most appropriate model to estimate the pre and post value sets, which takes into account the correlated nature of the data and provided comparable model fit to other models we initially examined using the developmental sample.

To illustrate the difference in pre and post values that would be estimated by each value set for a given set of health states, the scoring algorithms were applied to the 10 most common health states observed pre- and post-operatively. Analyses were carried out in SAS Version 9.4 (SAS Institute Inc., Cary, NC, USA) and STATA 14 (StataCorp LP, College Station, TX, USA).

RESULTS

There were 115,206 patients available in the PROMS dataset who received NHS-funded knee replacement between April 2009 and March 2012. The preoperative PROMs questionnaire was completed by 83.6% of eligible patients, and the postoperative questionnaire was completed by 79.5%.[23] After excluding patients who failed to complete all components of the EQ-5D, a total of 90,450 patients were included in our analysis (78.5% of patients in the dataset). Excluded patients were significantly older, were more likely to be female, had lower Oxford Knee Scores, EQ-5D index scores, and VAS scores both pre- and post-operatively ($p<0.001$) (Supplemental Table 1). The 90,450 patients were divided into a development sample of 72,360 and a validation sample of 18,090 through random selection (Table 1). There were 139 unique health states

described in the development sample preoperatively and 146 unique health states postoperatively. In the validation sample, there were 110 health states preoperatively and 118 postoperatively.

Model coefficients were consistent and logical, which means that larger value decrements were associated with more problems across the dimensions (i.e., value decrements for level 3 were greater than for level 2) (Table 2). Predicted VAS values ranged from 36.0 to 84.5 preoperatively, and from 21.4 to 84.0 postoperatively. The model had an overall R^2 of 0.39.

In examining the regression coefficients for each dimension of the EQ-5D, extreme problems with anxiety/depression (level 3) was associated with the largest negative values prior to surgery, with a level 3 coefficient of -14.9 (95% CI: -15.6 to -14.3), followed by mobility (-10.4; 95% CI: -12.9, -7.9) and then self-care (-10.2; 95% CI: -11.7, -8.6) (Table 2). Level 3 (extreme problems) with pain/discomfort (-5.3; 95% CI: -6.5, -4.1) and usual activities -7.7 (95% CI: -8.3, -7.2) had comparatively less negative impact. In general, the regression coefficients were more negative for the postoperative valuations (Table 2). There was no difference in the average rating of health if no problems were reported. However, after surgery, further value decrements were observed to a significant extent for the coefficients on several levels and dimensions, including level 2 of mobility, level 2 and 3 of usual activities, level 3 of pain/discomfort, and level 3 of anxiety/depression (all $p<0.001$). The most substantial difference related to level 3 pain/discomfort (-6.0; 95% CI: -7.4, -4.7), which indicated that reporting level 3 (extreme) pain/discomfort after surgery was associated with a more negative impact than prior to surgery. Inclusion of age and gender as main effects showed males to have slightly higher scores on average (2.0; 95%: 1.8-2.2); compared to >80 year olds, patients 60-80 years of age had slightly higher values, and patient 40-60 years of age slightly lower (Supplemental Table S2). However,

the same weights were significantly different post-operatively compared to prior to surgery (Table S2), and similar results were observed across patient subgroups stratified by age and gender (Supplemental Tables S3 to S6).

When applying the EHS value sets to the 10 most common health states among patients before and after surgery, scores were consistently higher when based on the preoperative value set (Table 3). Some health state scores were much lower when based on the post-operative value set (e.g., -12.5 for "22232") (Table 3). As reflected by the coefficients for each of the value sets, i.e. more negative coefficients post-operatively, the more severe the health state, the greater the difference between pre and post-operative value set-based scores.

DISCUSSION

In examining whether experience-based value sets varied with context, we found significant differences in EHS value sets before and after knee replacement surgery. Postoperative coefficients were generally more negative than those of preoperative valuations, and the difference was substantially greater for certain dimensions, particularly pain/discomfort. Results indicated that health state descriptors such as "some" and "extreme" problems were associated with systematically lower values after surgery than before surgery for several dimensions, which raises concerns about the susceptibility of EHS value sets to the context in which they were elicited, e.g. timing of assessment, as illustrated in knee replacement patients.

While patient preferences are typically not used for economic evaluations, social policies, and broad allocation of resources, they are arguably more appropriate than societal preferences in the context of clinically oriented research or individual decision making.[10,24,25] The present study

adds to the literature by demonstrating that EHS value sets depend on context, which can be informative to clinical decision making. In contrast, previous studies have tended to investigate differences between EHS and GHS[9,12-18] rather than comparing across EHS value sets. Some studies found that patients tended to have higher valuations for more severe health states and lower valuations for mild states, compared to the general public or other rater groups.[1,3,4,25] Mann et al. found patient self-rated VAS values were systematically lower compared to population VAS values for the same health states,[26] while the meta-analysis by Dolders et al. showed no difference between patient and population preferences.[27] The stability of value sets as posed in this study could be examined for generalizability by applying the research question to stage of disease, such as testing whether EHS differ by cancer stage for a particular type of cancer.

Comparing the regression coefficients across the EQ-5D dimensions provided valuable information about the relative values that patients assigned to different dimensions of health. Consistent with other findings that anxiety/depression dimension has the greatest impact on VAS values,[12,13] our study revealed that the largest value decrement was associated with problems with anxiety/depression and self-care in patients experiencing knee replacement. We also found that presence of extreme (i.e. level 3) pain/discomfort becomes more important in explaining negative ratings of health after surgery. Previous studies examining patient values of health using happiness or life satisfaction scores have concluded that patients tend to place more value weights on their mental health than other dimensions such as pain and physical function.[28,29] In addition, Mann et al. found that patients and general population valued their health significantly different with regard to pain/discomfort, mobility, and anxiety/depression.[26] In the pre-operative value set, the dimension of anxiety/depression was associated with the largest value decrement, with pain/discomfort initially having a small value decrement, then much more important post-

surgery. These results suggest that mental health and well-being is a primary factor in self-rating of health.

This study is among the first to examine the stability of EHS value sets. Unlike other valuation studies conducted in the general public or patients with stable chronic conditions, the evaluations of health in the present study were more dynamic and interceded by a planned intervention of knee replacement. Our results demonstrated that variation in VAS scores was only partially explained by the health states described by the EQ-5D descriptive system (R^2 = 0.39). The ability of EQ-5D dimension responses to explain variance (i.e., indicated by R^2) in patients undergoing knee replacement was lower as compared with other VAS-based EQ-5D valuation studies (R^2 ranged from 0.53 to 0.97).[11] A study by Whynes et al. that used VAS data from 4 patient groups found they do not value the same health state similarly, which supports our results concerning the lack of consistency of EHS value sets.[30] More studies are needed to see whether these results apply to other clinical contexts.

Our results showed that the reporting of problems on several of the dimensions of the EQ-5D after knee replacement has a more negative impact on the rating of overall health. Thus, we inferred that experience-based value sets can depend on timing, particularly when a medical intervention is performed. This finding further explained our previous study that patients self-rated their improvement following knee replacement systematically lower using VAS, compared to EQ-5D index and OKS.[31] Ratings of health like the VAS can reflect wants, hopes, and expectations for normal health and activity, capturing the difference in pre-surgical expectations and their post-surgical experiences.[32] The disagreement between pre- and post-operative valuations may be due to dissatisfaction with surgical outcomes especially if the intervention failed to meet patient needs and expectations.[32,33] Satisfaction with surgery may be a salient

considerations for procedures that may be elective such as knee replacement surgery. More research and effort are needed to understand the relationship between health needs, satisfaction, and quality of life, and to improve satisfaction and quality of life by minimizing unmet need.[34]

There were some limitations in this study. First, VAS is not grounded in economic theory,[35] and it is unclear if the lack of stability in EHS over time generalizes to utility-based techniques such as the TTO and standard gamble. As a scaling method, VAS may suffer from end-of-scale bias in which respondents tend to avoid using the extremes of the scale in their answers[36] and ignore the duration of disease during a VAS task.[37] Nonetheless, VAS can serve as a feasible and reliable method to elicit patient values and is the predominant basis for generating EHS. Second, age and gender have been identified to relate to differential values in health (i.e., men have higher values than women, and values of health states may increase or decrease with age),[9,38,39] but they did not improve the predictive accuracy of the models we explored (not reported here but results available upon request). Other factors such as socioeconomic status have also been reported to be associated with health valuation,[11,39,40] but were not available in the data so could not be examined. Although descriptive health state classifier items were used to anchor health objectively, there could have been shifts in the how people perceived those items, a possibility which further highlights concerns about deriving stable values sets from an EHS perspective. The change in values could be considered a form of response shift,[41-43] and methodologies such as structural equation modeling may help to further elucidate these relationships. Last, it was unclear how broadly the results can be generalized to conditions beyond knee replacement surgery.

In summary, our study showed how preference weights for health are sensitive to context by estimating and comparing value sets in patients undergoing knee replacement before and after

surgery. Lower values were estimated for patients postoperatively, possibly a result of dissatisfaction when surgery fails to improve health. Results of this study improve our understanding of the strengths and limitations of using patient experience-based value sets, and may be useful to improve patient-centered decision aids by illustrating the impact of timing of the assessment of patient experience on values associated with health status before and after knee replacement surgery.

REFERENCES

1. Dolan P. The effect of experience of illness on health state valuations. *Journal of clinical epidemiology.* 1996;49(5):551-564.
2. De Wit GA, Busschbach JJ, De Charro FT. Sensitivity and perspective in the valuation of health status: whose values count? *Health economics.* 2000;9(2):109-126.
3. Kind P, Dolan P. The effect of past and present illness experience on the valuations of health states. *Medical Care.* 1995:AS255-AS263.
4. Badia X, Diaz-Prieto A, Rue M, Patrick D. Measuring health and health state preferences among critically ill patients. *Intensive care medicine.* 1996;22(12):1379-1384.
5. National Institute of Health and Clinical Excellence (NICE). Guide to the methods of technology appraisal 2013. NICE article [PMG9] Published date: April 2013. Available at: https://www.nice.org.uk/article/pmg9/chapter/5-The-reference-case. Accessed November 24, 2015.
6. Canadian Agency for Drugs and Technologies in Health (CADTH). HTA Guidelines for the Economic Evaluation of Health Technologies: Canada. 3rd Edition, 2006. Available at: https://www.cadth.ca/sites/default/files/pdf/186_EconomicGuidelines_e.pdf. Accessed November 24, 2015.
7. Dobrez D, Cella D, Pickard AS, Lai JS, Nickolov A. Estimation of Patient Preference-Based Utility Weights from the Functional Assessment of Cancer Therapy—General. *Value in Health.* 2007;10(4):266-272.
8. Pickard AS, Shaw JW, Lin HW, et al. A Patient-Based Utility Measure of Health for Clinical Trials of Cancer Therapy Based on the European Organization for the Research

and Treatment of Cancer Quality of Life Questionnaire. *Value in Health.* 2009;12(6):977-988.

9. Leidl R, Reitmeir P. A value set for the EQ-5D based on experienced health states. *Pharmacoeconomics.* 2011;29(6):521-534.

10. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Quality of life Research.* 2003;12(6):599-607.

11. Xie F, Gaebel K, Perampaladas K, Doble B, Pullenayegum E. Comparing EQ-5D Valuation Studies: A Systematic Review and Methodological Reporting Checklist. *Med Decis Making.* Mar 22 2013.

12. Burström K, Sun S, Gerdtham U-G, et al. Swedish experience-based value sets for EQ-5D health states. *Quality of Life Research.* 2014;23(2):431-442.

13. Sun S, Chen J, Kind P, Xu L, Zhang Y, Burström K. Experience-based VAS values for EQ-5D-3L health states in a national general population health survey in China. *Quality of Life Research.* 2014;24(3):693-703.

14. Leidl R, Reitmeir P, König H-H, Stark R. The performance of a value set for the EQ-5D based on experienced health states in patients with inflammatory bowel disease. *Value in Health.* 2012;15(1):151-157.

15. Little MH, Reitmeir P, Peters A, Leidl R. The Impact of Differences between Patient and General Population EQ-5D-3L Values on the Mean Tariff Scores of Different Patient Groups. *Value in Health.* 2014;17(4):364-371.

16. 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(4):799-809.

17. Aronsson M, Husberg M, Kalkan A, Eckard N, Alwin J. Differences between hypothetical and experience-based value sets for EQ-5D used in Sweden: Implications for decision makers. *Scandinavian journal of public health.* 2015;43(8):848-854.

18. Nemes S, Burström K, Zethraeus N, Eneqvist T, Garellick G, Rolfson O. Assessment of the Swedish EQ-5D experience-based value sets in a total hip replacement population. *Quality of Life Research.* 2015;24(12):2963-2970.

19. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Annals of the rheumatic diseases.* 2001;60(2):91-97.

20. Health and Social Care Information Centre. Provisional monthly Patient Reported Outcome Measures (PROMs) in England: a guide to PROMs methodology. Available at: http://www.hscic.gov.uk/media/1537/A-Guide-to-PROMs-Methodology/pdf/PROMS_Guide_v5.pdf. Accessed November 24, 2015.

21. EuroQoL Group. EuroQoL. Available at: http://www.euroqol.org/. Accessed November 24, 2015.

22. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *The Journal of bone and joint surgery.* Jan 1998;80(1):63-69.

23. Resnik L, Dobrykowski E. Outcomes measurement for patients with low back pain. *Orthop Nurs.* Jan-Feb 2005;24(1):14-24.

24. Brazier J, Akehurst R, Brennan A, et al. Should patients have a greater role in valuing health states? *Appl Health Econ Health Policy.* 2005;4(4):201-208.

25. De Wit GA, Busschbach JJ, De Charro FT. Sensitivity and perspective in the valuation of health status: whose values count? *Health economics.* Mar 2000;9(2):109-126.

26. Mann R, Brazier J, Tsuchiya A. A comparison of patient and general population weightings of EQ-5D dimensions. *Health economics.* Mar 2009;18(3):363-372.

27. Dolders MG, Zeegers MP, Groot W, Ament A. A meta-analysis demonstrates no significant differences between patient and population preferences. *Journal of clinical epidemiology.* Jul 2006;59(7):653-664.

28. Dolan P, Lee H, Peasgood T. Losing sight of the wood for the trees: some issues in describing and valuing health, and another possible approach. *Pharmacoeconomics.* Nov 1 2012;30(11):1035-1049.

29. Mukuria C, Brazier J. Valuing the EQ-5D and the SF-6D health states using subjective well-being: a secondary analysis of patient data. *Soc Sci Med.* Jan 2013;77:97-105.

30. Whynes DK. Does the correspondence between EQ-5D health state description and VAS score vary by medical condition? *Health Qual Life Outcomes.* 2013;11:155.

31. Lin F-J, 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. *The European Journal of Health Economics.* 2014;15(5):489-496.

32. Calman KC. Quality of life in cancer patients--an hypothesis. *J Med Ethics.* Sep 1984;10(3):124-127.

33. Shim JK, Russ AJ, Kaufman SR. Clinical life: expectation and the double edge of medical promise. *Health (London).* Apr 2007;11(2):245-264.

34. Asadi-Lari M, Tamburini M, Gray D. Patients' needs, satisfaction, and health related quality of life: towards a comprehensive model. *Health and quality of life outcomes.* Jun 29 2004;2:32.

35. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press; 2005.

36. Torrance GW, Feeny D, Furlong W. Visual analog scales: do they have a role in the measurement of preferences for health states? *Med Decis Making.* Jul-Aug 2001;21(4):329-334.

37. Robinson A, Dolan P, Williams A. Valuing health status using VAS and TTO: what lies behind the numbers? *Soc Sci Med.* Oct 1997;45(8):1289-1297.

38. Badia X, Herdman M, Kind P. The influence of ill-health experience on the valuation of health. *Pharmacoeconomics.* Jun 1998;13(6):687-696.

39. Bernert S, Fernandez A, Haro JM, et al. Comparison of different valuation methods for population health status measured by the EQ-5D in three European countries. *Value Health.* Jul-Aug 2009;12(5):750-758.

40. Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the visual analogue scale. *Qual Life Res.* Dec 1996;5(6):521-531.

41. Ubel PA, Peeters Y, Smith D. Abandoning the language of "response shift": a plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. *Qual Life Res.* May 2010;19(4):465-471.

42. Sprangers MA, Schwartz CE. Do not throw out the baby with the bath water: build on current approaches to realize conceptual clarity. Response to Ubel, Peeters, and Smith. *Qual Life Res.* May 2010;19(4):477-479.

43. Sprangers MA, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med.* Jun 1999;48(11):1507-1515.

Table 1. Characteristics of Knee Replacement Patients in the Developmental, Validation and Total Sample

	Development sample (n=72,360)	Validation sample (n=18,090)	Total (n=90,450)
Age (n, %)			
40 – 50	157 (0.2)	41 (0.2)	198 (0.2)
50 – 60	7099 (9.8)	1751 (9.7)	8,850 (9.9)
60 – 70	25517 (35.3)	6435 (35.6)	31,952 (35.3)
70 – 80	26583 (36.7)	6678 (36.9)	33,261 (36.8)
> 80	7919 (10.9)	1950 (10.8)	9,869 (10.9)
Male (n, %)	30,003 (41.5)	7,451 (41.2)	37,454 (41.4)
Pre-op scores			
Oxford Knee Score (OKS)			
Number of missing (%)	392 (0.5)	107 (0.6)	499 (0.6)
Mean ± SD (range)	19.0 ± 7.7 (0-47)	18.9 ± 7.7 (0-48)	18.9 ± 7.7 (0-48)
EQ-5D index			
Mean ± SD (range)	0.41 ± 0.31 (-0.59-1)	0.41 ± 0.31 (-0.59-1)	0.41 ± 0.31 (-0.59-1)
EQ-VAS			
Mean ± SD (range)	68.1 ±19.8 (0-100)	68.0 ± 19.8 (0-100)	68.0 ± 19.8 (0-100)
Unique EQ-5D states valued	139	110	146
Post-op scores			
Oxford Knee Score (OKS)			
Number of missing (%)	1569 (2.2)	407 (2.2)	1976 (2.2)
Mean ± SD (range)	34.0 ± 10.1 (0-48)	34.0 ± 10.0 (0-48)	34.0 ± 10.1 (0-48)
EQ-5D index			
Mean ± SD (range)	0.71 ± 0.26 (-0.59 -1)	0.70 ± 0.27 (-0.59 -1)	0.71 ± 0.26 (-0.59 -1)
EQ-VAS			
Mean ± SD (range)	71.9 ± 18.9 (0-100)	71.6 ± 19.1 (0-100)	71.8 ± 19.0 (0-100)
Unique EQ-5D states valued	146	118	154

*Age and gender information were missing in 6,320 and 6,386 patients included in the analysis

Table 2. Comparison of preoperative and postoperative value sets using generalized least squares regression

	Preoperative			Postoperative (difference from preoperative)		
	Coefficients	**LL 95% CI**	**UL 95% CI**	**Coefficients**	**LL 95% CI**	**UL 95% CI**
Constant	84.5	83.3	85.7	-0.5	-1.7	0.7
MO2*	-4.6	-5.1	-4.1	-1.7	-2.3	-1.1
MO3	-10.4	-12.9	-7.9	-0.5	-4.8	3.7
SC2	-7.3	-7.6	-7.0	-0.1	-0.6	0.3
SC3	-10.2	-11.7	-8.6	-1.8	-4.1	0.4
UA2*	-3.6	-4.1	-3.2	-1.1	-1.7	-0.6
UA3*	-7.7	-8.3	-7.2	-3.4	-4.3	-2.6
PD2	-1.4	-2.6	-0.2	-0.8	-2.1	0.4
PD3*	-5.3	-6.5	-4.1	-6.0	-7.4	-4.7
AD2	-7.7	-7.9	-7.4	-0.1	-0.5	0.2
AD3*	-14.9	-15.6	-14.3	-2.2	-3.1	-1.2

MO2=mobility level 2; MO3=mobility level 3; SC2=self-care level 2; SC3=self-care level 3; UA2=usual activities level 2; UA3=usual activities level 3; PD2=pain and discomfort level 2; PD3=pain and discomfort level3; AD2=anxiety and depression level 2; AD3=anxiety and depression level 3.

Preoperative value set: VAS = 84.5 – 4.6 MO2 – 10.4 MO3 – 7.3 SC2 – 10.2 SC3 – 3.6 UA2 – 7.7 UA3 – 1.4 PD2 – 5.3 PD3 – 7.7 AD2 – 14.9 AD3. (range 36.0 to 84.5)

Postoperative value set: VAS = 84.0 – 6.3 MO2 – 10.9 MO3 – 7.4 SC2 – 12.0 SC3 – 4.8 UA2 – 11.2 UA3 – 2.2 PD2 – 11.4 PD3 – 7.8 AD2 – 17.1 AD3. (range 21.4 to 84.0)

* $P < 0.001$.

Table 3: Patient experienced health state-based scores for most common health states, based on pre and post-operative value sets

Most Common Health State		Proportion Observed (%)	Pre-operative EHS Value Set Score	Post-operative EHS Value Set Score	Difference (post-pre)
Pre-operative	21221	26.2	74.9	70.7	-4.2
	21231	10.2	71.0	62.1	-8.9
	21222	9.1	67.2	62.9	-4.3
	22232	6.0	56.0	46.3	-9.7
	22221	6.0	67.6	63.3	-4.3
	21232	5.6	63.3	53.8	-9.5
	22231	4.7	63.7	54.1	-9.5
	22222	4.6	59.9	55.4	-4.5
	21121	4.0	78.5	75.5	-3.0
	22332	3.0	51.9	39.9	-12.0
Post-operative	11111	22.8	84.5	84.0	-0.5
	21221	15.3	74.9	70.7	-4.2
	11121	10.8	83.1	81.7	-1.4
	22221	6.3	67.6	63.3	-4.3
	11221	6.1	79.5	77.0	-2.5
	21222	6.0	67.2	63.0	-4.2
	22222	5.8	59.9	55.4	-4.5
	21121	3.7	78.5	75.5	-3.0
	11211	2.9	80.9	79.2	-1.7
	21211	2.0	76.3	73.0	-3.3

EHS= patient experienced health state-based

Table S1. Comparing patients included in and excluded from the analysis

	Included patients	Excluded patients	*P* value
Number of unique patients	90,450	24,756	--
Age, no. (%)*			<.0001 †
40 – 50	198 (0.22)	52 (0.21)	
50 – 60	8,850 (9.87)	1908 (7.71)	
60 – 70	31,952 (35.33)	7429 (30.01)	
70 – 80	33,261 (36.77)	9978 (40.31)	
> 80	9,869 (10.91)	3,779 (15.26)	
Male, no. (%)*	37,454 (41.41)	8,288 (33.48)	<.0001 †
Pre-op scores			
Oxford Knee Score (OKS)			
Number of missing (%)	499 (0.55)	437 (1.77)	
Mean ± SD (range)	18.95 ± 7.73 (0-48)	17.90 ± 7.79 (0-48)	<.0001§
EQ-5D index			
Number of missing (%)	0 (0.00)	6582 (26.59)	
Mean ± SD (range)	0.41 ± 0.31 (-0.594-1)	0.36 ± 0.32 (-0.594-1)	<.0001§
EQ-VAS			
Number of missing (%)	0 (0.00)	11379 (45.96)	
Mean ± SD (range)	68.04 ± 19.83 (0-100)	65.95 ± 21.11 (0-100)	<.0001§
Post-op scores			
Oxford Knee Score (OKS)			
Number of missing (%)	1976 (2.18)	1963 (7.93)	
Mean ± SD (range)	33.99 ± 10.06 (0-48)	32.39 ± 10.61 (0-48)	<.0001§
EQ-5D index			
Number of missing (%)	0 (0.00)	8257 (33.43)	
Mean ± SD (range)	0.71 ± 0.26 (-0.594 -1)	0.68 ± 0.28 (-0.594-1)	<.0001§
EQ-VAS			
Number of missing (%)	0 (0.00)	7937 (32.06)	
Mean ± SD (range)	71.81 ± 18.98 (0-100)	68.95 ± 21.01 (0-100)	<.0001§
Pre-op valuation			
EQ-5D health states valued	146	--	--
Post-op valuation			
EQ-5D health states valued	154	--	--

Based on †Chi-square, or §*t*-test.

*Age and gender information were missing in 6,320 and 6,386 patients included in the analysis and were missing in 1,610 and 1,626 patients excluded from the analysis, respectively.