Cross-sectional validation of the PROMIS-Preference scoring system by its association with social determinants of health

Janel Hanmer[1]

Abstract

Purpose PROMIS-Preference (PROPr) is a generic, societal, preference-based summary score that uses seven domains from the Patient-Reported Outcomes Measurement Information System (PROMIS). This report evaluates construct validity of PROPr by its association with social determinants of health (SDoH).

Methods An online panel survey of the US adult population included PROPr, SDoH, demographics, chronic conditions, and four other scores: the EuroQol-5D-5L (EQ-5D-5L), Health Utilities Index (HUI) Mark 2 and Mark 3, and the Short Form-6D (SF-6D). Each score was regressed on age, gender, health conditions, and a single SDoH. The SDoH coefficient represents the strength of its association to PROPr and was used to assess known-groups validity. Convergent validity was evaluated using Pearson correlations between different summary scores and Spearman correlations between SDoH coefficients from different summary scores.

Results From 4142 participants, all summary scores had statistically significant differences for variables related to education, income, food and financial insecurity, and social interactions. Of the 42 SDoH variables tested, the number of statistically significant variables was 27 for EQ-5D-5L, 17 for HUI Mark 2, 23 for HUI Mark 3, 27 for PROPr, and 27 for SF-6D. The average SDoH coefficients were -0.086 for EQ-5D-5L, -0.039 for HUI Mark 2, -0.063 for HUI Mark 3, -0.064 for PROPr, and -0.037 for SF-6D. Despite the difference in magnitude across the measures, Pearson correlations were 0.60 to 0.76 and Spearman correlations were 0.74 to 0.87.

Conclusions These results provide evidence of construct validity supporting the use of PROPr monitor population health in the general US population.

Keywords Health-related quality of life · Patient-Reported Outcomes Measurement Information System (PROMIS) · PROMIS-Preference (PROPr) · Social determinants of health

Introduction

Quantifying the health of populations is important for designing and evaluating policies and setting research priorities [1, 2]. In the last two decades, there has been an increasing emphasis on using patient-reported outcomes of functioning and well-being, or health-related quality of life

✉ Janel Hanmer
hanmerjz@upmc.edu

1   Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

(HRQoL), to quantify and monitor population health [3]. For example, in the USA, either the SF-36 or the VR-12 has been used in the Medicare Health Outcomes Survey since 1998 [4, 5], and the SF-12 has been used in the Medical Panel Expenditures Survey since 2003 [6]. HRQoL measures selected for measuring population health have usually been disease agnostic (generic), providing an overall description of health not limited to one organ system or disease, thus making them applicable to all survey participants [7]. In addition to being generic, most measures used for population health evaluations have a single composite score to simplify analyses and comparisons. Multi-attribute utility theory provides a framework to develop a single preference-based score from multiple health domains [8, 9].

Because social determinants of health (SDoH) have many known direct and indirect effects on HRQoL, any measure

used for population evaluation should be sensitive to these effects [10–15]. Several frameworks for understanding the relationships between SDoH, medical care, health conditions, and HRQoL are available [16]. For example, the World Health Organization constructed an explanatory and action-oriented conceptual framework of the relationship between SDoH and health and well-being, which includes elements of structural determinants, social determinants, socioeconomic position, individual circumstances, and the health system [17]. Models such as the WHO's emphasize that HRQoL is not just a product of disease and healthcare, but is also directly and indirectly influenced by SDoH.

The most widely used generic preference-based measures in clinical intervention evaluations are the EuroQol-5D (EQ-5D) [18, 19], the Health Utilities Index (HUI) [20, 21], and the SF-6D [22]. These measures have also been shown to be sensitive to SDoH such as income, education, financial security, perceived discrimination, and gender [10–15]. Each of these measures, however, has limitations, including ceiling effects in the general population and an inability to robustly assess individual health domains (such as physical functioning or depression) [23]. The PROMIS-Preference (PROPr) score was built, in part, to address these limitations.

PROPr is based on the Patient-Reported Outcomes Measurement Information System (PROMIS) [24]. Using PROMIS as the foundation for a preference-based score addresses several of the limitations of other scoring systems because the PROMIS domain measures have been developed using item response theory (IRT) [25]. IRT is a modern psychometric measurement technique in a which a pool of items are calibrated over a latent construct to create an "item bank". In HRQoL measurement, the latent constructs include concepts such as depression, pain interference, physical function, and fatigue; PROMIS currently has over 90 adult item banks and over 20 pediatric item banks. All of the responses to all of items within a bank provide information about a respondent's location on the latent construct. This allows a single item bank to be adapted to fit many projects needs while maintaining the score on a latent construct that is comparable to any other project that used the same item bank and calibration. For example, physical function could be measured by a standardized six-item short form for a clinical setting, a custom two-item short form for a population study, and a computer adaptive test in a research study and the physical function scores can be compared. This is true even if each physical function measure used entirely different items from the item bank.

By linking directly to PROMIS domains, PROPr gains many of the advantages of an IRT-based descriptive system including flexible administration of items and the potential to avoid ceiling and floor effects [26]. PROPr also creates the possibility of collecting both health domain information and preference-based scores from the same questions. Moreover,

PROMIS and PROPr are both freely available, whereas the other measures require licensing.

Before any measure can be adopted for widespread use, its validity needs to be demonstrated [7]. PROPr has been shown to have construct validity for chronic conditions with minimal ceiling and floor effects in the general population [27]. This report evaluates the construct validity of PROPr by its association with SDoH in the general US adult population.

Methods

Data

Data used in this study are from a general population sample designed to be representative of US adults age 18 and older. Participants were selected from the AmeriSpeak Panel maintained by NORC at the University of Chicago using sampling strata based on age, race, Hispanic ethnicity, education, and gender [28]. The primary focus of the study was to better understand individuals' HRQoL and its relationship to SDoH. This survey was offered in English and Spanish and was made available on the web and by phone in December 2018. Amerispeak panelists are consented for research when they join the panel and study participants received "Ameripoints" (a cash equivalent point system) from AmeriSpeak for completing surveys. NORC provided demographic information about the sample, latitude and longitude of the participants' household, and the sampling weights necessary to make nationally representative estimates. Data used in these analyses are available online at the Open Science Framework [29].

HRQoL measures

Because of substantial redundancy in questions across the HRQoL measures, participants were randomly assigned to complete either (1) the EQ-5D-5L and SF-12 or (2) the HUI Marks 2 and 3. All participants received PROPr and all SDoH questions.

PROPr

PROPr is based on 7 Adult PROMIS domains: Cognitive Function—Abilities v2.0, Depression v1.0, Fatigue v1.0, Pain Interference v1.0, Physical Function v2.0, Sleep Disturbance v1.0, and Ability to Participate in Social Roles and Activities v2.0 [30]. The PROMIS questions refer to the participant's own health "in the past 7 days" and have five response options. Participants answered 2 questions per domain for a total of 14 questions. Responses were scored by uploading responses to the scoring service on Assessment

Center, and scoring used the default IRT parameters for each item [31]. The PROPr scoring algorithm was developed from standard gamble valuations from a US sample [26]. Possible PROPr scores range from − 0.022 (worst) to 1.0 (best).

To keep sample sizes consistent across measures, PROPr scores were split into two sets (set 1 and set 2) using the same randomization that assigned participants to the other HRQoL measures.

EQ-5D-5L

The EQ-5D-5L questions refer to "your health today". The EQ-5D descriptive system has five domains (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and the – 5 L version has five response options [32] (no problems, slight problems, moderate problems, severe problems, extreme problems/unable to do). This study used the US time trade-off value set [19]. Possible scores range from -0.699 (worst) to 1.0 (best).

Health utilities index (HUI) The original HUI Mark 2 classification system included seven attributes: sensation, mobility, emotion, cognition, self-care, pain, and fertility. Fertility is not currently used in the HUI Mark 2 questionnaire, so current versions have six attributes and the disutility of the fertility attribute is set at 0 [20]. The HUI Mark 3 defines health using eight attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain [21]. The 15 item self-administered HUI questionnaire allows scoring of both Mark 2 and Mark 3. HUI questions refer to "your level of ability or disability during the past week". Scoring algorithms for both HUI Mark 2 and HUI Mark 3 were derived from standard gamble assessments made by adults in community samples in Hamilton, Ontario. HUI Mark 2 scores range from -0.03 (worst) to 1.0 (best) [20]; HUI Mark 3 scores range from − 0.36 (worst) to 1.0 (best) [21].

Short form-6D (SF-6D) from the SF-12 The Short Form-12 (SF-12) version 1 has 12 multiple-choice items that relate to eight dimensions: physical functioning, physical role limitations, emotional role limitations, pain, general health, vitality, social functioning, and mental health [33]. The SF-6D scoring algorithm uses seven of the questions from the SF-12. These seven questions were evaluated using the standard gamble technique in a representative sample of the UK population [22]. Possible SF-6D scores range from 0.345 (worst) to 1.0 (best).

Self-reported SDoH

Participants had several demographic and survey variables provided by the panel company (age, gender, race, Hispanic ethnicity, survey language, survey mode). They were also asked questions about the SDoH listed below. Most of these questions were extracted from large US national surveys such as the National Health Interview Survey [34] or the National Health and Nutrition Examination Survey [35]. Specific sources, questions, responses, and coding are included in the Supplement. All SDoH were binary or categorical outcomes and coded such that the reference category would have the least hardship.

- Health literacy
- Educational attainment
- Has a usual medical provider
- Medical provider CAHPS scores
- Difficulty of getting to medical appointments
- Medical insurance type
- Difficulty paying medical deductibles
- Difficulty paying medical copays
- Difficulty paying monthly bills
- Experienced intimate partner violence in the last year
- Food insecurity
- Social isolation
- Stress
- Social support
- Living with partner
- Employment
- Household income
- House ownership

SDOH from geographic information system linkages

A participant's household latitude and longitude were linked to several data sources. Census tract information was drawn from the 2016 American Community Survey [36]. Air quality information was drawn from the 2017 Environmental Protection Agency Air Quality System [37]. County crime statistics were drawn from the 2014 Uniform Crime Reporting Data [38]. Variables are given as follows:

- Census tract percent education less than high school
- Census tract percent below poverty
- Census tract percent household income under $10,000
- Census tract percent household with food stamps
- Census tract percent unemployed
- Census tract percent vacant housing
- Census tract percent medically uninsured ages 18–34
- Air quality unhealthy for sensitive groups
- Air quality median score
- County rate of violent crime
- County rate of property crime
- County rate of drug possession crime
- County rate of drug sale crime

Self-reported chronic conditions

Participants were asked, "Have you ever been told by a doctor or other health professional that you have any of the following?":

- Coronary heart disease
- Angina (also called angina pectoris)
- A heart attack (also called myocardial infarction)
- A stroke
- Emphysema
- Chronic obstructive pulmonary disease (also called COPD)
- Asthma
- Cancer or malignancy of any kind
- Arthritis, gout, lupus, or fibromyalgia
- A seizure disorder or epilepsy
- A depressive disorder, including depression, major depression, dysthymia, or minor depression
- Diabetes or sugar diabetes

Statistical analysis

Information about sample demographics, SDoH exposure, and presence of chronic conditions is presented as percentages or means, as appropriate. Because all participants were given PROPr but participants were randomly assigned to complete either the HUI Mark 2 and 3 or the EQ-5D and SF-6D, PROPr scores were split into two sets. This ensures that any differences in the number of SDoH that are statistically significant are not due to different sample sizes across measures. Summary scores of each measure are presented as histograms. Percent of participants at the ceiling was calculated, and score skewness was quantified with Pearson's moment coefficient of skewness. Convergent validity was evaluated using Pearson correlations between summary scores within subjects.

Known-groups validity was evaluated using the coefficient associated with SDoH created using ordinary least squares regression in which a summary score was regressed on age, gender, health conditions, and a single SDoH. The coefficient(s) for the SDoH is the estimated effect on the preference score of being exposed to a SDoH vs. not being exposed. Categorical variables (such as education category or insurance category) were coded as dummy variables. Because the ranges of the census tract and county variables varied widely, their associated coefficients also varied widely. These variables were standardized by multiplying a variable's range to its coefficient and standard error. For example, "census tract education less than high school" values vary from 0 to 0.659 and its unstandardized coefficient is -0.237 with a standard error of 0.052; its standardized coefficient is $-0.237 \times 0.659 = -0.165$ with a standard error of

$0.052 \times 0.659 = 0.036$. A separate analysis was done for each SDoH. Since all SDoH were coded such that higher scores indicated more hardship, negative coefficients are expected. A $p < 0.05$ was considered statistically significant. For SDoH with statistically significant coefficients in at least four of the six measures, Spearman correlations were calculated using the SDoH coefficients between the different measures.

All analyses were performed using SAS 9.4 (The SAS Institute, Cary, NC). All analyses except the Spearman correlations of SDoH coeffficients were weighted to be nationally representative using person weights provided by the survey company.

Results

The sample had 4142 participants. The average age was 47 and 52% of participants were female. Seventy-three percent of the participants reported their race as white, 13% reported black, and 15% reported other races; 16% of the sample reported Hispanic ethnicity and 3% completed the survey in Spanish. Eleven percent of the sample completed the survey by phone and the rest completed it by web interface. Negative SDoH exposures were common. For example, 45% reported difficulty paying their bills, 15% reported intimate partner violence within the last year, 30% reported some food insecurity, and 40% reported social isolation. Full descriptive information is available in the supplement with the source of the question and coding information.

Figure 1 illustrates the distribution of summary scores. The percent of participants with a score at the ceiling was 11.8% for the EQ-5D, 4.8% for the HUI Mark 2, 3.6% for the HUI Mark 3, 0.4% for PROPr set 1, 0.7% for PROPr set 2, and 1.2% for the SF-6D. All distributions were negatively skewed; Pearson's moment coefficient of skewness was -2.7 for the EQ-5D, -2.2 for the HUI Mark 2, -1.9 for the HUI Mark 3, -0.1 for PROPr set 1, -0.1 for PROPr set 2, and -0.8 for the SF-6D. For comparison, this coefficient is 0 for a symmetric distribution (such as a normal distribution), around 1 for a half-normal distribution, and 2 for an exponential distribution.

Pearson correlations were 0.72 between EQ-5D-5L and PROPr set 1; 0.74 between EQ-5D-5L and SF-6D; 0.60 between HUI Mark 2 and PROPr set 2; 0.85 between HUI Mark 2 and HUI Mark 3; 0.62 between HUI Mark 3 and PROPr set 2; 0.71 between SF-6D and PROPr set 1. For context, Evans suggests interpreting correlations of 0.40 to 0.56 as moderate, 0.60 to 0.79 as strong, and 0.80 to 1.0 as very strong [39]. It should be noted that the correlation between HUI Mark 2 and HUI Mark 3 is inflated because these measures use some of the same questions.

Figures 2 and 3 illustrate the age-, gender-, and chronic condition-adjusted SDoH coefficients in self-reported SDoH
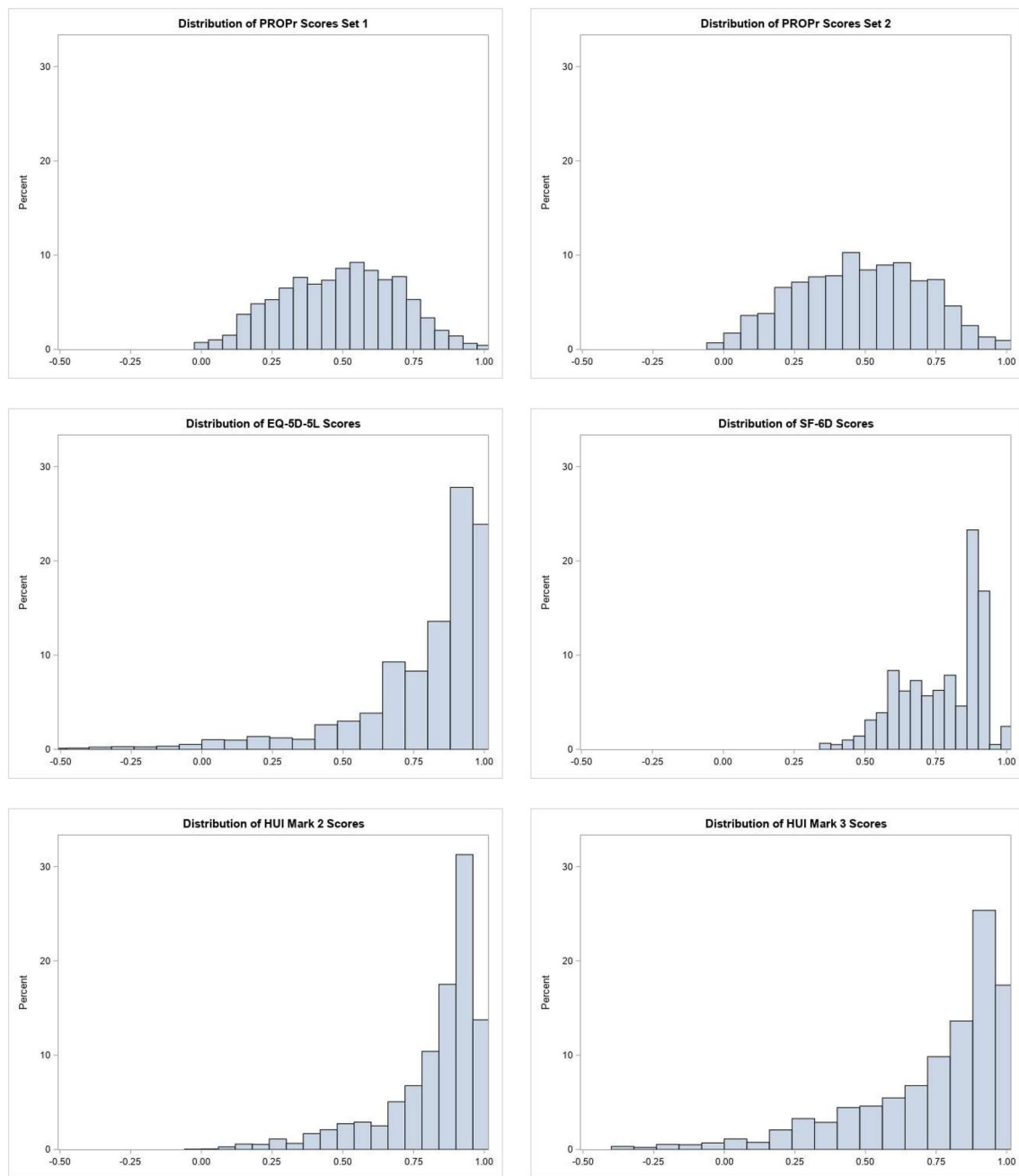
Fig. 1 Distribution of summary scores

variables and community SDoH variables. Results used to generate Figs. 2 and 3 are in Supplement 2. Of the 42 SDoH variables tested, 27 had statistically significant coefficients for EQ-5D, PROPr (both sets), and SF-6D. Statistical significance was reached for 17 SDoH variables with HUI Mark 2 and 23 variables with HUI Mark 3. As the figures illustrate, the measures tend to move together across SDoH. Fourteen of the SDoH had statistically significant coefficients in all of the measures and another six had statistically significant coefficients in five of the six measures. The measures show

significant differences in HRQoL for a variety of person-level and community-level variables, particularly related to education, income, food and financial insecurity, and social interactions. The measures used in this study are not associated with county variables.

The figures also illustrate that there is variance across the measures within each SDoH variable. The average SDoH coefficients by measure were -0.037 (95% CI − 0.051 to − 0.024) for the SF-6D, − 0.039 (95% CI − 0.053 to − 0.026) for the HUI Mark 2, − 0.063 (95% CI − 0.084

Self-Reported Variables

Census Tract and County Variables

to − 0.043) for the HUI Mark 3, − 0.059 (95% CI − 0.086 to − 0.048) for PROPr set 1, − 0.067 (95% CI − 0.077 to − 0.041) for PROPr set 2, and -0.086 (95% CI − 0.110 to − 0.063) for the EQ-5D-5L. Despite the difference in absolute values of SDoH coefficients across the different measures, the rankings of SDoH coefficients by different measures were often similar. There were 27 SDoH with statistically significant coefficients in at least four of the six measures. These SDoH are starred in Supplement 2. Table 1 shows the pairwise Spearman correlations between different measures across estimates for these 27 SDoH coefficients. All have Spearman correlations above 0.70.

Discussion

This report provides evidence of the validity of the PROPr score by testing its association with SDoH on HRQoL using a large cross-sectional dataset from the general US population. PROPr discriminates between those with and without a variety of SDoH exposures in the expected direction. PROPr was simultaneously administered with 4 other widely used preference-based measures that have been shown to be associated with SDoH: EQ-5D, HUI Mark 2, HUI Mark 3, and SF-6D. PROPr has good convergent validity with these other measures when measured as correlations of scores and as associations with SDoH. These findings support the use of PROPr to monitor population health in the general US population.

PROPr scores were associated with self-reported variables (education, income, employment, difficulty getting to medical appointments, food and financial insecurity, intimate partner violence, stress, and social isolation) and census tract variables (education, poverty, unemployment, vacant housing), but not county variables (air quality, crime). These county-level variables may be too coarse to apply to individual participants. Even when adjusted for demographic variables and chronic conditions, several expected SDoH exposures were not associated with a difference in scores (not having a usual medical provider, medical insurance type, employment categories, living with a partner, home ownership). This may be due to incomplete model adjustment. For example, though we may expect that those with a usual medical provider have better HRQoL than those

without, this comparison may be confounded by healthy people not establishing a relationship with a medical provider. To keep the many analyses in this report consistent, further analyses were not performed to better understand these null findings.

Given the wide range of variables to which they are sensitive, any of the measures used in this report appear appropriate for use in population health monitoring, including calculating aggregated indices of morbidity and mortality such as Quality Adjusted Life Years (QALYs). If used for monitoring population health, these measures would provide the added benefit of being able to be compared to clinical interventions because clinical research often uses these measures [40–42]. As shown in other studies, even when controlling for demographics and health conditions, SDoH are associated with a substantial difference in HRQoL. Using these measures to evaluate programs intended to improve overall population health by addressing SDoH will be able to show the value of these programs relative to clinical interventions.

This study found that the coefficients associated with SDoH vary substantially by measure. The measures with the smallest average coefficients were the HUI Mark 2 and SF-6D. The measure with the largest average coefficients was the EQ-5D-5L with US scoring. These findings are consistent with other regression-based analyses of co-administered measures, as well as analyses using item response theory, which show different measures assign varying decrements to the same change in a latent joint construct of health [43]. PROPr's average coefficients were in the middle of the tested measures. Despite the magnitude of coefficients varying across measures, the relative coefficients, as measured by Spearman correlation, of these SDoH was similar across these preference-based scores.

Findings in this report are limited by the data being cross-sectional, the unknown severity of chronic conditions, and relatively simple questions about SDoH. While there are several concerns about the representativeness of online panel surveys [44], the project was intended to show sensitivity to SDoH without quantifying the national impact of SDoH. The project's strengths include its use of standardized questions from large US surveys when available and co-administration of common generic HRQoL measures.

As with several other studies that co-administered these measures, these results support selecting a single

Table 1 Spearman correlations of measures using 27 social determinants of health coefficients

	HUI Mark 2	HUI Mark 3	PROPr Set 1	PROPr Set 2	SF-6D
EQ-5D-5L	0.77	0.70	0.87	0.80	0.94
HUI Mark 2		0.92	0.87	0.80	0.80
HUI Mark 3			0.83	0.74	0.71
PROPr Set 1				0.81	0.88
PROPr Set 2					0.83

preference-based measure that can be uniformly applied across research, clinical, and population health studies. Though use of any one measure will give similar health condition rankings to those of any other single measure, mixing health condition estimates from different measures can create substantial variation in condition rankings [45–47, 48]. PROPr has several arguments in its favor for use as a harmonizing measure. These include: PROPr is free to use, has less ceiling effect and skew in the general population than other measures, and is linked to a robust descriptive system (PROMIS). Though not presented in this report, collection of PROPr scores requires the collection of seven common PROMIS domains which can also be analyzed and compared to national samples. Arguments against the use of PROPr include its relative newness, limited number of validity studies, and its requirement to collect at least 14 questions.

Harmonization of measures has the potential to substantially improve the practice of clinical outcomes research, cost-effectiveness analysis, and population monitoring. The findings presented in this report provide evidence of construct validity for PROPr: it is correlated with other widely used generic preference-based summary scores and generates lower scores for those persons with negative SDoH exposures than for those without. Future work is needed to validate PROPr using longitudinal data collection, particularly with interventions that are expected to change HRQoL.

Data availability https://osf.io/63548/

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Consent to participate Voluntary consent was obtained before participants could proceed to the survey, and participants were free to withdraw their participation at any time.

Consent for publication The author hereby gives consent for publication this article in *Quality of Life Research*.

Ethical approval University of Pittsburgh IRB PRO17080294.

References

1. Ashraf, K., Ng, C. J., Teo, C. H., & Goh, K. L. (2019). Population indices measuring health outcomes: A scoping review. *Journal of Global Health*. https://doi.org/10.7189/jogh.09.010405
2. Kindig, D. A. (2000). *Purchasing population health: Paying for results*. Ann Arbor: University of Michigan Press.
3. Breslow, L. (2006). Health measurement in the third era of health. *American Journal of Public Health, 96*(1), 17–19. https://doi.org/10.2105/ajph.2004.055970
4. Jones, N., Jones, S. L., & Miller, N. A. (2004). The Medicare Health Outcomes Survey program: Overview, context, and near-term prospects. *Health and Quality of Life Outcomes*. https://doi.org/10.1186/1477-7525-2-33
5. *Medicare Health Outcomes Survey*. Retrieved April 10, 2020, from https://www.hosonline.org/
6. Cohen, S. B. (2003). Design strategies and innovations in the medical expenditure panel survey. *Medical Care*. https://doi.org/10.1097/01.mlr.0000076048.11549.71
7. McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires*. New York: Oxford University Press.
8. Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics, 5*(1), 1–30. https://doi.org/10.1016/0167-6296(86)90020-2
9. Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value trade-offs*. Cambridge: Cambridge University Press.
10. Cherepanov, D., Palta, M., Fryback, D. G., & Robert, S. A. (2010). Gender differences in health-related quality-of-life are partly explained by sociodemographic and socioeconomic variation between adult men and women in the US: Evidence from four US nationally representative data sets. *Quality of Life Research, 19*(8), 1115–1124. https://doi.org/10.1007/s11136-010-9673-x
11. Coley, S. L., Leon, C. F. M. D., Ward, E. C., Barnes, L. L., Skarupski, K. A., & Jacobs, E. A. (2017). Perceived discrimination and health-related quality-of-life: gender differences among older African Americans. *Quality of Life Research, 26*(12), 3449–3458. https://doi.org/10.1007/s11136-017-1663-9
12. Robert, S. A., Cherepanov, D., Palta, M., Dunham, N. C., Feeny, D., & Fryback, D. G. (2009). Socioeconomic status and age variations in health-related quality of life: Results from the national health measurement study. *The Journals of Gerontology Series B, 64B*(3), 378–389. https://doi.org/10.1093/geronb/gbp012
13. Mills, S. D., Fox, R. S., Bohan, S., Roesch, S. C., Sadler, G. R., & Malcarne, V. L. (2020). Psychosocial and neighborhood correlates of health-related quality of life: A multi-level study among Hispanic adults. *Cultural Diversity and Ethnic Minority Psychology, 26*(1), 1–10. https://doi.org/10.1037/cdp0000274
14. Sellers, S. L., Cherepanov, D., Hanmer, J., Fryback, D. G., & Palta, M. (2013). Interpersonal discrimination and health-related quality of life among black and white men and women in the United States. *Quality of Life Research, 22*(6), 1313–1318. https://doi.org/10.1007/s11136-012-0278-4
15. Twardzik, E., Clarke, P., Elliott, M. R., Haley, W. E., Judd, S., & Colabianchi, N. (2019). Neighborhood socioeconomic status and trajectories of physical health-related quality of life among stroke survivors. *Stroke, 50*(11), 3191–3197. https://doi.org/10.1161/strokeaha.119.025874
16. Committee on Educating Health Professionals to Address the Social Determinants of Health; Board on Global Health; Institute of Medicine; National Academies of Sciences, Engineering, and Medicine. *A framework for educating health professionals to address the social determinants of health*. Washington (DC):

National Academies Press (US); 2016 Oct 14. 3, Frameworks for addressing the social determinants of health. Available from: https://www.ncbi.nlm.nih.gov/books/NBK395979/

17. World Health Organization. (2010). A conceptual framework for action on the social determinants of health: debates, policy & practice, case studies. In *A conceptual framework for action on the social determinants of health: debates, policy & practice, case studies*. Geneva.

18. Brooks, R. G., Rabin, R., & Charro, F. D. (2010). *The measurement and valuation of health status using Eq-5D: a European perspective: Evidence from the EuroQol Biomed Research Programme*. Dordrecht: Kluwer Academic Publishers.

19. Pickard, A., Law, E., Jiang, R., Oppe, M., Shaw, J., Xie, F., … Balch, A. (2019). United States valuation of EQ-5D-5L health states: an initial model using a standardized protocol. *Value in Health*, 22(8), 931–941. doi:10.1016/j.jval.2019.02.009

20. Feeny, D., Torrance, G. W., & Furlong, W. (1996). Health Utilities Index. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., pp. 239–252). Philadelphia: Lippincott-Raven Press.

21. Feeny, D., Furlong, W., Torrance, G.W., Goldsmith, C.H., Zhu, Z., Depauw, S., … Boyle, M. (2002). Multiattribute and single-attribute utility functions for the Health Utilities Index Mark 3 System. *Medical Care*, 40(2), 113–128. doi:10.1097/00005650-200202000-00006

22. Brazier, J. E., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care, 42*(9), 851–859. https://doi.org/10.1097/01.mlr.0000135827.18610.0d

23. Hanmer, J., Feeny, D., Fischhoff, B., Hays, R.D., Hess, R., Pilkonis, P.A., … Yu, L. (2015). The PROMIS of QALYs. *Health and Quality of Life Outcomes*, 13(1). doi:https://doi.org/10.1186/s12955-015-0321-6

24. Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., … Rose, M. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*, 45(Suppl 1). doi:https://doi.org/10.1097/01.mlr.0000258615.42478.55

25. Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. New York: Psychology Press. https://doi.org/10.4324/9781410605269

26. Dewitt, B., Feeny, D., Fischhoff, B., Cella, D., Hays, R.D., Hess, R., … Hanmer, J. (2018). Estimation of a preference-based summary score for the Patient-Reported Outcomes Measurement Information System: The PROMIS®-Preference (PROPr) scoring system. *Medical Decision Making*, 38(6), 683–698. doi:10.1177/0272989x18776637

27. Hanmer, J., Dewitt, B., Yu, L., Tsevat, J., Roberts, M., Revicki, D., … Cella, D. (2018). Cross-sectional validation of the PROMIS-Preference scoring system. *Plos One*, 13(7). doi:https://doi.org/10.1371/journal.pone.0201093

28. NORC . N.d. "AmeriSpeak: NORC's Breakthrough Panel-Based Research Platform". Retrieved April 13, 2020. https://amerispeak.norc.org/Pages/default.aspx

29. Hanmer, J. Developing the PROMIS-preference score for monitoring population health outcomes. https://osf.io/63548/.

30. Hanmer, J., Cella, D., Feeny, D., Fischhoff, B., Hays, R.D., Hess, R., … Yu, L. (2017). Selection of key health domains from PROMIS® for a generic preference-based scoring system. *Quality of Life Research*, 26(12), 3377–3385. doi: 10.1007/s11136-017-1686-2

31. Hanmer, J., Cella, D., Feeny, D., Fischhoff, B., Hays, R. D., Hess, R., … Yu, L. (2018). Evaluation of options for presenting health-states from PROMIS® item banks for valuation exercises. *Quality of Life Research*, 27(7), 1835–1843. doi:10.1007/s11136-018-1852-1

32. Stolk, E., Ludwig, K., Rand, K., Hout, B. V., & Ramos-Goñi, J. M. (2019). Overview, update, and lessons learned from the International EQ-5D-5L valuation work: Version 2 of the EQ-5D-5L valuation protocol. *Value in Health, 22*(1), 23–30. https://doi.org/10.1016/j.jval.2018.05.010

33. -Item Short Form Survey from the RAND Medical Outcomes Study. (n.d.). Retrieved April 13, 2020. https://www.rand.org/health-care/surveys_tools/mos/12-item-short-form.html

34. NHIS - Questionnaires, Datasets, and Related Documentation. (2019). Retrieved April 13, 2020, from https://www.cdc.gov/nchs/nhis/nhis_questionnaires.htm

35. NHANES - National Health and Nutrition Examination Survey Homepage. (2020). Retrieved April 13, 2020, from https://www.cdc.gov/nchs/nhanes/index.htm

36. US Census Bureau. (2020). American Community Survey (ACS). Retrieved April 13, 2020, from https://www.census.gov/programs-surveys/acs/

37. Air Quality System (AQS). (2020). Retrieved April 13, 2020, from https://www.epa.gov/aqs

38. Uniform Crime Reporting (UCR) Program. (2018). Retrieved April 13, 2020, from https://www.fbi.gov/services/cjis/ucr

39. Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing Company.

40. Feeny, D. H., & Torrance, G. W. (1989). Incorporating utility-based quality-of-life assessment measures in clinical trials: Two examples. *Medical Care, 27*(S3), 190-S204. https://doi.org/10.1097/00005650-198903001-00016

41. Wolowacz, S. E., Briggs, A., Belozeroff, V., Clarke, P., Doward, L., Goeree, R., et al. (2016). Estimating health-state utility for economic models in clinical studies: An ISPOR good research practices task force report. *Value in Health, 19*(6), 704–719. https://doi.org/10.1016/j.jval.2016.06.001

42. Neumann, P.J., Sanders, G.D., Russell, L.B., Siegel, J.E. and Ganiats, T.G. eds. (2016). Cost-effectiveness in health and medicine. Oxford University Press. doi: 10.1093/acprof:oso/9780190492939.003

43. Fryback, D. G., Palta, M., Cherepanov, D., Bolt, D., & Kim, J. S. (2010). Comparison of 5 health-related quality-of-life indexes using item response theory analysis. *Medical Decision Making, 30*(1), 5–15. https://doi.org/10.1177/0272989X09347016

44. Liu, H., Cella, D., Gershon, R., Shen, J., Morales, L. S., Riley, W., & Hays, R. D. (2010). Representativeness of the patient-reported outcomes measurement information system internet panel. *Journal of Clinical Epidemiology, 63*(11), 1169–1178. https://doi.org/10.1016/j.jclinepi.2009.11.021

45. Hanmer, J., Cherepanov, D., Palta, M., Kaplan, R. M., Feeny, D., & Fryback, D. G. (2016). Health condition impacts in a nationally representative cross-sectional survey vary substantially by preference-based health index. *Medical Decision Making, 36*(2), 264–274. https://doi.org/10.1177/0272989x15599546

46. Tengs, T. O., & Wallace, A. (2000). One thousand health-related quality-of-life estimates. *Medical Care, 38*(6), 583–637. https://doi.org/10.1097/00005650-200006000-00004

47. Mcdonough, C. M., & Tosteson, A. N. A. (2007). Measuring preferences for cost-utility analysis. *PharmacoEconomics, 25*(2), 93–106. https://doi.org/10.2165/00019053-200725020-00003