



Published in final edited form as:

*Med Decis Making*. 2016 February ; 36(2): 264–274. doi:10.1177/0272989X15599546.

## Health Condition Impacts in a Nationally Representative Cross-Sectional Survey Vary Substantially by Preference-Based Health Index

Janel Hanmer, MD, PhD<sup>1</sup>, Dasha Cherepanov, PhD<sup>2</sup>, Mari Palta, PhD<sup>3</sup>, Robert M. Kaplan, PhD<sup>4</sup>, David Feeny, PhD<sup>5</sup>, and Dennis Fryback, PhD<sup>3</sup>

<sup>1</sup>Department of Internal Medicine, University of Pittsburgh, Pittsburgh, PA

<sup>2</sup>Partnership for Health Analytic Research, LLC, Beverly Hills, CA

<sup>3</sup>Population Health Sciences, University of Wisconsin-Madison, Madison, WI

<sup>4</sup>UCLA Department of Health Services, University of California, Los Angeles, CA

<sup>5</sup>Department of Economics, McMaster University, Hamilton, ON, Canada; Health Utilities Incorporated, Dundas, ON

### Abstract

**Importance**—Many cost-utility analyses rely on generic utility measures for estimates of disease impact. Commonly used generic preference-based indexes may generate different absolute estimates of disease burden despite sharing anchors of dead at 0 and full health at 1.0.

**Objective**—We compare the impact of 16 prevalent chronic health conditions using six utility-based indexes of health and a visual analog scale.

**Design**—Data were from the National Health Measurement Study (NHMS), a cross-sectional telephone survey of 3844 adults aged 35–89 in the United States.

**Main Outcome Measures**—The NHMS included the EuroQol-5D-3L, Health and Activities Limitation Index (HALex), Health Utilities Index Mark 2 (HUI2) and Mark 3 (HUI3), preference-based scoring for the SF-36 (SF-6D), Quality of Well-being Scale, and visual analog scale.

Respondents self-reported 16 chronic conditions. Survey-weighted regression analyses for each index with all health conditions, age, and gender were used to estimate health condition impact estimates in terms of quality-adjusted life-years (QALYs) lost over 10 years. All analyses were stratified by ages 35–69 and 70–89.

**Results**—There were significant differences between the indexes for estimates of the absolute impact of most conditions. On average, condition impacts were the smallest with the SF-6D and

---

CORRESPONDING AUTHOR: Janel Hanmer, MD, PhD. Department of General Internal Medicine, University of Pittsburgh Medical Center, 230 McKee Place, Suite 600, Pittsburgh, PA 15213. hanmerjz@upmc.edu. Phone: 412-586-9847, Fax: 412-692-4838.

Parts of these analyses were presented at the 31<sup>th</sup> Annual Meeting of the Society for Medical Decision Making, in Los Angeles, October 18–21, 2009.

It should be noted that David Feeny has a proprietary interest in Health Utilities Incorporated, Dundas, Ontario, Canada. HUInc. distributes copyrighted Health Utilities Index (HUI) materials and provides methodological advice on the use of HUI. None of the other authors declare a conflict of interest.

EQ-5D and the largest with the HALex and HUI3. Likewise, the estimated loss of QALYs varied across indexes. Condition impact estimates for EQ-5D, HUI2, HUI3, and SF-6D generally had strong Spearman correlations across conditions (i.e., > 0.69).

**Limitations**—This analysis uses cross-sectional data and lacks health condition severity information.

**Conclusions**—Health condition impact estimates vary substantially across the indexes. These results imply that it is difficult to standardize results across cost-utility analyses which use different utility measures.

---

## INTRODUCTION

The US Panel on Cost-Effectiveness in Health and Medicine discussed the desirability of having a uniform metric to measure the burden of different health conditions and to assess the value of health interventions. Measures with community-based preference weights have been recommended for use in cost-utility<sup>1</sup> and regulatory analyses.<sup>2</sup> Several such measures are available for these purposes,<sup>3</sup> and no one measure has become the de facto standard in the United States. It is important to ask how much health condition impacts might vary due to the choice of preference-based measure. If condition impacts vary considerably, then consistency in cost-utility analyses cannot be achieved solely by using a member of this family of measures. Also, when the financial acceptability of interventions is informed by a cost per quality-adjusted life-year (QALY) cut off, the selection of a preference-based measure may influence the decision regarding a given intervention's acceptability. Variation due to choice of index limits the usefulness of cost-utility analyses for use in shaping public policy.

In previous work, Franks et al.<sup>4</sup> studied the influence of choice for preference-based measure on estimates of health condition impact in the Medical Expenditures Panel Survey. They used seven different measures: EuroQol-5D-3L (EQ-5D) with UK weights, EQ-5D with US weights, the EQ-5D visual analog scale (VAS), the SF-6D based on SF-12 questions, and three imputed scores using regressions on the SF-12 questions. Franks et al. concluded that absolute (i.e., original) incremental cost-effectiveness analyses of a given problem would likely vary depending on the measure used, whereas the relative ordering of incremental cost-effectiveness analyses of a series of problems would likely be similar if a single measure was chosen.<sup>4</sup> One of the limitations of that investigation was that only two of the measures were directly administered while the others were imputed.

In this report, we present impact estimates for 16 health conditions based on six directly administered preference-based indexes and a visual analog scale, which were co-administered in a probability sample of the adult non-institutionalized US population. The measures in this report include the most widely used multi-attribute preference-based indexes. Direct rating on a visual analog scale is a common tool for estimating the value of health states.<sup>3</sup> We illustrate the importance of measure selection on QALY estimates and compare health condition impact rankings among these different measures.

## METHODS

### Data

We used data from the National Health Measurement Study (NHMS). The NHMS methods and measures are described elsewhere<sup>5</sup> and are publicly available at the National Archive of Computerized Data on Aging.<sup>6</sup> Fielded in 2005–2006, the NHMS was a cross-sectional, random, digit-dialed, computer-assisted telephone interview survey of community-dwelling US adults aged 35–89. The NHMS oversampled telephone exchanges with high proportions of African American households, as well as people aged 65 and older. The simple response rate was 56%.

The NHMS interview consisted of four health-related quality-of-life (HRQoL) questionnaires administered in random order to each individual: EQ-5D with VAS, Health Utilities Index (HUI) interviewer-administered form, SF-36v2™, and Quality of Well-being Scale (QWB-SA) with VAS. The only exception to the random order was that the five-category (excellent, very good, good, fair, poor), self-rated health question included in the SF-36v2™ was always administered before any of the HRQoL questionnaires and was not repeated when the SF-36v2™ questionnaire was administered. After the four questionnaires were administered, additional questions were asked to allow computation of the Health and Activities Limitations Index (HALex). Later in the survey, respondents were asked to self-report health conditions (see Independent Variables section).

### Dependent Variables

**Health Utility Scores**—Dependent variables include six health utility scores: the EQ-5D-3L<sup>7,8</sup>, Health and Activities Limitation Index (HALex)<sup>9</sup>, Health Utilities Index Mark 2 (HUI2) and Mark 3 (HUI3)<sup>10–12</sup>, Self-administered Quality of Well-being Scale (QWB-SA)<sup>13,14</sup>, and the SF-6D as calculated from the SF-36v2™<sup>15</sup>. Details of these measures and scores are provided in Table 1.

**Visual Analog Scale (VAS)**—Because data were collected via telephone, analog scale ratings of health were collected using words to describe the VAS scale. Hence, ratings were, strictly speaking, not visually based. With this understanding, we still refer to these ratings as “VAS” ratings for simplicity. A VAS is included in the QWB-SA instrument. The question reads: “Think about a scale of 0 to 100, with zero being the least desirable state of health that you could imagine and 100 being perfect health. What number, from 0 to 100, would you give the state of your health, on average, over the last 3 days?” Another VAS is included at the end of the EQ-5D instrument. The instruction reads: “Now I would like to ask you to rate your health. To help you say how good or bad your health is, I’d like you to picture in your mind a scale that looks like a thermometer. The best health state you can imagine is marked 100 at the top of the scale, and the worst state you can imagine is 0 at the bottom. Tell me the point on this scale where you would rate your own health state today.” VAS scores for both instruments ranged from 0 to 100. We averaged the two scores to form a single VAS score for each respondent.

## Independent Variables

Presence or absence of several self-reported health conditions were collected from survey participants. Questions were formulated as, “Has a doctor or other health professional ever told you that you had: coronary heart disease or a heart attack, also known as myocardial infarction or MI?; a stroke?; diabetes or high blood sugar?; arthritis?; any kind of eye disease, such as cataracts, macular degeneration or glaucoma?; a sleep disorder?; a chronic respiratory or lung disease, such as asthma, emphysema or chronic bronchitis?; clinical depression or anxiety disorder?; an ulcer? This could be a stomach, duodenal or peptic ulcer.; a thyroid disorder?; severe chronic back pain?”

Those reporting coronary artery disease were further subdivided into those who did and did not report currently taking medications for chest pain. Those reporting diabetes were further divided into those currently taking insulin, those currently taking other medications besides insulin, and those currently not taking any medications for diabetes. Those reporting chronic lung diseases were subdivided into those with asthma, emphysema, and chronic bronchitis. We limited those reporting a thyroid disorder to those currently taking prescription medications for the thyroid disorder. We also limited those reporting severe chronic back pain to those who had been told the “pain is caused by a herniated or bulging disk in [their] spine.”

Age in years was included as a continuous variable. As the models were stratified for ages 35–69 and 70–89, assuming that the effects of age are linear was reasonable. Health conditions and gender were dichotomous variables.

## Analyses

Respondents were stratified into ages 35–69 and 70–89 because HALex administration and the health states an individual can report change at age 70. Unweighted descriptive sample statistics were calculated for the two age groups. All other analyses employed survey weights and stratification to produce estimates reflecting the underlying US population. Weights reflect the sampling probability for each participant and post-stratification to the 2000 US Census population by age, gender, and race. A result was determined to be statistically significant if  $P < 0.05$ . The impact of a health condition was determined to be clinically important if it was  $> 0.03$  for all measures except the VAS.<sup>16</sup> Estimates of condition impacts were created by regressing each of the seven utility scores on the 16 health conditions, age, and sex within the two age groups, using PROC SURVEYREG in SAS 9.1 (The SAS Institute, Cary, NC). The coefficient estimate for a health condition is its impact and was tested for statistically significant difference from 0 by a robust  $t$  test. For example, one set of impact estimates for those aged 35–69 were created by regressing EQ-5D scores within this age group on age, sex, and the 16 health conditions.

The health condition impact estimates, based on each one of the six preference-based measures and VAS, were used to estimate the loss in QALYs for each health condition during a 10-year window with a 3% discount rate<sup>1</sup> using Equation 1.

$$10\text{-Year QALY Loss for Condition X} = \sum_{i=0}^9 (\text{Condition X impact estimate}) (0.97^i) \quad \text{EQ 1}$$

A difference of 10% across QALY estimates is considered important.<sup>17</sup> We estimated the percentage change in QALY estimates in four steps. First, we used the constant from the regression results as the “no condition utility estimate.” Second, we calculated the No Condition QALY estimate during a 10-year window with a 3% discount rate using Equation 2.

$$10\text{-Year No Condition QALY for Instrument Y} = \sum_{i=0}^9 (\text{No condition utility estimate for instrument Y}) (0.97^i) \quad \text{EQ 2}$$

Third, we calculated the total expected QALY for those with a condition as the No Condition QALY – 10-year QALY Loss for Condition X. Fourth, we used EQ-5D as the basis of comparison to calculate the percentage change in 10-year QALY for a health condition.

We also calculated the Spearman correlation coefficient between each pair of impact estimates across conditions for the two age groups as a measure of concordance of rankings of the conditions.

To test differences in impact between age groups, an additional model was fit for each instrument that included the scores from both age groups and interactions between age group and condition. A robust *t* test was used to evaluate the significance of the regression coefficient for each interaction term.

To determine if the health condition impact was significantly different across the different instruments, multivariate, repeated-measures models with pooled scores on all the instruments as the outcome were fit within an age group, with interaction terms between the health conditions and instrument indicators. Interactions for each condition were tested with a robust *F* test, taking into account correlation between measures on the same person using PROC SURVEYREG. One model included all seven instruments, recoding VAS to the 0–1 scale by dividing the score by 100. A reduced model included all instruments except HALex and VAS.

## RESULTS

Table 2 shows descriptive statistics on the NHMS sample. More than half the sample were women (56% in the younger group [n = 2710] and 60% in the older group [n=1134]). The table includes prevalence rates of reported health conditions which range from 3% (emphysema in the younger group) to 60% (eye disease in the older group). Rates were generally higher in the older group than the younger group. These results are similar to those of the 2005 National Health Interview Survey,<sup>18</sup> a large nationally representative survey collected by the Centers for Disease Control (Supplemental Table 1). A full description of the sample is provided elsewhere.<sup>5</sup>

Table 3 catalogs the estimated impact of the 16 health conditions for each of the seven measures in the younger and older age groups. Each column of the table shows the adjusted estimate of mean difference in utility for the respective instrument. Statistically significant health condition impacts are bolded. The number of health condition impact estimates that reached clinical significance varied across instruments. In the younger age group, the range of health condition impact estimates that reached clinical significance was 9 (EQ-5D) to 13 (HALex and QWB-SA). In the older age group, the range was 6 (QWB-SA) to 11 (HALex). All health condition impact estimates that were statistically significant were also clinically important.

Table 3 also includes the average of the impact estimates for the 16 health conditions for each measure. The measures with the smallest average impacts were EQ-5D (−0.047) and SF-6D (−0.040) in the younger age group and EQ-5D (−0.033) and QWB-SA (−0.034) in the older age group. The measures with the largest average impacts were HALex (−0.085 and −0.070) and HUI3 (−0.084 and −0.056) in both age groups. The bottom row of Table 3 shows each measure's average impact relative to the EQ-5D's average impact by dividing the average for a measure by −0.047 for those aged 35–69 and by −0.033 for those aged 70–89. For example, the average of the impact estimates for the 16 health conditions was 81% greater for HALex than for EQ5D in patients aged 35 to 89 years.

The health condition impact estimates were generally similar between the younger and older age groups. Statistically significant differences were present between age groups for the QWB-SA (depression/anxiety, GI, ulcer), VAS (depression/anxiety, emphysema, diabetes without medications), SF-6D (emphysema, diabetes without medications, GI ulcer, arthritis), and HALex (CHD without chest pain medications, eye disease, GI ulcer).

For the younger age group, when all instruments were modeled together, there was significant interaction terms between 10 of the 16 health conditions and the instrument used. This indicates that the impact of these health conditions is statistically significantly different when measured by different instruments. When the model was restricted to include only the five measures with community-based preference scoring algorithms (EQ-5D, HUI2, HUI3, SF-6D, QWB-SA), there was a statistically significant interaction between the health condition and instrument used for 6 of the 16 health conditions. For the older group, 5 of the 16 health conditions had statistically significant interaction between the health condition and instrument used in the model with all instruments. In the older age group with the restricted model, 4 of the 16 health conditions had this significant interaction.

Table 4 includes the estimated change in QALYs by the presence of a health condition in an individual during the course of 10 years with a 3% discount rate for each age group. The table illustrates that the choice of measure will influence the final QALY estimate. For example, for those with diabetes using insulin in the older group, the difference in QALYs ranged from −0.09 (QWB-SA) to −1.14 (HALex). There are 96 possible comparisons to the EQ-5D when looking at percentage change in QALY estimates within each age group. Of the 96 possible comparisons, there was a greater than 10% difference in 41 comparisons in the younger age group and 28 comparisons in the older age group.

Despite the difference in absolute values of estimates for a health condition impact across the different measures, the ranking of health condition impacts by different measures were often similar. Table 5 includes the pairwise Spearman correlations between the seven different instruments across estimates for the 16 health conditions for ages 35–69 and 70–89. The EQ-5D, HUI2, HUI3, and SF-6D have Spearman correlations above 0.69 with other instruments in both the younger and older age groups, except for EQ-5D with HUI2 and HUI3 in the older age group at 0.51 and 0.59, respectively.

## DISCUSSION

Based on simultaneous administration of seven measures to a sample of US adults, we found that the health-impact estimates, and QALY estimates calculated from them, may vary substantially by measurement system. The measures with the smallest average health-impact estimates were the EQ-5D and SF-6D. The measures with the largest average health-impact estimates were the HALex and HUI3. Consistent with differences in impact of health conditions shown in this report, a recent item response theory analysis using these same NHMS data indicates these measures assign different decrements to the same change in a latent joint construct of health.<sup>19</sup>

We found that the EQ-5D, HUI2, HUI3, and SF-6D generally rank conditions similarly in the younger and older age groups, except for EQ-5D with HUI2 and HUI3 in the older age group. The QWB-SA had strongly correlated health-impact estimates with the other community-based preference algorithms in the older age group. These correlations suggest all of these different measurement systems come close to giving health conditions a similar relative value but differ in absolute value assigned. The health condition impact estimates from the other preference-based measures included in this study, the HALex and VAS, did not correlate as highly with the measurement systems based on community-based preference scores.

The intent of this study is not to create condition impact estimates for clinical use but to compare health utility measures. The ideal source of QALY estimates for clinical use are from randomized head-to-head comparisons of treatment alternatives with pre- and post-intervention HRQoL measurement.

Our findings have implications for cost-utility analysis. In the present study, the incremental QALY change associated with common health conditions could well surpass or fall short of the 10% change threshold commonly used as an “important difference” depending on which utility measure is selected for the analysis. In health care systems where interventions may be rationed by cost per QALY cutoffs or comparisons, the choice of preference-based measurement system can affect the acceptability of a given intervention. For example, an intervention is more likely to reach acceptability or be more cost-effective than another intervention when estimates for QALYs are based on the HUI3 rather than the EQ-5D, as the average condition impact measured by HUI3 is 1.7 times larger than by the EQ-5D. This difference across measures can have important consequences when results from cost-utility analyses are used by decision makers.<sup>20,21</sup> Further work could explore the feasibility and



limitations of standardized transformations to make the comparisons across measures (both across diseases and within a disease) less problematic.

To maximize comparability across studies, a single preference-based measure could be agreed upon as having the best validity and measurement properties and then be uniformly applied. Organizations such as Great Britain's National Institute for Clinical Excellence request that the EQ-5D (and UK scoring algorithm) be used in the reference case.<sup>24</sup> In contrast, health technology agencies in Australia and Canada leave the choice of generic preference-based measure up to the analyst. For many reasons, no such recommendations have been put forth for US analyses, though an IOM panel recommended the ED-5D for regulatory analyses.<sup>2</sup> Thus the choice of measure is often based on an incentive to use the same measure in subsequent studies of a given clinical condition or disease to enhance comparability.<sup>1,2,25,24</sup> It should also be noted that while the use of a single measure in all studies would presumably enhance comparability, in many circumstances, it might also attenuate validity.<sup>25</sup>

Our study provides no guidance regarding which of the measures with community-based preference scores to use as the reference case. It is not clear which measure yields the most valid preference estimates for the general population or for specific subgroups, as the measures used different populations and techniques (such as standard gamble or time trade-off) to construct scoring algorithms. The various measures have different susceptibility to ceiling and floor effects in different populations, provide different definitions of what constitutes "full" or "perfect" health, and differ on whether "worse than dead" health states exist.<sup>19</sup> The measures vary by aspects of health captured, measurement error and reliability, readability, rates of missing values, time burden, and costs of administration.<sup>26-28</sup>

In the long run, to enhance comparability across studies, the development of a preference-based measure with high levels of test-retest reliability, cross-sectional construct validity, longitudinal construct validity and responsiveness, a lack of floor and ceiling effects, very broad applicability, and a sound theoretical and empirical basis for its preference-based scoring system should be a priority. Harmonization of measures has the potential to substantially improve the practice of clinical outcomes research, cost-effectiveness analysis, and population monitoring.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The National Health Measurement Study was funded by a grant from the National Institute on Aging (AG020679). Janel Hanmer was supported by the National Institutes of Health through Grant Number KL2 TR000146. The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

The authors would like to express their gratitude to the individuals who participated in the National Health Measurement Study. Parts of these analyses were presented at the 31<sup>th</sup> Annual Meeting of the Society for Medical Decision Making in Los Angeles, October 18–21, 2009. The National Health Measurement Study was funded by a grant from the National Institute on Aging (AG020679). Janel Hanmer was supported by the National Institutes of



Health through Grant Number KL2 TR000146. The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

## References

1. Gold, MR.; Russell, LB.; Weinstein, MC., editors. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.
2. Miller, W.; Robinson, LA.; Lawrence, RS. Valuing Health for Regulatory Analysis. Washington, DC: The National Academies Press; 2006.
3. Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the measurement of health utilities in published cost-effectiveness analyses. *Value Health*. 2006 Jul-Aug;9(4):213–8. [PubMed: 16903990]
4. 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 analyses. *Med Care*. 2006; 44:478–85. [PubMed: 16641667]
5. Fryback DG, Dunham N, Palta M, et al. The National Health Measurement Study: Simultaneous U.S. Norms for Six Generic Health-Related Quality-of-Life Instruments. *Med Care*. 2007; 45:1162–1170. [PubMed: 18007166]
6. National Archive of Computerized Data on Aging. [Accessed February 10, 2015] United States National Health Measurement Study, 2005–2006. <http://www.icpsr.umich.edu/icpsrweb/NACDA/studies/23263?geography%5b0%5d=United+States&paging.startRow=226>
7. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001; 33:337–343. [PubMed: 11491192]
8. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005; 43:203–220. [PubMed: 15725977]
9. Erickson P. Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Qual Life Res*. 1998; 7:101–114. [PubMed: 9523491]
10. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and singleattribute utility functions for the health utilities index mark 3 system. *Med Care*. 2002; 40:113–128. [PubMed: 11802084]
11. Feeny, D.; Torrance, G.; Furlong, W. Health Utilities Index. In: Spilker, B., editor. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia, PA: Lippincott-Raven Press; 1996.
12. Torrance, George W.; Feeny, David H.; Furlong, William J.; Barr, Ronald D.; Zhang, Yueming; Wang, Qinan. Multi-Attribute Preference Functions for A Comprehensive Health Status Classification System: Health Utilities Index Mark 2. *Med Care*. Jul; 1996 34(7):702–722. [PubMed: 8676608]
13. Kaplan RM, Sieber WJ, Ganiats TG. The Quality of Well-being Scale: comparison of the interviewer-administered version with a self-administered questionnaire. *Psychol Health*. 1997; 12:783–791.
14. Andresen EM, Rothenberg BM, Kaplan RM. Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults. *Med Care*. 1998; 36:1349–1360. [PubMed: 9749658]
15. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care*. 2004; 42:851–859. [PubMed: 15319610]
16. Feeny, David; Spritzer, Karen; Hays, Ron D.; Liu, Honghu; Ganiats, Theodore; Kaplan, Robert M.; Palta, Mari; Fryback, Dennis G. Agreement About Identifying Patients Who Change Over Time: Cautionary Results in Cataract and Heart Failure Patients. *Medical Decision Making*. Mar-Apr; 2012 32(2):273–286. published online October 18, 2011. [PubMed: 22009666]
17. Revicki, Dennis A.; Feeny, David; Hunt, Timothy L.; Cole, Bernard. Analyzing Oncology Clinical Trials Data using the Q-TWiST Method: Clinical Importance and Sources for Health State Preference Data. *Quality of Life Research*. Apr; 2006 15(3):411–423. [PubMed: 16547779]
18. Centers for Disease Control and Prevention. [Accessed February 10, 2015] National Health Interview Survey. 2005 Data Release. [http://www.cdc.gov/nchs/nhis/nhis\\_2005\\_data\\_release.htm](http://www.cdc.gov/nchs/nhis/nhis_2005_data_release.htm)

19. Fryback DG, Palta M, Cherepanov D, Bolt D, Kim JS. Comparison of 5 health-related quality-of-life indexes using item response theory analysis. *Med Decis Making*. 2010; 30:5–15. [PubMed: 19843961]
20. Laupacis, Andreas. Inclusion of drugs in provincial drug benefit programs: Who is making these decisions, and are they the right ones? *Canadian Medical Association Journal*. Jan 8; 2002 166(1): 44–47. [PubMed: 11800247]
21. Laupacis, Andreas. Incorporating Economic Evaluations into Decision-Making: The Ontario Experience. *Medical Care*. Jul; 2005 43(7 Supplement):II-15–II-19.
22. Guide to the methods of technology appraisal. National Institute for Health and Clinical Excellence; 2008.
23. Kopec JA, Willison KD. A comparative review of four preference-weighted measure of health-related quality of life. *J Clin Epi*. 2003; 56:317–325.
24. Paz, Sylvia H.; Liu, Honghu; Fongwa, Marie N.; Morales, Leo S.; Hays, Ron D. Readability Estimates for Commonly Used Health-Related Quality of Life Surveys. *Quality of Life Research*. Sep; 2009 18(7):889–900. E-publication July 10, 2009. [PubMed: 19590979]
25. Feeny, David. Standardization and Regulatory Guidelines May Inhibit Science and Reduce the Usefulness of Analyses Based on the Application of Preference-Based Measures for Policy Decisions. *Medical Decision Making*. Apr; 2013 33(3):316–319. published online November 26, 2012. [PubMed: 23184461]
26. Cherepanov D, Palta M, Fryback DG. Underlying Dimensions of the Five Health-Related Quality-of-Life Measures Used in Utility Assessment Evidence From the National Health Measurement Study. *Med Care*. 2010; 48(8):718–25. [PubMed: 20613664]
27. Palta, Mari, et al. Standard error of measurement of 5 health utility indexes across the range of health for use in estimating reliability and responsiveness. *Medical Decision Making*. 2011; 31(2): 260–269. [PubMed: 20935280]
28. Raisch, Dennis W., et al. Baseline comparison of three health utility measures and the feeling thermometer among participants in the action to control cardiovascular risk in diabetes trial. *Cardiovasc Diabetol*. 2012; 11(1):35. [PubMed: 22515638]

**Table 1**

Details of the Health Utility Measures Used in This Study

Measure	Health Domains (number of levels)	Score Range	Scoring Function Origin
EQ-5D-3L	Mobility (3), self-care (3), usual activities (3), pain/discomfort (3), anxiety/depression (3)	-0.11 to 1.0	Time tradeoff from a sample of United States adults
HALex	Self-reported health (5), and activity limitation (6 for those under age 70, 3 for those age 70 and older)	0.10 to 1.0	Correspondence analysis to Health Utilities Index Mark 1
Health Utilities Index Mark 2	Sensation (4), mobility (5), emotion (5), cognition (4), self-care (4), and pain (5)	-0.03 to 1.0	Standard gamble and visual analog scale from a sample of Canadian adults
Health Utilities Index Mark 3	Vision (6), hearing (6), speech (5), ambulation (6), dexterity (6), emotion (5), cognition (5), and pain (5)	-0.36 to 1.0	Standard gamble and visual analog scale from a sample of Canadian adults
Quality of Well-being Scale	Mobility (3), physical activity (3), social activity (5), symptom/problem complex (27)	0.09 to 1.0	Visual Analog Scale from a sample of United States adults
SF-6D	Physical functioning (6), role limitations (4), social functioning (5), pain (6), mental health (5), vitality (5)	0.30 to 1.0	Standard Gamble in a sample of United Kingdom adults

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Unweighted Descriptive Statistics in Patients Aged 35–69 Years and 70–89 Years in the National Health Measurement Study

Characteristic	Age, y, 35–69 (N = 2710)	Age, y, 70–89 (N = 1134)
Female (%)	56	60
CHD <sup>a</sup> with chest pain medications, No. (%)	102 (3.8)	116 (10.2)
CHD without chest pain medications, No. (%)	123 (4.5)	142 (12.5)
Stroke, No. (%)	106 (3.9)	116 (10.2)
Diabetes and using insulin, No. (%)	125 (4.6)	72 (6.3)
Diabetes and using medications which are not insulin, No. (%)	209 (7.7)	136 (12.0)
Diabetes without medication, No. (%)	112 (4.1)	72 (6.3)
Arthritis, No. (%)	919 (33.9)	638 (56.3)
Eye disease, No. (%)	440 (16.2)	687 (60.1)
Sleep disorder, No. (%)	268 (9.9)	88 (7.8)
Asthma, No. (%)	216 (8.0)	81 (7.1)
Emphysema, No. (%)	78 (2.9)	72 (6.3)
Chronic bronchitis, No. (%)	116 (4.3)	68 (6.0)
Depression/anxiety, No. (%)	459 (16.9)	100 (8.8)
GI <sup>b</sup> ulcer, No. (%)	299 (11.0)	193 (17.0)
Thyroid disorder currently taking medications, No. (%)	187 (6.9)	146 (12.9)
Chronic back pain from disk herniation, No. (%)	252 (9.3)	98 (8.6)
Reporting no conditions, No. (%)	969 (35.8)	110 (9.7)
Reporting 1 condition, No. (%)	741 (27.3)	274 (24.2)
Reporting 2 conditions, No. (%)	465 (17.2)	308 (27.2)
Reporting 3 conditions, No. (%)	255 (9.4)	225 (19.8)
Reporting 4 conditions, No. (%)	134 (4.9)	139 (12.3)
Reporting 5 or more conditions, No. (%)	146 (5.4)	78 (6.9)

<sup>a</sup>CHD = coronary heart disease

<sup>b</sup>GI = gastrointestinal

**Table 3**

Catalog of Health Condition Impacts for the Seven Health Utility Measures, Stratified by Age<sup>d</sup>

Each column of the table includes the regression results an individual instrument. The individual instrument was regressed on age, sex, and presence of the sixteen health conditions. Estimates represent the difference in mean utility based on each instrument between those with and without the specific condition, while adjusting for age, sex, and other conditions. Statistically significant health condition impacts are bolded. In general, a clinically important difference in these measures are between 0.03 and 0.05.

	Age, y, 35-69						
	EQ-5D <sup>b</sup> Estimate (SE)	HALex <sup>c</sup> Estimate (SE)	HUI2 <sup>d</sup> Estimate (SE)	HUI3 Estimate (SE)	SF-6D Estimate (SE)	QWB-SA <sup>e</sup> Estimate (SE)	VAS <sup>f</sup> Estimate (SE)
Constant	<b>0.93 (0.0078)</b>	<b>0.90 (0.0096)</b>	<b>0.90 (0.0079)</b>	<b>0.89 (0.012)</b>	<b>0.83 (0.0078)</b>	<b>0.73 (0.011)</b>	<b>89 (0.8)</b>
Age, y, 35	0.0004 (0.0004)	-0.0007 (0.0005)	0.0004 (0.0004)	0.0011 (0.00066)	0.0006 (0.0004)	-0.0006 (0.0004)	0.029 (0.048)
Male	0.005 (0.007)	0.009 (0.010)	0.015 (0.008)	0.008 (0.012)	0.014 (0.007)	<b>0.024 (0.009)</b>	-0.58 (0.84)
CHD <sup>g</sup> with chest pain medications	-0.052 (0.029)	<b>-0.18 (0.038)</b>	<b>-0.062 (0.029)</b>	<b>-0.10 (0.045)</b>	-0.043 (0.025)	<b>-0.043 (0.018)</b>	<b>-7.8 (3.3)</b>
CHD without chest pain medications	-0.026 (0.018)	<b>-0.12 (0.040)</b>	-0.049 (0.027)	-0.053 (0.037)	<b>-0.040 (0.016)</b>	<b>-0.062 (0.019)</b>	<b>-6.5 (2.5)</b>
Stroke	-0.029 (0.031)	-0.027 (0.068)	-0.041 (0.052)	-0.097 (0.072)	-0.017 (0.039)	<b>-0.094 (0.023)</b>	<b>-12 (3.8)</b>
Diabetes and using insulin	-0.025 (0.030)	<b>-0.16 (0.034)</b>	-0.051 (0.037)	-0.059 (0.052)	-0.036 (0.021)	-0.030 (0.021)	<b>-10 (4.3)</b>
Diabetes and using medications which are not insulin	0.00021 (0.021)	<b>-0.10 (0.029)</b>	-0.056 (0.030)	-0.070 (0.039)	-0.030 (0.016)	<b>-0.039 (0.016)</b>	<b>-5.2 (2.2)</b>
Diabetes without medication	-0.036 (0.024)	<b>-0.078 (0.035)</b>	<b>-0.055 (0.023)</b>	<b>-0.088 (0.032)</b>	<b>-0.038 (0.016)</b>	-0.044 (0.033)	<b>-7.1 (2.4)</b>
Arthritis	<b>-0.076 (0.0092)</b>	<b>-0.053 (0.014)</b>	<b>-0.062 (0.011)</b>	<b>-0.087 (0.016)</b>	<b>-0.044 (0.0090)</b>	<b>-0.057 (0.0089)</b>	<b>-5.2 (1.0)</b>
Eye disease	-0.021 (0.011)	<b>-0.035 (0.018)</b>	-0.023 (0.014)	<b>-0.044 (0.021)</b>	-0.011 (0.0099)	<b>-0.047 (0.011)</b>	-0.77 (1.4)
Sleep disorder	<b>-0.096 (0.021)</b>	<b>-0.098 (0.025)</b>	-0.10 (0.022)	<b>-0.13 (0.032)</b>	<b>-0.054 (0.014)</b>	<b>-0.054 (0.013)</b>	<b>-4.9 (1.8)</b>
Asthma	-0.019 (0.017)	<b>-0.061 (0.025)</b>	0.0025 (0.019)	-0.036 (0.033)	-0.024 (0.013)	<b>-0.046 (0.013)</b>	-1.0 (2.2)
Emphysema	<b>-0.073 (0.026)</b>	<b>-0.10 (0.038)</b>	<b>-0.091 (0.038)</b>	<b>-0.12 (0.058)</b>	<b>-0.051 (0.022)</b>	-0.026 (0.027)	<b>-16 (4.3)</b>

Med Decis Making. Author manuscript; available in PMC 2017 February 01.

Age, y, 35–69							
	EQ-5D <sup>v</sup> Estimate (SE)	HALex <sup>c</sup> Estimate (SE)	HUI2 <sup>d</sup> Estimate (SE)	HUI3 Estimate (SE)	SF-6D Estimate (SE)	QWB-SA <sup>e</sup> Estimate (SE)	VAS <sup>f</sup> Estimate (SE)
Chronic bronchitis	-0.063 (0.029)	-0.067 (0.038)	-0.058 (0.032)	-0.095 (0.050)	-0.035 (0.022)	-0.050 (0.020)	-9.1 (2.7)
Depression/anxiety	-0.099 (0.013)	-0.095 (0.019)	-0.12 (0.016)	-0.15 (0.023)	-0.083 (0.010)	-0.11 (0.0099)	-9.4 (1.5)
GI <sup>h</sup> ulcer	-0.031 (0.013)	-0.057 (0.022)	-0.040 (0.014)	-0.048 (0.023)	-0.044 (0.013)	-0.038 (0.012)	-4.4 (1.4)
Thyroid disorder currently taking medications	-0.0063 (0.014)	-0.016 (0.019)	0.0011 (0.016)	-0.021 (0.029)	-0.0065 (0.012)	-0.0011 (0.01)	0.33 (1.9)
Chronic back pain from disk herniation	-0.092 (0.019)	-0.12 (0.024)	-0.11 (0.022)	-0.15 (0.033)	-0.08 (0.015)	-0.054 (0.015)	-6.9 (2.0)
R <sup>2</sup>	0.33	0.37	0.34	0.31	0.28	0.30	0.31
Average of the estimates for the 16 health conditions	-0.047	-0.085	-0.057	-0.084	-0.040	-0.050	-0.066
Average of the estimates for the 16 health conditions divided by the average EQ-5D estimate of -0.047	1.00	1.81	1.21	1.79	0.85	1.06	1.40
Age, y, 70–89							
	EQ-5D Estimate (SE)	HALex Estimate (SE)	HUI2 Estimate (SE)	HUI3 Estimate (SE)	SF-6D Estimate (SE)	QWB-SA Estimate (SE)	VAS Estimate (SE)
Constant	0.92 (0.013)	0.88 (0.019)	0.93 (0.014)	0.92 (0.022)	0.85 (0.013)	0.73 (0.013)	88 (1.6)
Age, y, 70	-0.0009 (0.00010)	-0.0038 (0.0015)	-0.0018 (0.0011)	-0.0045 (0.0019)	-0.0016 (0.00094)	-0.0031 (0.0011)	-0.18 (0.14)
Male	0.030 (0.0099)	0.042 (0.016)	0.036 (0.012)	0.034 (0.019)	0.023 (0.011)	0.002 (0.011)	2.5 (1.4)
CHD with chest pain medications	-0.049 (0.019)	-0.14 (0.028)	-0.10 (0.029)	-0.15 (0.050)	-0.071 (0.017)	-0.071 (0.021)	-12 (2.6)
CHD without chest pain medications	-0.0077 (0.016)	-0.025 (0.022)	-0.0032 (0.017)	0.0078 (0.026)	-0.0099 (0.018)	-0.030 (0.017)	-2.4 (1.8)
Stroke	-0.051 (0.017)	-0.15 (0.032)	-0.069 (0.022)	-0.13 (0.038)	-0.051 (0.017)	-0.042 (0.02)	-7.1 (2.6)
Diabetes and using insulin	-0.047 (0.028)	-0.13 (0.039)	-0.015 (0.034)	-0.081 (0.062)	-0.044 (0.022)	-0.0099 (0.028)	-4.1 (4.0)

	Age, y, 70–89							
	EQ-5D Estimate (SE)	HALex Estimate (SE)	HUI2 Estimate (SE)	HUI3 Estimate (SE)	SF-6D Estimate (SE)	QWB-SA Estimate (SE)	VAS Estimate (SE)	
Diabetes and using medications which are not insulin	-0.030 (0.019)	<b>-0.092 (0.028)</b>	-0.0031 (0.021)	0.0032 (0.039)	<b>-0.033 (0.017)</b>	-0.020 (0.019)	<b>-5.9 (2.2)</b>	
Diabetes without medication	-0.0093 (0.020)	0.0041 (0.031)	-0.029 (0.040)	-0.031 (0.051)	0.013 (0.017)	-0.0079 (0.022)	3.8 (2.2)	
Arthritis	<b>-0.070 (0.0098)</b>	<b>-0.090 (0.016)</b>	<b>-0.064 (0.012)</b>	<b>-0.097 (0.019)</b>	<b>-0.074 (0.011)</b>	<b>-0.070 (0.010)</b>	<b>-6.9 (1.3)</b>	
Eye disease	-0.0045 (0.010)	0.0038 (0.016)	-0.018 (0.012)	-0.020 (0.020)	-0.013 (0.010)	<b>-0.025 (0.011)</b>	-1.1 (1.4)	
Sleep disorder	<b>-0.078 (0.022)</b>	<b>-0.11 (0.032)</b>	<b>-0.063 (0.030)</b>	-0.067 (0.048)	<b>-0.072 (0.021)</b>	<b>-0.087 (0.024)</b>	<b>-7.5 (3.0)</b>	
Asthma	-0.0037 (0.016)	-0.051 (0.029)	-0.024 (0.031)	-0.022 (0.045)	<b>-0.055 (0.017)</b>	-0.011 (0.021)	-4.1 (3.0)	
Emphysema	-0.0073 (0.021)	<b>-0.12 (0.041)</b>	-0.025 (0.037)	-0.023 (0.048)	-0.027 (0.018)	-0.020 (0.022)	-4.7 (3.2)	
Chronic bronchitis	0.014 (0.027)	<b>-0.077 (0.037)</b>	-0.034 (0.042)	-0.073 (0.055)	-0.023 (0.020)	-0.021 (0.024)	-5.3 (4.0)	
Depression/anxiety	<b>-0.071 (0.019)</b>	<b>-0.076 (0.026)</b>	<b>-0.085 (0.021)</b>	<b>-0.11 (0.038)</b>	<b>-0.072 (0.015)</b>	<b>-0.051 (0.015)</b>	-2.2 (2.3)	
GI ulcer	-0.0035 (0.012)	0.0071 (0.019)	-0.025 (0.017)	-0.0043 (0.025)	0.0056 (0.014)	-0.0024 (0.014)	-0.46 (1.9)	
Thyroid disorder currently taking medications	-0.012 (0.014)	-0.018 (0.024)	0.012 (0.016)	0.0058 (0.028)	0.018 (0.012)	-0.0024 (0.014)	0.63 (1.9)	
Chronic back pain from disk herniation	<b>-0.090 (0.019)</b>	-0.050 (0.028)	<b>-0.068 (0.025)</b>	<b>-0.11 (0.039)</b>	<b>-0.063 (0.017)</b>	<b>-0.067 (0.018)</b>	-4.6 (2.4)	
R <sup>2</sup>	0.29	0.36	0.27	0.25	0.34	0.28	0.24	
Average of the estimates for the 16 health conditions	-0.033	-0.070	-0.038	-0.056	-0.036	-0.034	-0.040	
Average of the estimates for the 16 health conditions divided by the average EQ-5D estimate of -0.033	1.00	2.12	1.15	1.70	1.09	1.03	1.21	

<sup>a</sup>Estimates in bold have  $P < 0.05$ . Each column contains the regression model results from the instrument which heads that column.

<sup>b</sup>EQ-5D = EuroQol-5D-3L

<sup>c</sup>HALex = Health and Activities Limitations Index



HUI = Health Utilities Index  
QWB-SA = Quality of Well-being Scale  
VAS = visual analog scale  
CHD = coronary heart disease  
GI = gastrointestinal

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4**

Estimated Change in Quality-Adjusted Life Years for Each Health Condition in Each Health Utility Measure During 10 Years with a 3% Discount Rate

	Age, y, 35–69						
	EQ-5D <sup>a</sup>	HALex <sup>b</sup>	HUI2 <sup>c</sup>	HUI3	SF-6D	QWB-SA <sup>d</sup>	VAS <sup>e</sup>
<b>CHD<sup>f</sup> with chest pain medications</b>	-0.46	-1.58	-0.60	-0.88	-0.38	-0.38	-0.68
<b>CHD without chest pain medications</b>	-0.23	-1.05	-0.48	-0.46	-0.35	-0.54	-0.57
<b>Stroke</b>	-0.25	-0.24	-0.40	-0.85	-0.15	-0.82	-1.05
<b>Diabetes and using insulin</b>	-0.22	-1.40	-0.50	-0.52	-0.32	-0.26	-0.88
<b>Diabetes and using medications which are not insulin</b>	0.00	-0.88	-0.55	-0.61	-0.26	-0.34	-0.46
<b>Diabetes without medication</b>	-0.32	-0.68	-0.54	-0.77	-0.33	-0.39	-0.62
<b>Arthritis</b>	-0.67	-0.46	-0.60	-0.76	-0.39	-0.50	-0.46
<b>Eye disease</b>	-0.18	-0.31	-0.22	-0.39	-0.10	-0.41	-0.07
<b>Sleep disorder</b>	-0.84	-0.86	-0.98	-1.14	-0.47	-0.47	-0.43
<b>Asthma</b>	-0.17	-0.53	0.02	-0.32	-0.21	-0.40	-0.09
<b>Emphysema</b>	-0.64	-0.88	-0.89	-1.05	-0.45	-0.23	-1.40
<b>Chronic bronchitis</b>	-0.55	-0.59	-0.57	-0.83	-0.31	-0.44	-0.80
<b>Depression/anxiety</b>	-0.87	-0.83	-1.17	-1.31	-0.73	-0.96	-0.82
<b>GI<sup>g</sup> ulcer</b>	-0.27	-0.50	-0.39	-0.42	-0.39	-0.33	-0.39
<b>Thyroid disorder currently taking medications</b>	-0.06	-0.14	0.01	-0.18	-0.06	-0.01	0.03
<b>Chronic back pain from disk herniation</b>	-0.81	-1.05	-1.07	-1.31	-0.70	-0.47	-0.60

  

	Age, y, 70–89						
	EQ-5D	HALex	HUI2	HUI3	SF-6D	QWB	VAS
<b>CHD with chest pain medications</b>	-0.43	-1.23	-0.98	-1.31	-0.62	-0.62	-1.05
<b>CHD without chest pain medications</b>	-0.07	-0.22	-0.03	0.07	-0.09	-0.26	-0.21
<b>Stroke</b>	-0.45	-1.31	-0.67	-1.14	-0.45	-0.37	-0.62
<b>Diabetes and using insulin</b>	-0.41	-1.14	-0.15	-0.71	-0.39	-0.09	-0.36
<b>Diabetes and using medications which are not insulin</b>	-0.26	-0.81	-0.03	0.03	-0.29	-0.18	-0.52
<b>Diabetes without medication</b>	-0.08	0.04	-0.28	-0.27	0.11	-0.07	0.33
<b>Arthritis</b>	-0.61	-0.79	-0.62	-0.85	-0.65	-0.61	-0.60

	<b>Age, y, 70–89</b>						
	<b>EQ-5D</b>	<b>HALex</b>	<b>HUI2</b>	<b>HUI3</b>	<b>SF-6D</b>	<b>QWB</b>	<b>VAS</b>
<b>Eye disease</b>	−0.04	0.03	−0.18	−0.18	−0.11	−0.22	−0.11
<b>Sleep disorder</b>	−0.68	−0.96	−0.61	−0.59	−0.63	−0.76	−0.66
<b>Asthma</b>	−0.03	−0.45	−0.23	−0.19	−0.48	−0.10	−0.36
<b>Emphysema</b>	−0.06	−1.05	−0.24	−0.20	−0.24	−0.18	−0.41
<b>Chronic bronchitis</b>	0.12	−0.67	−0.33	−0.64	−0.20	−0.18	−0.46
<b>Depression/anxiety</b>	−0.62	−0.67	−0.83	−0.96	−0.63	−0.45	−0.19
<b>GI ulcer</b>	−0.03	0.06	−0.24	−0.04	0.05	−0.02	−0.04
<b>Thyroid disorder currently taking medications</b>	−0.11	−0.16	0.12	0.05	0.17	−0.02	0.06
<b>Chronic back pain from disk herniation</b>	−0.79	−0.44	−0.66	−0.96	−0.55	−0.59	−0.40

<sup>a</sup>EQ-5D = EuroQol-5D-3L

<sup>b</sup>HALex = Health and Activities Limitations Index

<sup>c</sup>HUI = Health Utilities Index

<sup>d</sup>QWB-SA = Quality of Well-being Scale

<sup>e</sup>VAS = visual analog scale

<sup>f</sup>CHD = coronary heart disease

<sup>g</sup>GI = gastrointestinal

Spearman Correlation of Health Condition Impacts as Measured by Each of the Health Utility Measures

**Table 5**

Age, y, 35–69						
	EQ-5D <sup>a</sup>	HALex <sup>b</sup>	HUI2 <sup>c</sup>	HUI3	SF-6D	QWB-SA <sup>d</sup>
HALex	0.23					
HUI2	0.84	0.53				
HUI3	0.84	0.44	0.92			
SF-6D	0.86	0.74	0.83	0.74		
QWB-SA	0.48	-0.08	0.33	0.43	0.29	
VAS <sup>e</sup>	0.42	0.44	0.51	0.66	0.33	0.16

  

Age 70–89						
	EQ-5D	HALex	HUI2	HUI3	SF-6D	QWB-SA
HALex	0.42					
HUI2	0.51	0.37				
HUI3	0.59	0.59	0.89			
SF-6D	0.69	0.60	0.67	0.70		
QWB-SA	0.66	0.47	0.69	0.61	0.79	
VAS	0.42	0.84	0.48	0.54	0.70	0.71

<sup>a</sup>EQ-5D = EuroQol-5D-3L

<sup>b</sup>HALex = Health and Activities Limitations Index

<sup>c</sup>HUI = Health Utilities Index

<sup>d</sup>QWB-SA = Quality of Well-being Scale

<sup>e</sup>VAS = visual analog scale