Title

Selection of key health domains from PROMIS® for a generic preference-based scoring system

Journal

Quality of Life Research, 26(12)

ISSN

0962-9343

Authors

Hanmer, Janel
Cella, David
Feeny, David
et al.

Publication Date

2017-12-01

DOI

10.1007/s11136-017-1686-2

Peer reviewed

Qual Life Res (2017) 26:3377–3385
DOI 10.1007/s11136-017-1686-2

Selection of key health domains from PROMIS® for a generic preference-based scoring system

Janel Hanmer[1] · David Cella[2] · David Feeny[3,4] · Baruch Fischhoff[5] · Ron D. Hays[6] · Rachel Hess[7] · Paul A. Pilkonis[8] · Dennis Revicki[9] · Mark Roberts[10,11] · Joel Tsevat[12] · Lan Yu[1]

Accepted: 11 August 2017 / Published online: 19 August 2017

Abstract

Purpose We sought to select a parsimonious subset of domains from the patient-reported outcomes measurement information system (PROMIS®) that could be used for preference-based valuation. Domain selection criteria included face validity, comprehensiveness, and structural independence.

Methods First, 9 health outcomes measurement experts selected domains appropriate for a general health measure using a modified Delphi procedure. Second, 50 adult community members assessed structural independence of domain pairs. For each pair, the participant was asked if it were possible to have simultaneously good functioning in domain 1 but poor functioning in domain 2, and vice versa. The community members also rated the relative importance of the domains. Finally, the experts selected domains, guided by community members' judgments of structural independence and importance.

Results After 3 rounds of surveys, the experts agreed on 10 potential domains. The percent of pairs deemed structurally independent by community members ranged from 50 to 95 (mean = 78). Physical Function, Pain Interference, and Depression were retained because of their inclusion in existing preference-based measures and their importance to community members. Four other domains were added because they were important to community members and

✉ Janel Hanmer
hanmerjz@upmc.edu

David Cella
D-cella@northwestern.edu

David Feeny
feeny@mcmaster.ca

Baruch Fischhoff
baruch@cmu.edu

Ron D. Hays
drhays@ucla.edu

Rachel Hess
rachel.hess@hsc.utah.edu

Paul A. Pilkonis
pilkonispa@upmc.edu

Dennis Revicki
Dennis.Revicki@evidera.com

Mark Roberts
mroberts@pitt.edu

Joel Tsevat
joel.tsevat@uc.edu

Lan Yu
Yul2@upmc.edu

[1] Department of General Internal Medicine, University of Pittsburgh Medical Center, 230 McKee Place, Suite 600, Pittsburgh, PA 15213, USA

[2] Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[3] Department of Economics, McMaster University, Hamilton, ON, Canada

[4] Health Utilities Incorporated, Dundas, ON, Canada

[5] Department of Engineering and Public Policy and Institute for Politics and Strategy, Carnegie Mellon University, Pittsburgh, PA, USA

[6] Division of General Internal Medicine & Health Services Research, UCLA, Los Angeles, CA, USA

[7] Division of Health System Innovation and Research, University of Utah Schools of the Health Sciences, Salt Lake City, UT, USA

[8] Department of Psychiatry, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

judged to be independent by at least 67% of respondents: Cognitive Function—Abilities; Fatigue; Ability to Participate in Social Roles and Activities; and Sleep Disturbance. *Conclusion* With input from measurement experts and community members, we selected 7 PROMIS domains that can be used to create a preference-based score.

Keywords Health-related quality of life · Utility · Multi-attribute utility instrument · Health domains · PROMIS® · Health status

Introduction

Self-reports about health-related quality of life (HRQL) provide information about the health of individuals and populations, the impact of chronic medical conditions, and the effectiveness of health care interventions [1]. A variety of health outcome measures have been developed and applied in population-based and clinical studies over the past 40 years [2].

The Patient-Reported Outcomes Measurement Information System (PROMIS®) is a U.S. National Institutes of Health funded resource of self-reported health measures assessing multiple domains (e.g., pain, fatigue, emotional distress, physical function, social function) relevant for individuals from the general population and those with chronic illnesses [3, 4]. PROMIS was developed as a health descriptive system in which a score is calculated for each domain collected [3]. There are currently over 70 different domains relevant to adult health and illness available in the PROMIS Assessment Center [see www.healthmeasures.net]. Because items within a domain are calibrated on the same underlying metric using item response theory, any subset of items can be used to estimate the domain score. This subset may be a short for or may be selected using computer adaptive testing.

Despite its strengths, the PROMIS measurement system does not provide a single summary score for overall health. A comprehensive summary score would be useful for many descriptive purposes (e.g., to track groups over time, to compare across groups). A summary score constructed using preference-based methods and anchored on a 0 ("dead") to 1 ("full health") scale could also be used in decision analyses and cost-effectiveness analyses [5–8]. Such preference-based scores allow both morbidity and mortality to be captured in the form of quality-adjusted life years (QALYs) [7–9]. Cost-effectiveness analyses are ideally suited to incorporate such metrics with resource use data to compare the value of healthcare interventions [10–13].

Developing preference-based scores requires having representatives from the population making explicit trade-offs between different health states using methods such as the standard gamble [6], time trade-off [6], or discrete choice experiments [14]. Because these trade-offs are cognitively complex, score developers need to limit the number of domains included in a given preference-based measure.

For explicit trade-offs to make sense, domains must be structurally independent. Structural independence means that the range of possible outcomes on domain A is potentially independent of the outcome on domain B (and vice versa). For example, physical function and depression are structurally independent if one can imagine an individual with good levels of physical function and very severe depression as well as an individual who is not depressed with very poor physical function. Domains can be structurally independent even if they are highly correlated (e.g., depressed individuals tend to have poor physical functioning).

For the sake of comparability, and the accumulation of evidence across studies, it would be useful to have a standard set of domains from PROMIS used to develop a preference-based score. Here, we derive such a set, driven by the criteria of face validity, comprehensiveness, and structural independence. In this paper, we describe a process involving input from both measurement experts and community members to select a subset of the adult domains available in the PROMIS domain framework [4, 15]. This is the first step in a larger project to develop a preference-based scoring system for PROMIS [16]. Subsequent steps include selecting representative items from the domains, eliciting valuations from the general population, and estimating a scoring function using the valuations. The preference-based score constructed from PROMIS domains will provide an important extension to PROMIS.

Methods

We used a 4-step procedure to select a subset of PROMIS domains for the PROMIS-preference scoring system.

[9] Outcomes Research, Evidera, Bethesda, MD, USA

[10] Department of General Internal Medicine, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

[11] Department of Health Policy and Management, University of Pittsburgh, Pittsburgh, PA, USA

[12] Division of General Internal Medicine, University of Cincinnati College of Medicine and Cincinnati VA Medical Center, Cincinnati, OH, USA

Step 1: reduction of available domains by a modified delphi procedure with experts

We recruited 9 health status measurement experts to participate in a modified Delphi procedure. A list of all 37 domains available at the time of the study was provided to the experts ("Appendix A"). They were asked to exclude domains that "were either not appropriate for inclusion in a general health utility measure or redundant with another PROMIS domain." We removed a domain if 7 or more experts voted for its removal. For domains that were considered to be very similar or redundant (e.g., Cognitive Function—General Concerns and Cognitive Function—Abilities), a single domain was selected by a simple majority rule. Responses were combined and presented again in further rounds until consensus was reached.

Step 2: pairwise independence evaluation by a community sample

We next recruited a convenience sample of 50 community dwelling adults from the University of Pittsburgh's Clinical and Translational Science Institute research registry website (https://pittplusme.org/; http://www.ctsi.pitt.edu/research-rsp-pitt.html). Eligible participants were at least 18 years old and spoke English; there were no other exclusion criteria. Participants were compensated $20 for their time.

We conducted face-to-face interviews in which participants evaluated 20 randomly assigned pairs of domains selected in Step 1. The 20 domain pairs were generated using Microsoft Excel's randomization function. The pages with the domain pair (as in Fig. 1) were generated in Word which was linked to the Excel file using "field" link. A research assistant manually checked each set of 20 pairs to ensure that: (1) all 10 domains were included in the set, and (2) all pairs in the set were unique.

Participants were asked to determine whether pairs of the identified domains were structurally independent with the following procedure. The participants were given the name and description of each domain on a piece of paper (example in Fig. 1). The content of these descriptions was based on the definitions for the different domains in PROMIS [4, 15]. A research assistant provided an example of good functioning on domain 1 and poor functioning on domain 2 and asked the participant if this combination was possible. Participants were then asked if the opposite combination was also possible. The research assistant explained that we wished to know whether the combinations were *possible* even if not necessarily *probable*. That is, could they imagine such a combination of health states ever happening? Each participant reviewed each of the health domains in at least one pair.

For each domain pair, we calculated the percent of respondents who reported its members to be structurally independent.

Step 3: domain importance in a community sample

After comparing pairs of health domains (Step 2), participants were asked "I'd like you to consider a person's overall quality of life. Between these two domains of health, which do you think is more important in overall quality of life?" The participant could select a domain or indicate that they were equally important.

Participants were given cards with each domain from Step 2. They could remove any domain if they did not think it was important for overall HRQL. The participant was then asked to rank the remaining domains from most important to least important. The most important domain was placed at 100 on a visual analog scale (VAS) board and the least important at 0. The participant was asked to place the remaining domain cards along the VAS and their scores were recorded by the research assistant.

After completing the rating task, participants provided the following demographic and clinical information: age, gender, race/ethnicity, education, number of times they see a doctor in a year, number of medications they take daily, number of times they've ever been hospitalized (excluding labor and delivery), and experience with health problems that limit a person's ability to take care of him or herself.

Step 4: final domain selection

We used information from Steps 2 and 3 to create the final set of domains for the PROMIS-preference measure. The expert panel agreed upon a minimum of three specific domains essential for face validity because they are present in many generic HRQL profile measures: Physical Functioning, Depression, and Pain. We then used information about structural independence from the community sample to determine which domains could be added to this core set and maintain the same level of structural independence. Among the three pairs, the lowest level of structurally independence was for pain and physical functioning: 67% reported these two domains to be structurally independent. We accepted the same level of structural independence (or better) when including additional domains. When structural independence could be maintained with either but not both of a domain pair, we were guided by the community participants' importance ratings.

IRB approval for the project was obtained from the University of Pittsburgh (PRO14070021 and PRO14100533).

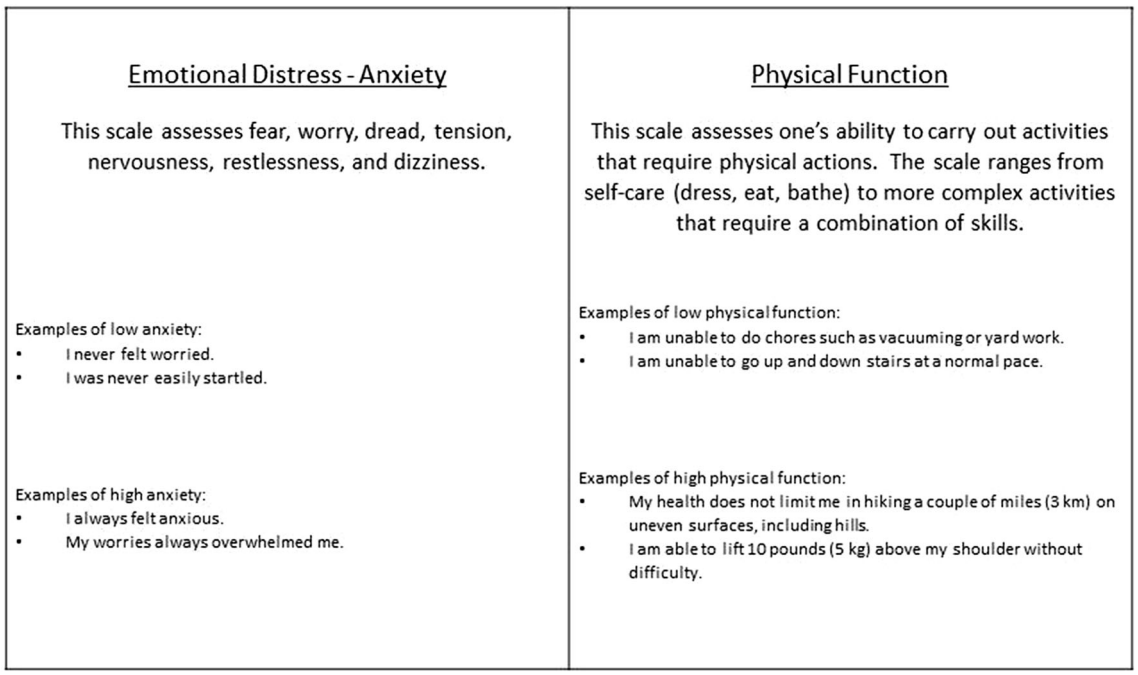

Fig. 1 Example of a 2-domain comparison. In this example, the participant was asked to determine whether Anxiety and Physical Function are structurally independent. The research assistant uncovered the names and descriptions first. The research assistant then uncovered good functioning (low anxiety) on the *left* side and bad functioning (low physical function) on the *right* side and asked the participant if he/she thought this combination was possible. The research assistant then uncovered the rest of the page, with bad functioning on the *left* side and good functioning on the *right* side, and asked the participant if this combination was possible. If both combinations were possible, the pair was recorded as structurally independent

Results

Step 1: reduction of available domains by a modified delphi procedure with experts

In the first round, all 9 experts voted to include 5 and 16 domains were removed by at least 7 of the expert respondents. This left 16 domains for further evaluation in the second round. In the second round, a single cognitive function domain (Cognitive Function—Abilities) was selected from a set of 2 such domains (Cognitive Function—Abilities and Cognitive Function—General Concerns). Likewise, a single Social Roles Domain (Ability to Participate in Social Roles and Activities v2.0) was selected from a set of 3 domains about social roles. One other domain (Sleep Disturbance) was included and 9 domains were removed. For 2 domains, there was no clear vote for exclusion or inclusion—Pain Intensity and Sexual Function and Satisfaction: Global Satisfaction with Sex Life—so these were included for further testing in the community sample. Details of the Delphi results are summarized in Table 1.

In the third round, all of the experts endorsed a final set of 10 potential domains. These 10 domains were: Cognitive Function—Abilities (referred to subsequently as Cognition), Emotional Distress—Anxiety (Anxiety), Emotional Distress—Depression (Depression), Fatigue, Pain Intensity, Pain Interference, Physical Function, Ability to Participate in Social Roles and Activities (Social Roles), Sleep Disturbance, and Sexual Function and Satisfaction: Global Satisfaction with Sex Life (Sexual Function).

Step 2: pairwise independence evaluation by a community sample

Convenient sample of 50 adults were recruited from the Pittsburgh area. The sample were 60% female and with a mean age of 44 (range 22–70). Fifty-two percent of the sample were White, 32% were Black, and 7% identified their race as Other. Their self-rated health ratings were: excellent (27%), very good (41%), good (29%), fair (4%), and poor (0%).

Domain pairs were evaluated 20–27 times (mode = 22) (Table 2). Across all comparisons, 78% of domain pairs were judged to be structurally independent. The pair that was least often considered as structurally independent was Anxiety and Sleep Disturbance (50%). The pairs that were most often reported as structurally independent were Depression and Physical Function, Pain Intensity and Physical Function, and Sexual Function and Social Roles (95% for all 3 pairs).

Table 1 Summary of Delphi Results

Included	Undecided	Excluded
Delphi round 1 results		
Emotional distress—anxiety	Cognitive Function—abilities	Alcohol—alcohol use
Emotional distress—depression	Cognitive Function—general concerns	Alcohol—positive consequences
Fatigue	Satisfaction with social roles and activities v1.0	Alcohol—negative consequences
Pain interference	Satisfaction with social roles and activities v2.0	Alcohol—positive expectancies
Physical function	Ability to participate in social roles and activities v2.0	Alcohol—negative expectancies
	Companionship	Informational support
	Emotional distress—anger	Mobility
	Emotional support, instrumental support	Pain—behavior
	Global health	Upper extremity
	Pain intensity	Satisfaction with social roles and activities v1.0
	Physical function for mobility aid users	Satisfaction with participation in Discretionary social activities v1.0
	Psychosocial illness impact—positive	Sexual function and satisfaction: interest in sexual activity
	Psychosocial illness impact—negative	Sexual function and satisfaction: lubrication
	Social isolation	Sexual function and satisfaction: vaginal discomfort
	Sexual function satisfaction: global satisfaction with sex life	Sexual function and satisfaction: erectile function
	Sleep disturbance	Sleep-related impairment

Delphi round 2 results		
Cognitive Function—abilities	Pain intensity	Emotional distress—anger
Ability to participate in social roles and activities v2.0	Sexual function satisfaction: global satisfaction with sex life	Physical function for mobility aid users
Sleep disturbance		Psychosocial illness impact—positive
		Psychosocial illness impact—negative
		Companionship
		Emotional support
		Instrumental support
		Social isolation
		Global health
		Cognitive Function - general concerns

Confirmed
Delphi round 3 results
Cogntivie Function—abilities v2.0
Ability to participate in social roles and activities v2.0
Emotional distress—anxiety v1.0
Emotional distress—depression v1.0
Fatigue v1.0
Pain intensity v1.0
Pain interference v1.1
Physical function v1.2
Sexual function satisfaction: global satisfaction with sex life v1.0
Sleep disturbance v1.0

Step 3: domain importance in a community sample

One participant, while able to complete Step 2, was unable to comprehend and complete the ranking and rating task, leaving 49 participants in this analysis. Before ranking, 1 respondent removed Fatigue, Pain Intensity, Pain Interference, and Sexual Function; 3 respondents removed Sexual Function.

Table 2 Percent of community respondents reporting structural independence of domain pairs

Domain	Cognition	Anxiety	Depression	Fatigue	Pain interference	Pain intensity	Physical function	Social roles	Sexual function
Cognition	–	–	–	–	–	–	–	–	–
Anxiety	68	–	–	–	–	–	–	–	–
Depression	68	81	–	–	–	–	–	–	–
Fatigue	68	91	85	–	–	–	–	–	–
Pain interference	90	91	76	86	–	–	–	–	–
Pain intensity	68	68	80	88	**61**	–	–	–	–
Physical function	91	74	95	68	67	95	–	–	–
Social roles	81	79	71	72	81	**55**	87	–	–
Sexual function	91	91	73	**64**	77	**52**	86	95	–
Sleep disturbance	68	**50**	86	90	76	77	77	86	74

Pairs with less than 67% of respondents reporting the pair to be structurally independent are bolded

Table 3 Visual analog scale score and rank of health domains

Domain	N	VAS score, mean (SD)	Rank, mean (SD)*
Physical function	49	75 (30)	3.3 (2.5)
Cognition	49	75 (30)	3.3 (2.7)
Pain intensity	48	68 (32)	3.6 (2.4)
Sleep disturbance	49	67 (27)	4.0 (2.5)
Depression	49	60 (33)	4.2 (2.8)
Pain interference	48	55 (36)	4.6 (2.7)
Anxiety	49	50 (37)	5.1 (3.1)
Social roles	49	53 (34)	5.7 (2.8)
Fatigue	48	48 (35)	5.9 (3.0)
Sexual function	45	28 (34)	7.4 (2.7)

VAS visual analog scale

* Ranks are ordered from lowest (most important) to highest (least important)

Physical Function and Cognition were tied for the most important domain by both mean VAS score and rank order (Table 3). Sexual Function was the least important on the basis of number of responses, VAS score, and rank order.

Step 4: final domain selection

The expert panel had initially selected a small set of health domains necessary for face validity in a generic health measurement system: Physical Functioning, Depression, and Pain. We had the option of using Pain Interference or Pain Intensity to describe Pain. Pain Interference was chosen over Pain Intensity because it was assessed as more structurally independent when paired with other domains. Seventy-six percent of respondents reported Depression and Pain Interference to be structurally independent, 95% reported Depression and Physical Functioning to be structurally independent, and 67% reported Pain Interference and Physical Functioning to be structurally independent. We considered including any other domains that had structural independence of at least 67% with other domains in the set.

To maintain this level of structural independence, there were 2 cases where the expert panel could include 1, but not both of 2, domains. The first case involved Fatigue and Sexual Function. We chose Fatigue because the community sample rated it as more important for overall HRQL. The second case involved Sleep Disturbance and Anxiety. We selected Sleep Disturbance because it was viewed as more important than Anxiety by the community sample.

The final set of 7 domains, the PROMIS-preference measure, and their structural independence is summarized in Table 4.

Discussion

We used judgments from an expert panel and a community sample to select 7 domains appropriate for a PROMIS-Preference scoring system. The 7 domains are: Cognitive Function—Abilities, Depression, Fatigue, Pain Interference, Physical Function, Ability to Participate in Social Roles and Activities, and Sleep Disturbance. These domains will be used to develop a descriptive system for preference valuation studies and development of a preference-based scoring function.

These 7 domains have substantial overlap with the domains in other generic health measurement systems. Six were also selected by a group of PROMIS experts for the 8 domain PROMIS Profile measures (PROMIS-29, PROMIS-43, PROMIS-57; www.healthmeasures.net). Previously developed generic preference-based measures such as the EQ-5D [17], Health Utilities Index [18, 19], Quality

Table 4 The percent of community dwelling respondents who reported that a particular combination of domains is structurally independent for the final set of domains

Domain	Cognition	Depression	Fatigue	Pain interference	Physical function	Social roles
Cognition	–	–	–	–	–	–
Depression	68	–	–	–	–	–
Fatigue	68	85	–	–	–	–
Pain interference	90	76	86	–	–	–
Physical function	91	95	68	67	–	–
Social roles	81	71	72	81	87	–
Sleep disturbance	68	86	90	76	77	86

of Well-being Index [20], and SF-6D [21, 22] all include physical functioning, pain, and mental health in some form; we therefore designed PROMIS-preference to include physical function, pain, and depression. Most existing generic preference-based measures also include social functioning; similarly, PROMIS-preference includes ability to participate in social roles and activities. PROMIS-preference, reflecting the input of its community sample includes 2 domains that have rarely been included: Cognitive Function has only been included in the Health Utilities Index Mark 2 and Mark 3; Fatigue has only been included in the SF-6D (as Vitality). Sleep Disturbance has not appeared in any of these other measures.

We believe that these differences reflect our selection process. Whereas domain selection for prior measures was primarily driven by the measure developers with informal input from other sources, we have attempted to formalize the domain selection process with both experts and community members. We first engaged a wide variety of measurement experts in a modified Delphi process to select a range of potentially relevant domains. We then engaged community members to assess the structural independence and importance of potential domains. To the best of our knowledge, this is the first time a formal evaluation of domains' structural independence by a community sample has been incorporated in the development of a preference-based health measure.

One potential limitation of these results is using a small geographically limited community sample in order to allow in-person interviews. Although this sample does not accurately represent the broader US general population, it was diverse in age, gender, race, educational background, and prior healthcare experience and should represent a substantial range of experiences and values. A second potential limitation is that the experts engaged in the Delphi procedure cannot be representative of the scientific community although they were chosen for the diversity of their opinions. A third potential limitation is community participants' lack of familiarity with the structural independence task which may have posed a barrier to obtaining informed responses. We believe that the in-person interview format allowed for needed clarification.

The motivation for this study is the opportunity to develop a preference-based score for the PROMIS domains and measures. This preference-based score can be estimated whenever the 7 PROMIS domains selected in this study have been collected, regardless of the method of collection (i.e., short form or computer adaptive testing). PROMIS measures are increasingly incorporated into clinical trials and practice based studies in the US [23]. PROMIS measures are also being translated and adopted throughout Europe and other countries [24]. A preference-based score developed from PROMIS domains would allow simultaneous collection of descriptive health status and a single summary score without increasing respondent burden. Moreover, the item response theory underpinnings of the PROMIS domains and measures will address problems in prior preference-based measurement systems such as ceiling and floor effects and limited precision in tracking individual outcomes [16].

To date, other preference-based scores have been estimated from PROMIS measures in 2 ways. One uses regression or linear equating methods to estimate EQ-5D-3L and HUI3 scores from PROMIS-29 and PROMIS Global scores, despite known limitations [25, 26]. The other used discrete choice experiments to elicit evaluations of health profiles constructed from the PROMIS-29 [27] without testing for structural independence. Neither method takes advantage of the improvements in health description provided by PROMIS, such as ensuring that the full range of health is measured and improving the precision of measurement.

Selecting a set of domains is the first step in developing a rigorous preference-based score for PROMIS [16]. Next steps include developing and implement a valuation elicitation method that links preference to PROMIS domain scores. This link will allow the use of techniques from multi-attribute utility theory to estimate the relative weight of each domain and an overall preference-based scoring function. The final result of this project will allow reporting

of both descriptive health outcomes and a preference-based score for studies that have included the 7 PROMIS domains selected in this study.

Acknowledgements Janel Hanmer was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number KL2TR001856. Participant recruitment was completed by the Clinical and Translational Science Institute at the University of Pittsburgh, which is supported by the National Institutes of Health Clinical and Translational Science Award program, Grants UL1RR024153 and UL1TR000005. The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to report.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. IRB approval for the project was obtained from the University of Pittsburgh (PRO14070021 and PRO14100533).

Appendix A: 37 adult PROMIS item banks considered for inclusion

Emotional Distress—Anger
Emotional Distress—Anxiety
Emotional Distress—Depression
Cognitive Function—Abilities
Cognitive Function—General Concerns
Psychosocial Illness Impact—Positive
Psychosocial Illness Impact—Negative
Alcohol—Alcohol Use
Alcohol—Positive Consequences
Alcohol—Negative Consequences
Alcohol—Positive Expectancies
Alcohol—Negative Expectancies
Fatigue
Pain—Behavior
Pain—Interference
Pain Intensity
Physical Function
Physical Function for Mobility Aid Users
Mobility
Upper Extremity
Sleep Disturbance
Sleep-Related Impairment
Sexual Function and Satisfaction: Global Satisfaction with Sex Life
Sexual Function and Satisfaction: Interest in Sexual Activity
Sexual Function and Satisfaction: Lubrication
Sexual Function and Satisfaction: Vaginal Discomfort
Sexual Function and Satisfaction: Erectile Function
Satisfaction with Participation in Discretionary Social Activities (v1.0)
Satisfaction with Social Roles and Activities (v1.0)
Satisfaction with Social Roles and Activities (v2.0)
Ability to Participate in Social Roles and Activities
Companionship
Informational Support
Emotional Support
Instrumental Support
Social Isolation
Global Health

References

1. McHorney, C. A. (1999). Health status assessment methods for adults: Past accomplishments and future challenges. *Annual Review of Public Health, 20,* 309–335.
2. Fryback, D., Dunham, N. C., Palta, M., Hanmer, J., Buechner, J., Cherepanov, D., et al. (2007). US norms for six generic health-related quality of life indexes from the National Health Measurement Study. *Medical Care, 45,* 1162–1170.
3. Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Medical Care, 45*(5), S3–11.
4. Cella, D., Riley, W., Reeve, B., Stone, A., Young, S., Rothrock, N., et al. (2010). Initial item banks and first wave testing of the patient-reported outcomes measurement information system (PROMIS) network: 2005–2008. *Journal of Clinical Epidemiology, 63*(11), 1179–1194.
5. Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). Oxford: Oxford University Press.
6. Torrance, G. (1986). Measurement of health-state utilities for economic appraisal: A review. *Journal of Health Economics, 5,* 1–30.
7. Gold, M. R., Siegel, J. E., & Russell, L. B. (Eds.). (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.
8. Neumann, Peter J., Sanders, Gillian D., Russell, Louise B., Siegel, Joanna E., & Ganiats, Theodore G. (Eds.). (2016). *Cost-Effectiveness in Health and Medicine* (2nd ed.). New York: Oxford University Press.
9. Weinstein, M. C., & Stason, W. B. (1977). Foundations of cost-effectiveness analysis for health and medical practice. *New England Journal of Medicine, 296,* 716–721.
10. Brazier, J., Ratcliff, J., Salomon, J. A., & Tsuchiya, A. (2007). *Measuring and valuing health benefits for economic evaluation*. Oxford: Oxford University Press.
11. Mitchell, A. S., & Viney, R. (2010). Meeting the information needs of a national drug payer: Aspirations of the guidelines from Australia. *Drug Development Research, 71*(8), 463–469.
12. Mittmann, N., Evans, W. K., Rocchi, A., Longo, C. J., Au, H.-J., Husereau, D., et al. (2009). *Addendum to CADTH's guidelines for the economic evaluation of health technologies: Specific*

guidance for oncology products. Ottawa: Canadian Agency for Drugs and Technologies. in Health
13. National Institute for Health and Clinical Excellence. (2013). *Guide to the methods of technology appraisal*. London: NICE.
14. Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., et al. (2013). Constructing experimental designs for discrete-choice experiments: Report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health, 16*(1), 3–13.
15. Riley, W. T., Rothrock, N., Bruce, B., Christodolou, C., Cook, K., Hahn, E. A., et al. (2010). Patient-reported outcomes measurement information system (PROMIS) domain names and definitions revisions: Further evaluation of content validity in IRT-derived item banks. *Quality of Life Research, 19*(9), 1311–1321.
16. Hanmer, Janel, Feeny, David, Fischhoff, Baruch, Hays, Ron D., Hess, Rachel, Pilkonis, Paul A., et al. (2015). The PROMIS of QALYs. *Health and Quality of Life Outcomes, 13*(1), 1.
17. Brooks, R., Rabin, R., & de Charro, F. (2003). *The measurement and valuation of health status using EQ-5D: A European perspective*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
18. Feeny, D., Furlong, W., Torrance, G. W., et al. (2002). Multiattribute and singleattribute utility functions for the health utilities index mark 3 system. *Medical Care, 40,* 113–128.
19. Feeny, D., Torrance, G., & Furlong, W. (1996). Health Utilities Index. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials*. Philadelphia, PA: Lippincott-Raven Press.
20. Kaplan, R. M., Sieber, W. J., & Ganiats, T. G. (1997). The quality of well-being scale: Comparison of the interviewer-administered version with a self-administered questionnaire. *Psychology and Health, 12*(6), 783–791.
21. Brazier, J. E., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care, 42,* 851–859.
22. Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics, 21,* 271–292.
23. Collins, F. S., & Riley, W. T. (2016). NIH's transformative opportunities for the behavioral and social sciences. *Science Translational Medicine, 23*(8), 366.
24. Alonso, J., Bartlett, S. J., Rose, R., Aaronson, N. K., Chaplin, J. E., Efficace, F., et al. (2013). The case for an international patient-reported outcomes measurement information system (PROMIS(R)) initiative. *Health Quality Life Outcomes, 11,* 210.
25. Hays, R.D., Revicki, D.A., Feeny, D., Fayers, P., Spritzer, K.L., Cella, D. (2016). Using linear equating to map PROMIS global health items and the PROMIS-29 V2.0-profile measure to the Health Utilities Index Mark 3. Pharmacoeconomics (ePub).
26. Revicki, D. A., Kawata, A. K., Harnam, N., Chen, W. H., Hays, R. D., & Cella, D. (2009). 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. *Quality of Life Research, 18*(6), 783–791.
27. Craig, B. M., Reeve, B. B., Brown, P. M., Cella, D., Lipscomb, J., Pickard, S., et al. (2014). US valuation of health outcomes measured using the PROMIS-29. *Value in Health, 17*(8), 846–853.