



Published in final edited form as:

*Med Care*. 2015 October ; 53(10): 894–900. doi:10.1097/MLR.0000000000000413.

## Confirmatory Factor Analysis of the Patient-Reported Outcomes Measurement Information System (PROMIS) Adult Domain Framework Using Item Response Theory Scores

**Adam Christopher Carle, MA, PhD [Associate Professor of Pediatrics]**

Cincinnati Children's Hospital Medical Center, College of Medicine, Associate Professor of Psychology, College of Arts and Sciences, University of Cincinnati

**William Riley, PhD [Chief, Science of Research and Technology Branch]**

Science of Research and Technology Branch, Behavioral Research Program, National Cancer Institute, National Institutes of Health, william.riley@nih.gov, 240-276-6973

**Ron D. Hays, PhD [Professor of Medicine]**

UCLA Department of Medicine, Division of General Internal Medicine & Health Services Research, 310-794-2294, drhays@ucla.edu

**David Cella, PhD [Chair, Department of Medical Social Sciences]**

Department of Medical Social Sciences, Director, Center for Patient-Centered Outcomes - Institute for Public Health and Medicine, Northwestern University, 312-503-1086, d-cella@northwestern.edu

### Abstract

**Background**—To guide measure development, National Institutes of Health (NIH)-supported Patient Reported Outcomes Measurement Information System (PROMIS®) investigators developed a hierarchical domain framework. The framework specifies health domains at multiple levels. The initial PROMIS domain framework specified that physical function and symptoms such as pain and fatigue indicate Physical Health (PH); depression, anxiety, and anger indicate Mental Health (MH); and social role performance and social satisfaction indicate Social Health (SH). We used confirmatory factor analyses (CFA) to evaluate the fit of the hypothesized framework to data collected from a large sample.

**Methods**—We used data ( $n = 14,098$ ) from PROMIS's wave 1 field test and estimated domain scores using the PROMIS item response theory parameters. We then used CFA to test whether the domains corresponded to the PROMIS domain framework as expected.

**Results**—A model corresponding to the domain framework did not provide ideal fit (RMSEA=0.13; CFI=0.92; TLI=0.88; SRMR=0.09). Based on modification indices and EFA, we allowed Fatigue to load on both PH and MH. This model fit the data **acceptably** (RMSEA=0.08; CFI=0.97; TLI=0.96; SRMR=0.03).

**Discussion**—Our findings generally support the PROMIS domain framework. Allowing fatigue to load on both PH and MH improved fit considerably.

## Keywords

Construct Validity; Conceptual Framework; Domain Definition; Confirmatory Factor Analysis; Item Response Theory; Patient-Reported Outcomes Measurement Information System; PROMIS

---

## Introduction

The Patient Reported Outcomes Measurement Information System (PROMIS<sup>®</sup>) is an NIH-supported project to advance the science of patient-reported outcomes (PROs). PROMIS consists of a network of collaborative researchers who seek to develop flexible and dynamic PROs applicable across a wide array of disease groups. To guide measure development, PROMIS investigators developed a domain framework based on literature reviews, analyses of archival data, and a modified Delphi procedure that included PRO measurement experts.<sup>1</sup> The resulting framework followed the World Health Organization's (WHO's) "tripartite model," which defines health as consisting of physical, mental, and social aspects of health.<sup>2</sup>

The PROMIS framework (see Figure 1) specifies that the three correlated aspects of self-reported health are comprised of domains. Within each aspect of self-reported health, Figure 1 further groups the domains according to whether they were tested in PROMIS Wave I or II. From Figure 1, one can see that physical function, fatigue, pain interference, and pain behavior correspond to Physical Health (PH); anger, anxiety, and depression correspond to Mental Health (MH); and social role performance and social role satisfaction both correspond to Social Health (SH). In this way, the PROMIS domain framework provides both a way of organizing measures and posits a theoretically testable model of health (see Riley, et al.<sup>3</sup> and Cella, et al.,<sup>4</sup> for a more detailed presentation).

PROMIS uses sets of items called item banks to measure each domain. The framework provided guidance on which domains to focus on for initial development, defined the boundaries of what PROMIS intended to measure, and provided an empirically testable conceptual model of self-reported health. Given its primary role in guiding PROMIS, an empirical test of the domain framework's structure is important. If empirical analyses do not support the framework, this might indicate that the framework does not capture some important domains. It might also indicate that PROMIS has inappropriately separated specific aspects and/or domains. Alternatively, support for the framework would support its continued use as a guide and tool for understanding the multidimensional nature of health more generally. However, to date, research has not examined whether empirical data support this framework. In this paper, we address this. We used confirmatory factor analyses (CFA)<sup>5</sup> to test whether the data collected during the PROMIS Wave 1 field test (2006-2007)<sup>4</sup> corresponds to the theoretical expectations generated by the initial framework. Data including Wave 2 domains were not available. Thus, we examined the PROMIS domain framework using Wave 1 data as illustrated in the Wave 1 component of Figure 1.

## Methods

### Overview

We used a subsample of data ( $n=14,098$ ) from the PROMIS Wave 1 testing sample.<sup>4</sup> The PROMIS Wave 1 data were collected to achieve several goals: (1) obtain item calibrations for each domains' items; (2) estimate profile scores for various disease populations; (3) create linking metrics to legacy questionnaires; (4) confirm the factor structure of the domains; and (5) conduct item and bank analyses. We selected the subset of participants for whom we were able to generate scores on the health domains (see details below). For each participant, we used item response theory (IRT) to estimate a score on each of the 9 domains included in the Wave 1 field test and measured by PROMIS at that point in time (2007): Physical Function, Fatigue, Pain Interference, Pain Behaviors, Depression, Anxiety, Anger, Social Role Performance, and Social Role Satisfaction. We then used CFA to test whether these domains corresponded to the PROMIS domain framework as expected.

### Sample

The Wave 1 sample included 21,133 adult respondents from the general US population. PROMIS investigators recruited 1,532 from PROMIS network sites. 19,601 came from YouGovPolimetrix's online panel.<sup>4</sup> All respondents completed the PROMIS measures online, though no respondent was administered all items because the total number of items evaluated in Wave 1 data collection exceeded 1,000. To reduce respondent burden, the sampling design used both a "full-bank" and "block" administration approach. Full and block administration provided data for dimensionality evaluation and IRT calibration (full administration) and data for examining associations among item banks (block administration). A subset of individuals answered all items for each of two domain item banks ( $n = 7,005$ ), but did not answer questions from any other domain. Another subset ( $n = 14,128$ ) received one of 16 blocks. Individuals in a given block received the same set of 7 items selected from each of the candidate domain item banks. Thus, participants in the block subsample answered questions about each domain, but individuals in different blocks answered different sets of questions. Our analyses included all of the individuals in the block subsample. Because we could not estimate domain scores for 30 individuals (see below), our final sample included 14,098 individuals. Cella, et al.,<sup>4</sup> describe the entire procedure and sample in detail.

### Measures

The sampling design described above limited our ability to do item level analyses. As noted, although individuals responded to questions measuring each domain, individuals responded to different sets of questions. Because the ratio of estimated parameters to number of individuals was too small to estimate an item level model for any specific block of participants and because participants did not answer the same sets of items across blocks, we could not conduct item level analyses. Instead, we used PROMIS calibrated IRT parameters and estimated individuals' scores for each domain, and used these scores in our CFA.

PROMIS has calibrated item parameters available for each of the domains' item banks: <http://www.nihpromis.org/software/assessmentcenter>. Extensive psychometric analyses,

including dimensionality and IRT fit evaluation, have been performed for each domain item bank, and each has met the standards adopted by PROMIS.<sup>6</sup> Like all IRT models,<sup>7</sup> one can use the calibrated parameters and an individual's responses to any given set of domain items to estimate an IRT score (i.e., a "theta" score). In this way, one can estimate an IRT score for a domain that is on the same standardized metric (and thus directly comparable across individuals) regardless of which items in a domain's bank individuals have answered. We used Mplus (version 7)<sup>8</sup> to estimate the IRT scores. We ran models in Mplus for each block of individuals, fixed the parameters for each item to their PROMIS value, and output IRT scores for the individuals in each block. This allowed us to estimate a score for any individual who responded to at least one item that had a calibrated parameter (**see the section on Missing Data below**). However, not all items tested during Wave 1 had acceptable psychometric properties (**e.g., local independence**), thus not all items fielded during Wave 1 have calibrated IRT parameters associated with them.<sup>4</sup> Thus, we could only estimate IRT scores for individuals who answered at least one of the items that had calibrated parameters associated with it. Thus, our final sample included 14,098 individuals for whom we estimated a score on at least one of the domains. From among these 14,098 individuals, 8,659 had a score on all of the domains.

Finally, we multiplied scores estimated from the PROMIS IRT parameters for Fatigue, Pain Interference, Pain Behaviors, Depression, Anxiety, and Anger by negative one so that higher scores represented better health. We made this change to align the meaning of high scores across all domains (i.e., high scores indicate better health). Other than reversing the meaning of low and high scores, this transformation did not otherwise change the distribution of the scores. Note that this type of transformation has no influence on CFA fit indices.<sup>5</sup>

### Confirmatory factor analyses

We tested the hypothesized domain framework using each individuals' estimated IRT score and CFA.<sup>5</sup> All analyses used Mplus (version 7). The hypothesized domain framework posited a three factor model with Physical Function, Fatigue, Pain Interference, and Pain Behaviors measuring PH; Depression, Anxiety, and Anger measuring MH; and Social Role Performance, and Social Role Satisfaction measuring SH. The framework also hypothesized that MH, PH, and SH correlate. We evaluated fit using fit index levels identified in the literature.<sup>9,10</sup> These included the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and the Tucker Lewis Index (TLI). For the RMSEA and SRMR, we considered values  $\leq 0.05$  ideal and values  $\leq 0.08$  acceptable. For the CFI and TLI, we considered values  $\geq 0.95$  ideal and  $\geq 0.90$  acceptable. Fit evaluation focused on the index set.

In addition to the hypothesized framework, a priori we also planned to test a first-order single factor model. This first-order single factor model specified that all 9 domains included in our analyses loaded on a single "General Health" factor. Though substantively appealing, we did not test a higher-order model built on the domain framework and hypothesized that a single higher-order "General Health" factor accounted for the hypothesized relationships among the PH, MH, and SH factors. This is because the "General

Health" factor would be just identified given only 3 first order factors.<sup>5</sup> As a result, the fit of this model would equal the first order model's fit.

Finally, we split our sample into random halves for our CFA analyses. We did this given the potential that neither the hypothesized model nor the unidimensional model would demonstrate acceptable fit and we would need to develop an alternative model. The second split half served as a validation sample. A priori, we planned to use the model's modification indices (MIs) and expected parameter change indices (EPCs) to identify sources of misfit. MIs give the expected change in the  $\chi^2$  when freeing a constraint (e.g., freely estimating a factor loading previously fixed to zero). To avoid relaxing constraints inconsistent with theory, we used EPCs. EPCs give the expected change in a given parameter when freeing the constraint associated with the parameter.<sup>11,12</sup> We used them to avoid relaxing constraints that would lead to theoretically inconsistent estimates. In addition, we used exploratory factor analyses (EFA) to identify a potential alternative model.

### Missing Data

We **performed** the CFAs in the first split half twice. First, we conducted analyses among individuals with an estimated IRT score on all of the domains ( $n=4,821$ ). Second, we repeated our analyses using Mplus' missing data function,<sup>8</sup> and included all individuals with a score on at least one of the domains ( $n=6,988$ ). Mplus does not impute individual level responses, but instead uses all available data to estimate the model using full information maximum likelihood. This approach assumes data are either missing completely at random or missing at random. Given that scores were missing simply due to no calibrated item parameters existing for the items an individual answered and that individuals were randomly assigned to receive different blocks of items, we felt it was reasonable to assume the data were missing at random. However, we conducted analyses both ways to examine whether the results differed for the listwise deletion and imputation approaches. The results did not differ so we report only the results based on the larger sample that utilized a missing data approach. We will provide interested readers with full results upon request.

### Results

The PROMIS network and online panel samples did not differ significantly in their mean age. However, the network sample had significantly more males and was significantly less diverse, better educated, and wealthier (see Table 1). Table 2 gives the means, standard deviations, and correlations for the domain scores. The single factor model did not provide ideal fit (RMSEA = 0.20; SRMR = 0.09; CFI = 0.77; TLI = 0.70;  $\chi^2 = \chi^2 7,904.11$ ,  $df=27$ ,  $p<0.01$ ; Normed  $\chi^2 = 1.14$ ). The majority of fit indices also indicated that the hypothesized three factor model did not demonstrate ideal fit (RMSEA = 0.13; SRMR = 0.06; CFI = 0.92; TLI = 0.89;  $\chi^2 = \chi^2 2,679.02$ ,  $df=24$ ,  $p<0.01$ ; Normed  $\chi^2 = 0.38$ ). Thus, we sought to develop an alternative model.

First, a review of the MIs suggested allowing Fatigue to load on both PH and MH (MI = 3,345.72;  $df=1$ ;  $p < 0.01$ ). In addition, the EPC (0.54) indicated a change consistent with theory. The MI constraining Fatigue to load only on PH was dramatically larger than all other MIs (which were nearly all  $< 100$ ). Given our desire to seek parsimony, we

hypothesized this single modification (allowing Fatigue to cross-load on PH and MH) might result in appropriate fit. We also conducted 2 and 3 factor EFAs, each with an oblique rotation and each allowing all items to cross-load. The two factor EFA did not fit acceptably (RMSEA=0.14; CFI=0.93; TLI=0.86; SRMR=0.05). The three factor model fit very well (RMSEA=0.04; CFI>0.99; TLI=0.99; SRMR<0.01). The three factors corresponded to our three hypothesized factors. Physical Function, Fatigue, Pain Interference, and Pain Behaviors loaded most highly on one factor. Depression, Anxiety, and Anger loaded most highly on the second. Social Role Performance and Social Role Satisfaction load most highly on the third. For all but Fatigue, the cross-loadings were < 0.1 and near zero. Fatigue had two relatively equal sized loadings on MH and PH (~0.4) and a smaller loading on SH (~0.2).

As such, we hypothesized that a model consistent with hypothesized model that also allowed Fatigue to cross-load on PH and MH would provide a suitable and theoretically consistent alternative (see Figure 2). We tested this model in the second split half. The single modification resulted in a model that met our criteria for acceptable fit (RMSEA = 0.08; SRMR = 0.03; CFI = 0.97; TLI = 0.96;  $\chi^2 = 950.343$ ,  $df=23$ ,  $p<0.01$ ; Normed  $\chi^2 = 0.13$ ). Thus, we considered this our final model. Table 3 presents this model's standardized parameters.

## Discussion

In this study, we set out to test PROMIS's hypothesized domain framework (see Figure 1).<sup>3</sup> We tested the hypothesis that the physical function, fatigue, pain interference, and pain behavior domains measure PH; anger, anxiety, and depression domains measure MH; and the social role performance and social role satisfaction domains measure SH. Our results suggest that fatigue represents both PH and MH. With this modification, the results support the hypothesized PROMIS domain framework. Fatigue's dual loading aligns with the SF-36's vitality subscale loadings and the factor scoring coefficients used to estimate the physical and mental health summary scores.<sup>13</sup>

The domain framework follows WHO's definition of health that describes three aspects of health: Physical, Mental, and Social (i.e., the tripartite model of health).<sup>2</sup> Though widely accepted,<sup>14</sup> some previous research failed to find sufficient evidence for a social dimension.<sup>15-17</sup> The earlier work suggested two rather than three health aspects. Our findings support the tripartite model, bolstering PROMIS's decision to adopt the WHO tripartite model. This supports the continued use of the PROMIS domain framework to understand how developed PROMIS item banks relate to each other and how they relate to higher order aspects of health. The results also support the continued use of the domain framework to develop new item banks measuring other domains not yet assessed by PROMIS network measures.

To our knowledge, this is the first time that an investigation using a large sample of diverse individuals has found empirical support for the tripartite model that specifies SH as a distinct aspect from PH and MH. We believe this occurred because our measures of SH domains resulted from PROMIS investigators' efforts to specifically measure SH domains.

Thus, they produced and tested a large, targeted, and substantial set of SH questions. Previous work used items that measured social features of physical and mental health (e.g., interference with social activities because of physical or mental health) and substantially fewer items (e.g., 2). These differences likely explain why earlier work did not find evidence for a SH aspect.

Although our results support the continued use of the domain framework as a guiding framework for PROMIS, we emphasize that this framework is not intended to serve as a classification structure like the International Classification of Diseases (ICD) or Functioning (ICF)<sup>18</sup> or the American Psychiatric Association's Diagnostic and Statistical Manual-IV.<sup>19</sup> Use of the PROMIS domain framework assures that measures developed for inclusion in PROMIS address the core aspects of health that apply to a diverse clinical research community. However, PROMIS investigators developed the framework with consideration to the ICF<sup>20</sup> and the DSM-V field trials included selected PROMIS short forms as cross-cutting dimensional measures.<sup>21</sup> Thus, while the framework should not replace the ICF or DSM systems, it can complement them.

Relatedly, the domain framework may serve as a useful starting point for developing a PRO conceptual framework when submitting an application for new treatments to regulatory authorities such as the United States Food and Drug Administration (FDA).<sup>22</sup> FDA guidance indicates that a conceptual framework should provide context-specific relationships among the various measures in a clinical trial. PROMIS measures offer a generic starting point for several common symptoms and functional status concerns relevant to this guidance.<sup>23</sup> In addition, the procedures used to develop the PROMIS domain framework align in many ways with those used to develop disease-specific measures.<sup>24</sup> Likewise, the PROMIS domain framework does provide a nested listing of possible health domains one might consider when developing a conceptual framework for FDA submission purposes.

Finally, a single factor model that specified a General Health dimension measured by physical function, fatigue, pain interference, pain behavior, anger, anxiety, depression, social role performance, and social role satisfaction, did not fit the data well. This suggests that a single score to represent self-reported health may necessarily be more complex than a simple summation of the diverse domains. However, the relatively high correlations among the three factors do indicate that individuals tend to report similar health across the MH, PH, and SH aspects. Though PROMIS has not yet determined an accepted method of estimating a single score, our data do not rule out the possibility of a higher order General Health score.

## Limitations

Before concluding, we note some study limitations. First, given the Wave 1 sampling design, we could not conduct item level analyses. Second, we did not examine measurement invariance for our model across different race/ethnic groups. Third, we had to exclude some participants from our analyses because calibrated item parameters did not exist for the items they answered. However, given participants' random assignment to blocks, this is unlikely to have biased our results. Fourth, PROMIS has not yet developed item banks measuring all of the framework's potential domains. The possibility exists that including more domains might result in different conclusions. Relatedly, SH only had two indicators in the model,

limiting the extent to which our findings broadly support a SH construct. Future research in which a substantially larger sample of individuals all answer calibrated items for all of the domains eventually to be measured by PROMIS can address these issues.

## Conclusion

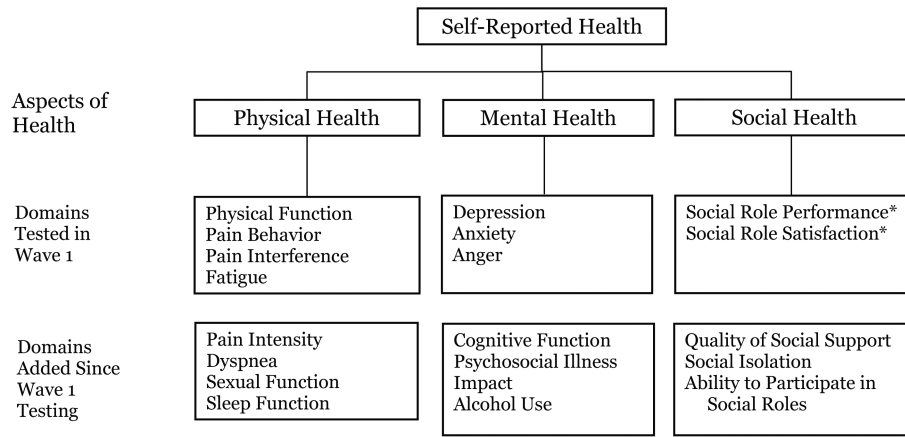
Using PROMIS Wave 1 data,<sup>4</sup> we found evidence supporting the initial PROMIS domain framework. Specifically, the physical function, fatigue, pain interference, and pain behavior domains measured by PROMIS item banks measure PH; the fatigue, anger, anxiety, and depression domains measured by PROMIS item banks measure MH; and the social role performance and social role satisfaction domains measured by PROMIS item banks measure SH. Other than finding that fatigue appears to measure both PH and MH, our findings do not diverge from the hypothesized domain framework. This constitutes the first large-scale demonstration of the validity of a tripartite model of health that specifies SH as a separate aspect along with the more traditionally included PH and MH. As the PROMIS network continues to develop measures corresponding to currently unmeasured (by PROMIS) domains, additional analyses will need to continue to evaluate the placement of these domains within the framework.

## References

1. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*. 2007; 45(5 Suppl 1):S3. [PubMed: 17443116]
2. Organization WH. Constitution of the World Health Organization basic document. 1948
3. Riley WT, Rothrock N, Bruce B, et al. 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*. 2010; 19(9):1311–1321. [PubMed: 20593306]
4. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of clinical epidemiology*. 2010; 63(11):1179–1194. [PubMed: 20685078]
5. Bollen, K. *Structural equations with latent variables*. Wiley; New York, NY: 1989.
6. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical care*. 2007; 45(5 Suppl 1):S22–31. [PubMed: 17443115]
7. Embretson, S.; Reise, S. *Item response theory for psychologists*. Lawrence Erlbaum; Mahwah, N.J.: 2000.
8. Muthén, LK.; Muthén, BO. *Mplus User's Guide*. Seventh Edition. Muthén & Muthén; Los Angeles, CA: 1998-2012.
9. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. 1999; 6(1):1–55.
10. Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*. 1998; 3(4):424–453.
11. Byrne, BM. *Structural Equation Modeling with Mplus : Basic Concepts, Applications, and Programming*. Routledge; New York, NY: 2012.
12. Bentler, PM. *EQS 6 Structural Equation Program Manual*. Multivariate Software; Encino, CA: 2005.
13. Farivar SS, Cunningham WE, Hays RD. Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. Health and quality of life outcomes. 2007; 5:54. *Journal Article*. [PubMed: 17825096]

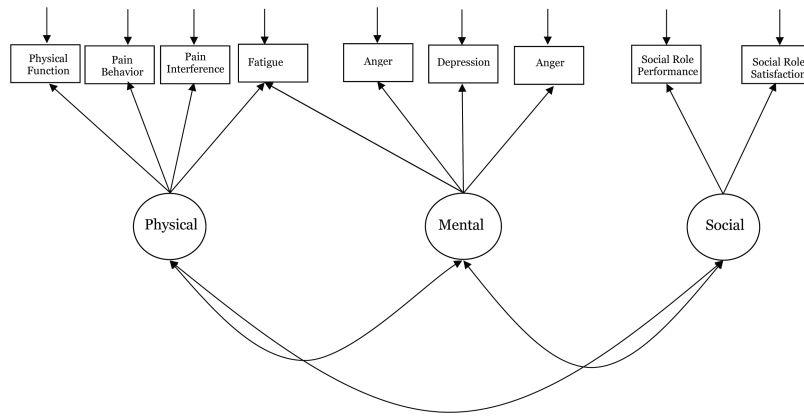


14. Larson JS. The conceptualization of health. *Medical Care Research and Review*. 1999; 56(2):123–136. [PubMed: 10373720]
15. Ware JE Jr, Brook RH, Davies AR, Lohr KN. Choosing measures of health status for individuals in general populations. *American Journal of public health*. 1981; 71(6):620–625. [PubMed: 7235100]
16. Hays RD, Stewart AL. The structure of self-reported health in chronic disease patients. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*. 1990; 2(1):22.
17. Chang C-H, Wright BD, Cella D, Hays RD. The SF-36 physical and mental health factors were confirmed in cancer and HIV/AIDS patients. *Journal of clinical epidemiology*. 2007; 60(1):68–72. [PubMed: 17161756]
18. Organization WH. International classification of functioning disability and health (ICF). 2001
19. APA. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. American Psychiatric Association; Washington, DC: 1994.
20. Tucker, C.; Riley, A.; Lai, J.; Cella, D.; Riley, W.; Forrest, C. *Conceptual Frameworks and Synergies for Measurement: PROMIS and ICF*. WHO ICF Meeting. Toronto, CA: 2010.
21. Narrow WE, Kuhl EA. Dimensional approaches to psychiatric diagnosis in DSM-5. *The journal of mental health policy and economics*. 2011; 14(4):197. [PubMed: 22345361]
22. Rothman ML, Beltran P, Cappelleri JC, Lipscomb J, Teschendorf B. Patient-Reported Outcomes: Conceptual Issues. *Value in Health*. 2007; 10(s2):S66–S75. [PubMed: 17995476]
23. Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: Use of medical product development to support labeling claims: draft guidance. 2009. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>
24. Turner RR, Quittner AL, Parasuraman BM, Kallich JD, Cleeland CS. Patient-reported outcomes: instrument development and selection issues. *Value in Health*. 2007; 10:S86–S93. [PubMed: 17995478]



\*Subsequent to Wave 1, Social Role Performance and Social Role Satisfaction were merged into a single domain named "Satisfaction with Social Roles and Activities"

**Figure 1.**  
PROMIS Domain Framework – Adult Banks.



**Figure 2.** Depiction of the Final Confirmatory Factor Analytic Model of the Patient-Reported Outcomes Measurement Information System Domain Framework.

**Table 1**

Descriptive Statistics for the Analytic Sample.

<u>Variables</u>	<u>Categories</u>	<u>Frequency</u>	<u>%</u>
Source	YouGovPolimetrix	12,925	91.68
	PROMIS	1,173	8.32
Race/Ethnicity	White Non-Hispanic	11,081	78.6
	Black Non-Hispanic	1,130	8.02
	Hispanic	1,174	8.33
	Other Non-Hispanic	503	3.57
	Multi Non-Hispanic	185	1.31
	Missing	25	0.18
	Gender	Male	6,762
	Female	7,335	52.03
	Missing	1	0.01
Education	<High School Graduate	390	2.77
	High School Graduate	2,111	15
	> High School Graduate	11,587	82
	Missing	10	0.07
Household Income	\$20,000	1,661	11.78
	>\$20,000 \$50,000	4,497	31.9
	>\$50,000 \$100,000	4,856	34.44
	>\$100,000	2,597	18.42
	Missing	487	3.45
	<u>Minimum</u>	<u>Maximum</u>	<u>Median</u>
Age	18	89	56

**Table 2**

Domain Means (first column), Standard Deviations (diagonal elements), and Correlations (off diagonal elements).

	<u>Means</u>	Anger	Anxiety	Depression	Fatigue	Physical Function	Pain Behavior	Social Role Satisfaction	Social Role Performance
Anger	-0.18	0.87	-	-	-	-	-	-	-
Anxiety	-0.21	0.59	0.84	-	-	-	-	-	-
Depression	-0.22	0.61	0.72	0.82	-	-	-	-	-
Fatigue	-0.17	0.46	0.62	0.62	0.91	-	-	-	-
Physical Function	-0.28	0.18	0.30	0.29	0.51	0.93	-	-	-
Pain Interference	-0.38	0.32	0.46	0.44	0.66	0.62	-	-	-
Pain Behavior	-0.23	0.31	0.39	0.38	0.56	0.53	0.90	-	-
Social Role Satisfaction	0.06	0.35	0.47	0.51	0.58	0.38	0.37	0.83	-
Social Role Performance	-0.08	0.36	0.48	0.52	0.65	0.47	0.42	0.74	0.86

**Table 3**

Standardized Parameters from the Final Confirmatory Factor Analytic Model of the Patient-Reported Outcomes Measurement Information System Domain Framework (standard errors in parentheses).

	<b>Loadings</b>	<b>Intercepts</b>	<b>Uniquenesses</b>
<i>Mental Health</i>			
Anger	0.68 (0.01)	-0.21 (0.01)	0.54 (0.01)
Anxiety	0.84 (0.01)	-0.26 (0.01)	0.29 (0.01)
Depression	0.86 (0.01)	-0.27 (0.01)	0.26 (0.01)
Fatigue	0.47 (0.01)	-	-
<i>Physical Health</i>			
Physical Function	0.68 (0.01)	-0.32 (0.01)	0.54 (0.01)
Pain Interference	0.93 (0.00)	-0.52 (0.01)	0.14 (0.01)
Pain Behavior	0.81 (0.01)	-0.27 (0.01)	0.35 (0.01)
Fatigue	0.46 (0.01)	-0.19 (0.01)	0.33 (0.01)
<i>Social Health</i>			
Social Role Satisfaction	0.83 (0.01)	0.07 (0.01)	0.31 (0.01)
Social Role Performance	0.90 (0.01)	-0.07 (0.01)	0.19 (0.01)
<b>Factor Correlations</b>			
	Mental Health	Physical Health	Social Health
Mental Health	1	-	-
Physical Health	0.57 (0.00)	1	-
Social Health	0.68 (0.01)	0.67 (0.01)	1