HHS Public Access

Author manuscript

Pharmacoeconomics. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Pharmacoeconomics. 2016 October ; 34(10): 1015–1022. doi:10.1007/s40273-016-0408-x.

Using Linear Equating to Map PROMIS Global Health Items and the PROMIS-29 V. 2 Profile Measure to the Health Utilities Index —Mark 3

Ron D. Hays[1], **Dennis A. Revicki**[2], **David Feeny**[3], **Peter Fayers**[4], **Karen L. Spritzer**[1], and **David Cella**[5]

[1]Division of General Internal Medicine, Department of Medicine, UCLA, Los Angeles, CA

[2]Outcomes Research, Evidera, Bethesda, MD

[3]Department of Economics, McMaster University, Hamilton, Ontario, Canada; Health Utilities Incorporated, Dundas, Ontario, Canada

[4]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, U.K; Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway

[5]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, Illinois

Abstract

Background—Preference-based health-related quality of life scores are useful as outcome measures in clinical studies, for monitoring health of populations, and for estimating quality-adjusted life years.

Methods—This was a secondary analysis of data collected in an internet survey as part of the Patient-Reported Outcomes Measurement Information System (PROMIS®) project. We used the 10 PROMIS global health items, the PROMIS-29 V. 2 single pain intensity item and 7 multi-item scales (physical functioning, fatigue, pain interference, depressive symptoms, anxiety, ability to participate in social roles and activities, sleep disturbance), and the PROMIS-29 V. 2 items to estimate HUI-3 preference scores. Linear regression analyses were used to identify significant predictors, followed by simple linear equating to avoid regression to the mean.

Results—The regression models explained 48% (global health items), 61% (PROMIS-29 V. 2 scales) and 64% (PROMIS-29 V. 2 items) of the variance in the HUI-3 preference score. Linear equated scores were similar to observed scores, although differences tended to be larger for older study participants.

Conclusions—HUI-3 preference scores can be estimated from the PROMIS global health items or the PROMIS-29 V. 2. The estimated HUI-3 scores from the PROMIS health measures can be

Corresponding Author: Ron D. Hays, Ph.D., UCLA, 911 Broxton Avenue, Los Angeles, CA 90024, drhays@g.ucla.edu; Phone: 310-794-2294; Fax: 310-794-0732.

Author Contributions. Hays drafted the article and supervised analyses of the data. All other authors provided edits to the draft article. Feeny and Fayers provided input on the statistical analyses. Spritzer implemented the analyses.

used for economic applications and as a measure of overall health-related quality of life in research.

Keywords

HUI-3; PROMIS®; mapping; linear equating

1. Introduction

Health-related quality of life (HRQOL) measures are often used to examine the effects of medical interventions. Generic HRQOL profile measures provide multiple health domains scores, but not an overall index score [1–4]. Preference-based measures provide a single summary score assessing overall HRQOL and are useful as an outcome measure [5], for monitoring the health of populations [6], and for estimating quality-adjusted life years for economic evaluations [7]. They provide information on the value of different health states and can be used to estimate health outcomes for cost-effectiveness analyses.

Preference-based measures include the EuroQoL EQ-5D-3L [8], the Quality of Well-Being Scale [9], the SF-6D [10], and the Health Utilities Index Mark 3 (HUI-3) [11]. Although each of these health indexes provides valuations on a 0 (dead) to 1 (perfect health/best imaginable health) scale (3 of the 4 indexes include health states rated less than 0), they differ in health state classification systems, methods for preference assessment, and scoring algorithms. U.S. normative data for these measures was reported in the National Health Measurement Study [12]. The different health indexes vary in their precision along the range of the underlying health status concept, but they are all related [13].

The National Institutes of Health (NIH) launched the Patient-Reported Outcomes Measurement Information System (PROMIS®) in 2004 with the goal of developing, evaluating, and disseminating publicly available item banks assessing HRQOL (www.nihpromis.gov). The PROMIS project developed global health items and profile measures to assess multiple HRQOL domains that are now widely used in the U.S. These measures are designed to be administered efficiently and provide a common-language across conditions. The PROMIS-29 and global health items have been mapped to the EQ-5D-3L [14], but not to other widely used preference-based measures. Estimated health preference scores from the PROMIS measures is useful when preferences for health states have not been assessed in a study.

The HUI-3 has 8 attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain and discomfort [11]. Three of the attributes have 5 levels (speech, emotion, pain) and the other five have 6 levels (vision, hearing, ambulation, dexterity, cognition). The objective of this study was to estimate HUI-3 scores from the PROMIS global health items and PROMIS-29 V.2 profile measure. We also compare the estimated health preference scores to HUI-3 index scores by age and gender groups. We followed recommendations for the reporting of mapping studies [15,16]. Our completed Mapping onto Preference-based measures reporting Standards (MAPS) checklist is available upon request.

2. Methods

2.1 Measures

The HUI-3 yields a preference score based on a multi-attribute utility function derived using visual analog scale and standard gamble elicited preferences from a general population sample in Hamilton (Ontario, Canada). PROMIS has 10 global health questions or items [17] including the widely used *excellent* to *poor* general health rating question [18]. The remaining 9 global health items assess physical health (two items), pain, fatigue, general mental health, emotional distress, overall quality of life, and social function (two items). The PROMIS-29 V. 2 profile measure assesses pain intensity using a single 0–10 numeric rating scale item and 7 health domains using 4 items each: physical functioning, fatigue, pain interference, depressive symptoms, anxiety, ability to participate in social roles and activities, and sleep disturbance. The PROMIS items and scales (see Appendix) in this study are conceptually similar to the HUI-3 attributes but do not include direct measures of cognition and sensation (vision, hearing, and speech).

Study participants completed the 10 PROMIS global health items, PROMIS-29 profile measure, HUI-3, and demographic questions on the web. They received nominal incentives from Op4G for completing the survey. The specific nature and value of the incentive varies, but did not exceed 10 U.S. dollars.

2.2 Sample

We analyzed data collected from members of the Op4G internet panel: https://op4g.com/our-panel/. Op4G maintains a U.S. national sample, and participants are required to update demographic information regularly. We specified quotas (fulfilled by Op4G) for region (18% Northeast, 20% Midwest, 37% South, 33% West), race/ethnicity (500 Hispanics, 500 African Americans, and 200 Asians), and education (14% less than high school, 31% high school degree, 28% some college, 27% college degree). Quotas were also set for 24 age-gender subgroups.

2.3 Analysis Plan

We estimated Spearman correlations between HUI-3 attribute levels with corresponding PROMIS domain scores. We estimated ordinary least squares regression equations predicting the HUI-3 preference scores from the PROMIS global health items, PROMIS-29 V.2 domain scores, and PROMIS-29 V. 2 items. First, we regressed the HUI-3 preference scores on the PROMIS global health items, retaining items that were statistically significantly ($p < 0.05$) associated with HUI-3 preference scores. The global health items were scored as equal interval with a higher score representing better health. The 0–10 global pain item was recoded in accordance with PROMIS convention [4, 17] into 5 categories based on grouping of the 0–10 response scales for the Sheehan Disability Scale and the Flushing Questionnaire: 10 = 1 (worst pain), 7–9 = 2, 4–6 = 3, 1–3 = 4, and 0 = 5 (no pain).

We regressed the HUI-3 on the PROMIS-29 V.2 scales. We scored these following the PROMIS convention that larger scale scores correspond to more of the concept depicted in the name. Thus, higher scores sometimes represent better health and sometimes worse

health, depending on the name of the scale. For physical functioning and ability to participate in social roles and activities a higher score indicates better health while a higher score on anxiety, depression, fatigue, sleep disturbance, pain interference, and pain intensity indicates worse health. We recoded the 0–10 global pain item into 5 categories: 0 =1 (no pain), 1–3 = 2, 4–6 = 3, 7–9 = 4, 10 =5 (worst pain). Next, we regressed the HUI-3 preference scores on PROMIS-29 V. 2 items, with higher scores corresponding to the item name (11 items were coded so that a higher score is worse health and 6 items were coded so that a higher score is better health). For this model, we used forward stepwise regression to identify items with statistically significant ($p < 0.05$) unique associations with the HUI-3 preference scores.

Regression-based prediction results in biased estimates due to regression to the mean. Linear equating reduces the typical problem of over prediction of low scores and under prediction of high scores [19]. Because our objective was to map PROMIS scores to the equivalent HUI-3 preference-based scores, we transformed predicted scores from each of the three regression models linearly to have the same mean and standard deviation as the observed HUI-3 preference-based scores (i.e., linear equating). Then, we recoded mapped (equivalent) scores that were outside of the observed −0.359 to 1.000 range to the nearest minimum or maximum observed scores [19].

To obtain an estimate of capitalization on chance in our regression models, we split the sample into two random halves and derived regression equations on the first random half and applied those equations to the second random half sample. We estimated product-moment and intraclass correlations between predicted and observed HUI-3 preference scores in the first random half sample and compared them to the correlations of observed HUI-3 preference scores with predicted scores in the second half.

We compared estimated HUI-3 preference scores with observed scores overall and by age and gender subgroups. In addition, we estimated HUI-3 preference scores in the original PROMIS Wave 1 data collected in 2007–2008 [20] by using used the regression equation based on the PROMIS-29 v. 2 scales and added a constant (product of the difference in the U.S. general population and the current study's HUI-3 preference score means and the ratio of their SDs).

3. Results

The sample consisted of 3,000 individuals: 51% female; 17% Hispanic, 60% non-Hispanic white, 14% non-Hispanic black, and 9% Asian; 14% less than a high school education, 31% high school graduates, and 55% education beyond high school. Age was distributed as 30% 18–34, 18% 35–44, 19% 45–54, 16% 55–64, 9% 65–74, and 8% 75–88. Fifty-six percent of the sample were married or living with a partner. The demographic characteristics of the sample was similar to that of the U.S. general population, but respondents reported worse health by about a half-standard deviation on PROMIS domains compared to the PROMIS wave 1 general population sample, which is comparable to the 2000 U.S. Census [21]. Thirty-four percent of the sample reported having been told by a doctor that they have high blood pressure, 20% arthritis or rheumatism, 17% asthma, 16% migraines, 11% diabetes,

10% angina, 5% heart attack, 5% cancer (other than non-melanoma skin cancer), 5% chronic lung disease, 4% congestive heart failure, 3% liver disease, and 3% kidney disease. Moreover, the relatively poor health of the sample was indicated by an average HUI-3 preference score of 0.544 (SD = 0.400) compared to a U. S. mean of 0.87 (SD = 0.21) in the Joint Canada/United States survey of health [22–23].

Spearman correlations between the PROMIS physical functioning scale with the HUI-3 ambulation and dexterity attributes were 0.70 and 0.55, respectively. The PROMIS depressive symptoms scale (4 items) correlated −0.62 with HUI-3 emotion. The PROMIS pain interference scale (4 items) correlated −0.68 with the HUI-3 pain attribute.

Item missing rates were less than 0.2%; sample sizes for multivariate analyses reported below were 2,994 or larger.

3.1 Global Health Items

Six of the global health items were significantly associated and accounted for 48% of the variance (adjusted R^2) in the HUI-3 preference score (Table 1). The strongest unique associations (standardized beta) with the HUI-3 preference scores were observed for the physical functioning and the pain rating items. The resulting equated HUI-3 preference scores had a mean of 0.530 and SD of 0.377 compared to the observed HUI-3 preference score mean of 0.544 and SD of 0.400. The product-moment correlation of the equated with observed HUI-3 preference scores was 0.70 ($n=2994$, $p<0.0001$); the intraclass correlation between equated and observed scores was also 0.70.

3.2 PROMIS-29 V.2 Scales

Six of the PROMIS-29 V.2 scales were significantly associated and accounted for 61% of the variance in the HUI-3 preference scores (Table 2). Because of the suppression effects for the global pain rating item (i.e., zero-order correlation was negative but regression coefficient was positive), we reran the regression model, dropping the item (Table 3). The variance explained by the model did not change (i.e., was 61%). The strongest unique associations with the HUI-3 preference scores were observed for the physical functioning and depressive symptoms scales.

The equated HUI-3 preference scores had a mean of 0.524 and SD of 0.371 compared to the observed HUI-3 preference score mean of 0.544 and SD of 0.400. The equated HUI-3 preference scores correlated (product-moment) 0.78 ($n=2996$, $p<0.0001$) with the observed HUI-3 preference scores; the intraclass correlation between equated and observed HUI-3 preference scores was also 0.78.

3.3 PROMIS-29 V.2 Items

The regression model for the PROMIS-29 V.2 items showed that 17 items had significant unique associations and accounted for 64% of the variance in the HUI-3 preference scores (Table 4). Among the 17 items, 2 displayed suppression effects (*sleep quality* and *feel fatigued*). The four strongest unique associations with the HUI-3 preference scores were found for three physical functioning items (do chores such as vacuuming or yard work, run

errands and shop, walk at least 15 minutes) and one depressive symptoms item (felt hopeless).

The equated HUI-3 preference scores had a mean of 0.542 and SD of 0.391 compared to the observed HUI-3 preference score mean of 0.544 and SD of 0.400. The equated HUI-3 preference scores correlated (product-moment) 0.80 (n = 2994, p <0.0001) with the observed HUI-3 preference scores; the intraclass correlation between equated and observed scores was also 0.80.

3.4 Cross Validation of Regression Equations

The product-moment and intraclass correlations between estimated and observed HUI-3 preference scores from a regression equation of the global health items in the first random half were 0.72 and 0.68 (n = 1513), respectively, compared to 0.67 and 0.63 (n = 1481) when applying the equation to the second random half sample. The product-moment and intraclass correlations between estimated and observed HUI-3 preference scores from a regression equation of the PROMIS-29 scales in the first random half were 0.79 and 0.77 (n = 1515), respectively, compared to 0.76 and 0.74 (n = 1481) when applying the equation to the second random half sample.

3.5 Estimated Versus Observed HUI-3 PreferenceScores by Age and Gender

The correspondence between observed HUI-3 and equated preference scores overall and by age and gender groups is summarized in Table 5. Average equated scores were within 0.02 of observed scores for the overall sample; less than the 0.03 difference in scores that is regarded as minimally important [24]. Equated scores tended to be more discrepant from observed scores for the oldest study participants. For example, equated scores based on the PROMIS-29 V.2 scales were 0.13 higher than observed scores for males 75–88 years old (0.25 vs. 0.12). The general pattern of equated HUI-3 preference scores showed a decline by age, but those aged 55–74 (55–64 and 65–74 age subgroups) tended to have higher scores than other age groups.

3.6 Estimated HUI-3 Preference Scores in General Population from PROMIS-29 V.2 Scales

The estimated HUI-3 preference scores in the PROMIS Wave 1 sample using the PROMIS-29 V. 2 scales are similar to U.S. general population norms reported for males by Fryback et al. [12], but HUI-3 preference estimates derived from the current study were higher (more positive) for females (Table 6).

4. Discussion

The PROMIS measures were rigorously developed and allow flexibility in administration using either targeted short forms or computerized adaptive testing [20]. The availability of HUI-3 preference scores based on the PROMIS global items and PROMIS-29 V.2 profile measure enables potential application of these measures to population-based studies and economic evaluations.

The regression models estimated here accounted for between 48% and 64% of the variance in the HUI-3 preference scores. The best prediction was obtained for the PROMIS-29 V. 2 items, followed closely by the PROMIS-29 V. 2 scale scores, and then the PROMIS global health items. In comparison, PROMIS wave 1 scale scores and global health items accounted for 57% and 65%, respectively, of the variance in the EQ-5D-3L [14]. The equated HUI-3 preference scores based on PROMIS measures were comparable to those directly assessed using the HUI-3 in this sample. Intraclass correlations were *good* according to the *poor*, *fair*, *moderate*, *good* or *very good* categorization suggested by Altman [25]. The largest differences between average equated and observed scores were found for older individuals, especially 75–88 year-old males. The higher mean equated scores for those 55–74 years old is consistent with the observed HUI-3 preference scores reported in Fryback et al. [12].

We recommend use of the PROMIS-29 V.2 scales to estimate HUI-3 preference scores in cases where only one approach is desired, because the variance explained was similar to that of the best regression prediction equation (PROMIS-29 V. 2 items) and the PROMIS-29 V. 2 scales allow for greater flexibility in choice of items in a study. That is, the HUI-3 can be estimated from any subset of PROMIS items that yield an estimate of physical functioning, depressive symptoms, pain interference, ability to participate in social roles and activities, and anxiety scale scores. IRT scores for the PROMIS V. 2 scales can be estimated within Assessment Center (www.assessmentcenter.net). Predicted HUI-3 preference scores can be obtained using: 0.42094 + (0.01704 × Physical functioning) + (−0.00793 × Depressive symptoms) + (−0.00505 × Pain interference) + (0.00451 × Ability to participate in social roles and activities) + (−0.00313 × Anxiety). These predicted scores can then be adjusted to the U.S. general population by adding 0.17103– the product of the difference in the U.S. general population and the current study's HUI-3 preference score means and the ratio of their SDs. Any scores below −0.359 should be recoded to −0.359, and scores greater than 1.000 recoded to 1.000.

There are several limitations associated with these analyses. First, the participants in this study were from an internet panel and had worse average HRQOL than U.S. national probability-based samples, indicating that the sample is not representative of the U.S. general population [26]. However, the sample included a wide range of HUI-3 preference scores and is therefore useful for equating PROMIS scores to the HUI-3 preference score. Second, the analyses are based on only a single dataset and variance explained in a derivation random half subsample was inflated by 5–7% compared to a cross-validated random half subsample. Third, the PROMIS and HUI items were self-administered by web-based methods and responses could differ for other modes of administration [27]. But a comparison of responses to PROMIS items administered by different modes (interactive voice response, paper questionnaire, personal digital assistant, or personal computer) showed method equivalence [28]. Fourth, the PROMIS measures were collected in the U.S. but the HUI-3 scoring function was derived from a representative sample of Canadians. However, estimated scoring functions for the HUI-3 are very similar between Canada [11], the Netherlands [29], France [30], and Spain [31]. Finally, it is preferable to include the HUI-3 itself or assess preferences directly (i.e., time trade-off, standard gamble) rather than estimate the HUI-3 preference scores However, when either of these is not possible, the

estimates provided here can provide a second-best approach. A previous study used discrete choice experiments to derive preferences for health states derived from the PROMIS-29 V.1, but the estimates produced were implausible—the mean was 0.16 in a sample drawn from the U.S. general population [32]. Potentially better alternative methods for directing eliciting preferences in PROMIS have been proposed [33].

5. Conclusion

We estimated HUI-3 preference scores accurately from PROMIS global health items and the PROMIS-29 V.2 scales, and these mapped preference scores varied as expected by demographic characteristics in the PROMIS sample. Additional research is needed to further evaluate the validity of the estimated index scores. In addition, studies are needed to examine other possible approaches to deriving preference-based scores from the PROMIS measures. These mapped HUI-3 preference scores have applications in measuring the health of populations and estimating quality-adjusted life years for economic evaluations. We recommend that these estimated HUI-3 preference scores be used only for group-level (not individual level) applications. Given the flexibility of multi-domain short forms and computerized adaptive testing, the PROMIS domain item banks and domain scores may be very useful in clinical studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from National Cancer Institute (1U2-CCA186878-01) and a supplement to the PROMIS statistical center grant (3U54AR057951-04S4). Ron D. Hays, Dennis A. Revicki, Peter Fayers, Karen L. Spritzer and David Cella declare no conflicts of interest. David Feeny has a proprietary interest in Health Utilities Incorporated, Dundas, Ontario, Canada.

References

1. Hays, RD.; Alonso, J.; Coons, SJ. Possibilities for summarizing health-related quality of life when using a profile instrument. In: Staquet, M.; Hays, RD.; Fayers, P., editors. Quality of Life Assessment in Clinical Trials: Methods and Practice. Oxford: Oxford University Press; 1998.
2. Kaplan, RM. Measuring quality of life for policy analysis. In: Lenderking, WR.; Revicki, DA., editors. Advancing health outcome research methods and clinical applications. McLean, VA: International Society for Quality of Life Research; 2005.
3. Kaplan RM, Feeny D, Revicki DA. Methods for assessing relative importance in preference-based outcome measures. Qual Life Res. 1993; 2:467–475. [PubMed: 8161981]
4. Revicki DA, Kaplan RM. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. Qual Life Res. 1993; 2:477–487. [PubMed: 8161982]
5. Hays RD, Eastwood J, Kotlerman J, Spritzer KL, et al. Health-related quality of life and patient reports about care outcomes in a multidisciplinary hospital intervention. Ann Behav Med. 2006; 31:173–178. [PubMed: 16542132]
6. Hays RD, Reeve BB, Smith AW, Clauser SB. Associations of cancer and other chronic medical conditions with SF-6D preference-based scores in Medicare beneficiaries. Qual Life Res. 2014; 23:385–391. [PubMed: 23990395]
7. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Pharmacoeconomics. 1995; 7:490–502. [PubMed: 10155335]

8. Brooks, R. The EuroQol group after 25 years. Springer; New York: NY: 2013.
9. Hector RD, Anderson JP, Paul RC, Weiss RE, et al. Health state preferences are equivalent in the United States and Trinidad and Tobago. Qual Life Res. 2010; 19:729–738. [PubMed: 20237958]
10. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002; 21:271–292. [PubMed: 11939242]
11. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the Health Utilities Index Mark 3 System. Med Care. 2002; 40:113–128. [PubMed: 11802084]
12. Fryback DG, Dunham NC, Palta M, Hanmer J, et al. U.S. Norms for Six Generic Health-Related Quality-of-Life Indexes from the National Health Measurement Study. Med Care. 2007; 45:1162–1170. [PubMed: 18007166]
13. Fryback DG, Palta M, Cherepanov D, Bolt D, et al. Comparison of five health-related quality-of-life indexes using item response theory. Med Dec Making. 2010; 30:5–15.
14. Revicki DA, Kawata AK, Harnam N, Chen WH, et al. Predicting EuroQol (EQ-5D) scores from the Patient-Reported Outcomes Measurement Information System (PROMIS) Global items and domain item banks in a United States sample. Qual Life Res. 2009; 18:783–791. [PubMed: 19472072]
15. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluations in health care. Annu Rev Public Health. 2000; 21:587–611. [PubMed: 10884966]
16. Petrou S, Rivero-Aria O, Dakin H, Longworth L, et al. Preferred reporting items for studies mapping onto preference-based outcome measures: The MAPS statement. PharmacoEconomics. 2015; 33:985–991. [PubMed: 26232201]
17. Hays RD, Bjorner J, Revicki DA, Spritzer KL, et al. Development of physical and mental health summary scores from the Patient-Reported Outcomes Measurement Information System (PROMIS) global items. Qual Life Res. 2009; 18:873–80. [PubMed: 19543809]
18. Hays RD, Spritzer KL, Thompson WW, Cella D. U.S. general population estimate for “excellent” to “poor” self-rated health item. J Gen Intern Med. 2015; 30:1511–6. [PubMed: 25832617]
19. Fayers PM, Hays RD. Should linking replace regression when mapping from profile to preference-based measures? Value in Health. 2014; 17:261–265. [PubMed: 24636385]
20. Cella D, Riley W, Stone A, Rothrock N, et al. Initial item banks and first wave testing of the Patient-Reported Outcomes Measurement Information System (PROMIS) network: 2005–2008. J Clin Epidemiol. 2010; 63:1179–1194. [PubMed: 20685078]
21. Liu H, Cella D, Gershon R, Shen J, Morales Riley W, Hays RD. Representativeness of the Patient-Reported Outcomes Measurement Information System Internet panel. J Clinical Epidemiology. 2010; 63(11):1169–1178.
22. Sanmartin C, Berthelot JM, Ng E, et al. Comparing health and health care in Canada and the United States. Health Faa. 2006; 25:1133–42.
23. Feeny D, Kaplan MS, Huguet N, McFarland BH. Comparing population health in the United States and Canada. Population Health Metrics. 2010; 8:8. [PubMed: 20429875]
24. Feeny D, Spritzer KL, Hays RD, Liu H, et al. Agreement about identifying patients who change over time: Cautionary results in cataract and heart failure patients. Med Decis Making. 2011; 32:273–286. [PubMed: 22009666]
25. Altman, DG. Practical statistics for medical research. London: Chapman & Hall; 1991.
26. Hays RD, Liu H, Kapteyn A. Use of internet panels to conduct surveys. Behav Res Methods. 2015; 47:685–90. [PubMed: 26170052]
27. Hanmer J, Hays RD, Fryback DG. Mode of administration is important in U.S. national estimates of health-related quality of life. Med Care. 2007; 45:1171–1179. [PubMed: 18007167]
28. Bjorner JB, Rose M, Gandek B, Stone AA, et al. Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. J Clin Epidemiol. 2014; 67:108–113. [PubMed: 24262772]
29. Raat H, Bonsel GJ, Hoogeveen WC, Essink-Bot ML. Feasibility and reliability of a mailed questionnaire to obtain visual analogue scale valuations for health states defined by the Health Utilities Index Mark 3. Medical Care. 2004; 42(1):13–18. [PubMed: 14713735]

30. Le Galès C, Buron C, Costet N, Rosman S, Slama PRG. Development of a preference-weighted health status classification system in France: The Health Utilities Index 3. Health Care Manag Sci. 2002; 5(1):41–51. [PubMed: 11862978]

31. Ruiz M, Rejas J, Soto J, Pardo A, Rebollo I. Adaptation and validation of the Health Utilities Index Mark 3 into Spanish and correction norms for Spanish population. Medicina Clinica. 2003; 120:89–96. [PubMed: 12605729]

32. Craig BM, Reeve BB, Brown PM, Cella D, et al. US valuation of health outcomes measured using the PROMIS-29. Value Health. 2014; 17:846–853. [PubMed: 25498780]

33. Hanmer J, Feeny D, Fischoff B, Hays RD, et al. The PROMIS of QALYs. Health and Quality of Life Outcomes. 2015; 3:122. [PubMed: 26260491]

Key Points

1. The HUI-3 preference score is estimated from the PROMIS-29 V. 2 scales;
2. The estimated HUI-3 preference scores can be used for economic applications; and
3. Future research is needed to derive preference scores directly from the PROMIS measures.

Table 1

Regression of HUI-3 Preference Scores on Global Health Items

Variable	Standardized beta	Unstandardized beta	Standard Error	t	p
Physical function (Global06)	0.40817	0.13657	0.00561	24.35	<.0001
Pain rating (Global07)	0.14099	0.05669	0.00610	9.29	<.0001
Mental health (Global04)	0.10539	0.03773	0.00705	5.35	<.0001
Fatigue (Global08)	0.10276	0.04250	0.00705	6.03	<.0001
Bothered by emotional problems (Global10)	0.09162	0.03271	0.00591	5.54	<.0001
Social (Global05)	0.06335	0.02227	0.00672	3.31	0.0009

Intercept = −0.61341; Adjusted R^2 = 0.4824, n = 2994; All items were scored so that higher score is better health. Global07 administered as 0–10 but recoded: 10 = 1, 7–9 = 2, 4–6 = 3, 1–3 = 4, 0 = 5.

Pharmacoeconomics. Author manuscript; available in PMC 2017 October 01.

Table 2

Regression of HUI-3 Preference Scores on PROMIS-29 V. 2 Scales

Variable	Standardized beta	Unstandardized beta	Standard Error	t	p
Physical functioning	0.40007	0.01737	0.00080	21.62	<.0001
Depressive symptoms	−0.20530	−0.00787	0.00079	−9.92	<.0001
Pain interference	−0.15187	−0.00637	0.00088	−7.21	<.0001
Ability to participate in social roles and activities	0.11359	0.00455	0.00080	5.67	<.0001
Anxiety	−0.08622	−0.00337	0.00081	−4.15	<.0001
Pain rating (Global07)*	0.06239	0.02509	0.00610	4.12	<.0001

Intercept = 0.42086; Adjusted R^2 = 0.6091, n = 2995. Two of the PROMIS-29 scales above were scored so that a higher score is better health (physical functioning, ability to participate in social roles and activities) while the other four scales were scored so that higher score is worse health. Pain rating administered as 0–10 response scale but recoded: 10 = 5, 7–9 = 4, 4–6 = 3, 1–3 = 2, 0 = 1

*Suppression effect (i.e., product-moment correlation of pain rating with HUI-3 = −0.43506 and pain rating with pain interference = 0.64893, n = 2995).

Table 3

Regression of HUI-3 Preference Scores on PROMIS-29 V. 2 Scales, dropping global pain rating item

Variable	Standardized beta	Unstandardized beta	Standard Error	T	p
Physical functioning	0.39256	0.01704	0.00080	21.22	<.0001
Depressive symptoms	−0.20678	−0.00793	0.00080	−9.96	<.0001
Pain interference	−0.12040	−0.00505	0.00082	−6.31	<.0001
Ability to participate in social roles and activities	0.11252	0.00451	0.00081	5.59	<.0001
Anxiety	−0.07990	−0.00313	0.00081	−3.84	<.0001

Intercept = 0.42094; Adjusted R^2 = 0.6062, n = 2996. Two of the PROMIS-29 scales above were scored so that a higher score is better health (physical functioning, ability to participate in social roles and activities) while the other three scales were scored so that higher score is worse health.

Table 4

Regression of HUI-3 Preference Scores on PROMIS-29 V. 2 Items

Variable	Standardized beta	Unstandardized beta	Standard Error	t	p
Do chores such as vacuuming or yard word (PFA11)	0.15566	0.05342	0.00638	8.37	<.0001
Run errands and shop (PFA53)	0.15336	0.05623	0.00701	8.03	<.0001
Felt hopeless (EDDEP41)	−0.14266	−0.04640	0.00637	−7.29	<.0001
Walk at least 15 minutes (PFA23)	0.10416	0.03572	0.00640	5.58	<.0001
Sleep quality (SLEEP109) *	0.07783	0.02811	0.00502	5.60	<.0001
Pain interfered with ability to participate in social activities (PAININ31)	−0.06333	−0.02054	0.00636	−3.23	0.0013
Hard to focus on anything other than my anxiety (EDANX40)	−0.06299	−0.02186	0.00621	−3.52	0.0004
Trouble starting things because tired (AN3)	−0.06265	−0.02083	0.00570	−3.65	0.0003
Trouble doing regular leisure activities with others (SRPPER11_CaPs)	0.06008	0.01948	0.00632	3.08	0.0021
Pain interfered with day-to-day activities (PAININ9)	−0.05667	−0.01893	0.00673	−2.81	0.0050
Up and down stairs at normal pace (PFA21)	0.05317	0.01786	0.00620	2.88	0.0040
Felt helpless (EDDEP06)	−0.05299	−0.01733	0.00650	−2.67	0.0077
Felt fearful (EDANX01)	−0.04968	−0.01741	0.00590	−2.95	0.0032
Trouble doing usual work (SRPPER23_CaPs)	0.04907	0.01587	0.00631	2.52	0.0119
Feel fatigued (HI7) *	0.03495	0.01205	0.00549	2.19	0.0283
Pain rating, recoded to 1–5 (Global07)	0.03254	0.01309	0.00577	2.27	0.0233
Refreshing sleep (SLEEP116)	−0.02850	−0.00913	0.00450	−2.03	0.0426

Intercept = −0.01283; Adjusted R^2 = 0.6406, n = 2994. 11 items were coded so that a higher score is worse health (EDANX01, EDANX40, EDDEP06, EDDEP41, HI7, AN3, PAININ9, PAININ31, SLEEP109, SLEEP116, Global07) and 6 items were coded so that a higher score is better health (PFA11, PFA21, PFA23, PFA53, SRPPER11_CaPs, SRPPER23_CaPs).

* Suppression effect.

Table 5

Observed Versus Equated HUI-3 Preference Scores by Age and Gender (Standard Error)

Global Health	Observed Males	Equated Males	Observed Females	Equated Females	Observed Overall	Equated Overall
18–34 years	0.50 (0.02) n = 465	0.55 (0.02)	0.61 (0.02) n = 453	0.60 (0.02)	0.55 (0.01) n = 918	0.57 (0.01)
35–44 years	0.53 (0.03) n = 261	0.54 (0.02)	0.62 (0.02) n = 263	0.57 (0.02)	0.58 (0.02) n = 524	0.55 (0.02)
45–54 years	0.53 (0.02) n = 280	0.48 (0.02)	0.58 (0.02) n = 294	0.49 (0.02)	0.56 (0.02) n = 574	0.49 (0.02)
55–64 years	0.55 (0.02) n = 228	0.52 (0.02)	0.71 (0.02) n = 240	0.66 (0.02)	0.64 (0.02) n = 468	0.59 (0.02)
65–74 years	0.47 (0.04) n = 128	0.46 (0.04)	0.69 (0.02) n = 148	0.67 (0.03)	0.59 (0.02) n = 276	0.57 (0.02)
75–88 years	0.12 (0.04) (n = 93)	0.18 (0.04)	0.20 (0.04) n = 140	0.25 (0.04)	0.17 (0.03) n = 233	0.23 (0.03)
Overall	0.49 (0.01) n = 1455	0.50 (0.01)	0.59 (0.01) n = 1539	0.56 (0.01)	0.54 (0.01) n = 2994	0.53 (0.01)
PROMIS-29 V. 2 Scales	Observed Males	Equated Males	Observed Females	Equated Females	Observed Overall	Equated Overall
18–34 years	0.50 (0.02) n = 465	0.48 (0.02)	0.61 (0.02) n = 453	0.58 (0.02)	0.55 (0.01) n = 918	0.53 (0.01)
35–44 years	0.53 (0.03) n = 261	0.51 (0.02)	0.62 (0.02) n = 264	0.58 (0.02)	0.58 (0.02) n = 525	0.55 (0.02)
45–54 years	0.53 (0.02) n = 280	0.49 (0.02)	0.58 (0.02) n = 294	0.51 (0.02)	0.56 (0.02) n = 574	0.50 (0.02)
55–64 years	0.55 (0.02) n = 228	0.55 (0.02)	0.71 (0.02) n = 240	0.64 (0.02)	0.64 (0.02) n = 468	0.60 (0.02)
65–74 years	0.47 (0.04) n = 128	0.50 (0.03)	0.69 (0.02) n = 149	0.64 (0.03)	0.59 (0.02) n = 277	0.57 (0.02)
75–88 years	0.12 (0.04) n = 93	0.25 (0.03)	0.20 (0.04) n = 140	0.32 (0.03)	0.17 (0.03) n = 233	0.29 (0.02)
Overall	0.49 (0.01) n = 1455	0.49 (0.01)	0.59 (0.01) n = 1541	0.56 (0.01)	0.54 (0.01) n = 2996	0.52 (0.01)
PROMIS-29 V2 Items	Observed Males	Equated Males	Observed Females	Equated Females	Observed Overall	Equated Overall
18–34 years	0.50 (0.02) n = 465	0.50 (0.02)	0.61 (0.02) n = 453	0.59 (0.02)	0.55 (0.01) n = 918	0.54 (0.01)
35–44 years	0.53 (0.03) n = 261	0.53 (0.02)	0.62 (0.02) n = 263	0.60 (0.02)	0.58 (0.02) n = 524	0.57 (0.02)

Global Health	Observed Males	Equated Males	Observed Females	Equated Females	Observed Overall	Equated Overall
45–54 years	0.53 (0.02) n = 280	0.53 (0.02)	0.58 (0.02) n = 294	0.54 (0.02)	0.56 (0.02) n = 574	0.53 (0.02)
55–64 years	0.55 (0.02) n = 228	0.58 (0.03)	0.71 (0.02) n = 240	0.68 (0.02)	0.64 (0.02) n = 468	0.63 (0.02)
65–74 years	0.47 (0.04) n = 128	0.52 (0.04)	0.69 (0.02) n = 149	0.68 (0.03)	0.59 (0.02) n = 277	0.60 (0.02)
75–88 years	0.12 (0.04) n = 93	0.23 (0.04)	0.20 (0.04) n = 139	0.29 (0.04)	0.17 (0.03) n = 232	0.27 (0.03)
Overall	0.49 (0.01) n = 1455	0.51 (0.01)	0.59 (0.01) n = 1539	0.58 (0.01)	0.54 (0.01) n = 2994	0.54 (0.01)

Table 6

Estimated HUI-3 Preference Scores in PROMIS Wave-1 General population sample using algorithm derived from PROMIS-29 V.2 scales in Op4G sample (Standard Error)

	Estimated Males	Estimated Females	Estimated Overall
18–34 years	0.92 (0.01) n=446	0.87 (0.01) n=780	0.88 (< 0.01) n=1226
35–44 years	0.89 (0.01) n=377	0.83 (0.01) n=511	0.85 (0.01) n=888
45–54 years	0.88 (0.01) n=340	0.76 (0.01) n=520	0.81 (0.01) n=860
55–64 years	0.85 (0.01) n=313	0.79 (0.01) n=494	0.81 (0.01) n=807
65–74 years	0.88 (0.01) n=227	0.79 (0.01) n=344	0.83 (0.01) n=571
75–100 years	0.86 (0.01) n=404	0.79 (0.01) n=278	0.83 (0.01) n=682
Overall	0.88 (<0.01) n=2107	0.81 (<0.01) n=2927	0.84 (<0.01) n=5034