The Patient Reported Outcomes Measurement Information System (PROMIS): *A Walk Through the First Four Years*

TABLE OF CONTENTS

I. Overview

II. PROMIS Domains

A. Domain framework
B. Domain definitions

III. Item Bank Development

A. Literature review for items
B. Archival data analysis
C. Item classification
D. Qualitative item review
E. PROMIS preliminary item pools

IV. Item Bank Testing

A. Mode of Testing Study to inform item administration format
B. Wave I testing
C. Wave II testing

V. Item Banks and Tools

A. Version 1.0 Item Banks
B. PROMIS User Manual Version 1.1
C. Assessment Center
D. Assessment Center User Manual
E. CAT Demonstration

VI. Publications and Presentations

VII. References

I. Overview

In late 2004, a group of outcomes scientists from seven institutions and the National Institutes of Health (NIH) formed a cooperative group funded under the NIH Roadmap for Medical Research Initiative to re-engineer the clinical research enterprise (http://www.nihroadmap.nih.gov). This initiative - the Patient-Reported Outcomes Measurement Information System (PROMIS) - aims to revolutionize the way patient-reported outcome tools are selected and employed in clinical research and practice evaluation. It will also establish a national resource for accurate and efficient measurement of patient-reported symptoms and other health outcomes in clinical practice. The NIH Roadmap is a series of far-reaching initiatives designed to transform the nation's medical research capabilities and speed the movement of research discoveries from the bench to the bedside. It provides a framework of the priorities NIH must address to optimize its entire research portfolio, and it lays out a vision for a more efficient and productive system of medical research. For more information about the NIH Roadmap, please visit the Web site at http://www.nihroadmap.nih.gov. For more information about PROMIS, visit http://www.nihpromis.org.

II. PROMIS Domains

A. Domain framework

The first task of the PROMIS network was to create a protocol for developing a domain map (framework) that portrayed the structure of each target domain and its conceptual framework or, where applicable, hierarchical structure. Existing outcome assessment questionnaires use an explicit or implicit framework that typically includes the concepts of physical function or limitation, mental health or distress, and social function, with many also including symptoms (eg, fatigue, pain). The Steering Committee (SC)-approved protocol for the domain mapping activity specified that the preliminary PROMIS framework would be developed through independent literature reviews by the Stanford PRS, the Pittsburgh PRS, and the Statistical Coordinating Center (SCC), followed by a consensus-building Delphi process and statistical analysis of available data regarding dimensionality of health status assessment. Early in the first year, the SC endorsed the World Health Organization (WHO) physical, mental, and social health framework. Other organizing frameworks were considered, such as the WHO International Classification of Functioning and a 2-factor model of physical and mental health. However, after discussion and careful consideration, the SC opted to retain the WHO tripartite framework as compelling and sufficiently broad and inclusive to enable important social dimensions of health to be developed further than has been done to date. After achieving consensus on the broad WHO framework, the SC launched the Domain Mapping Protocol. Under this protocol, PROMIS network investigators used a modified Delphi approach combined with quantitative analysis of existing relevant data, to inform multiple rounds of framework review and revision until consensus was reached on a detailed articulation of subordinate domains beneath the broad physical, mental, and social headings. The PROMIS SC then resolved to begin item development and testing in pain, fatigue and at least one subdomain in each of physical, mental and social health.

See the Domain Framework: http://www.nihpromis.org/measures/domainframework.

For more detail refer to the following manuscript:

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J. F., Bruce, B., Matthias, R., & on behalf of the PROMIS cooperative group. (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. (PMID: 17443116)

B. Domain definitions

Domain definitions were created for physical function, fatigue, pain, emotional distress (including depression, anxiety, and anger), social health (including social function and social support), and global health.

Briefly, **Physical function** is defined as one's ability to carry out various activities that require physical capability, ranging from self-care (activities of daily living) to more vigorous activities that require increasing degrees of mobility, strength, or endurance (Haley et al, 1994a; 1994b; Stewart & Kamberg, 1992; Wilson & Cleary, 1995). **Fatigue** is defined as an overwhelming, debilitating, and sustained sense of exhaustion that decreases one's ability to carry out daily activities, including the ability to work effectively and to function at one's usual level in family or social roles (Glaus, 1998; North American Nursing Diagnostic Association, 1996; Stewart, Hays & Ware, 1992). **Pain** is an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage (Chang, 1999; Mersky & Bogduk, 1994; Meuser et al, 2001; Sherbourne, 1992)**. Emotional distress** commonly refers to unpleasant feelings or emotions that are experienced privately and, therefore, are good candidates for assessment as patient-reported outcomes. Emotional distress is comprised typically of aspects of depression, anxiety, and anger. Symptoms specific to **depression** are those that reflect low levels of positive affect. In addition, depression is often characterized by the experience of loss and feelings of hopelessness, helplessness, and worthlessness. Symptoms that best differentiate **anxiety** are those that reflect autonomic arousal and the experience of threat. **Anger** is distinguished by attitudes of hostility and cynicism and is often associated with experiences of frustration when goal-directed behavior is impeded. In general, our PROMIS item banks emphasize the cognitive and affective components of these concepts. **Social health** is defined as perceived well-being regarding social activities and relationships, including the ability to relate to individuals, groups, communities and society as a whole. Components of social functioning include understanding and communication, getting along with people, participation in society and performance of social roles. The two broad patient-reported outcomes under social function within the PROMIS framework are social function and social support. **Social function** is defined by PROMIS as involvement in, and satisfaction with, one's usual social roles in life's situations and activities. These roles may exist in marital relationships, parental responsibilities, work responsibilities and social activities (Dijkers, Whiteneck & El-Jaroudi, 2000; McDowell & Newell, 1996). There are two broad types of **social support**: quantitative and qualitative (Cohen & Syme, 1985; Wills, 1985; Wortman & Conway, 1985). Quantitative social support refers to the existence of, and interconnections between, social ties, e.g., marital status, number of relationships, frequency of contacts with friends and relatives, church membership, and volunteer participation. Qualitative social support refers to functional aspects of supportive relationships, i.e., interpersonal relationships that serve particular functions. Measures of social support generally seek information about a person's perception of the availability or adequacy of resources provided by other

persons (Cohen & Syme, 1985). In this context, perceived social support is a subdomain of social health.

Finally, **Global health** refers to evaluations of health in general rather than specific elements of health. The global health items include global ratings of the five primary PROMIS domains (physical function, fatigue, pain, emotional distress, social health) and general health perceptions that cut across domains. Global items allow respondents to weight together different aspects of health to arrive at a 'bottom-line" indicator of their health status.

III. Item Bank Development

For its first set of item banks, PROMIS chose to focus on pain, fatigue, emotional distress, physical function, and social function. Six phases of item development are documented: *identification of existing items*, *item classification and selection*, *item review and revision*, *focus group input on domain coverage*, *cognitive interviews with individual items*, and *final revision before field testing*. Item banks were therefore developed using mixed qualitative and quantitative methods as described below.

A. Literature review for items

Drawing from decades of experience reflected in published literature and previous work of the investigators, PROMIS began the item bank development by cataloguing items from well-established instruments that had been extensively tested and had excellent track records. PROMIS investigators across the network conducted inclusive searches and evaluations of existing instruments to enrich the pool of domain-relevant items that were considered potential candidates for the PROMIS item banks. Searches started with MEDLINE and Health and Psychosocial Instruments, but also included proprietary databases like Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID). Each domain work-group constructed their own search strategy based upon the specific needs identified within the domain. For example, the Emotional Distress domain group identified 4 general areas (referred to as "subdomains") for starting bank development: depression, anxiety, anger, and substance misuse. They created search strategies to identify all known items covering these topics as a starting point for PROMIS banks. Importantly, the process allowed manual searches of files by investigators to identify items that were not found through the database searches. For example, the Statistical Coordinating Center (SCC) had accumulated databases of items in a variety of domains in which they had researched previously. The SCC made these item lists available to PROMIS researchers. At this stage of the process, items were not filtered out if they applied to a specific population; such items were kept for further qualitative analysis. By performing these searches, PROMIS investigators identified thousands of items relevant to the domains PROMIS was trying to measure. At that point, no judgment was made regarding the quality or redundancy of the items; they were only selected if their content fit or were deemed proximal to the domain definitions.

After identifying more than 7,000 existing items related to the five broad domains selected for PROMIS wave one development and testing, each item was entered into a common item library at the SCC. Item characteristics recorded in the library included: (1) **context**: the instructions associated with answering the item; (2) **stem**: the part of the item that makes it unique from others in the same scale; (3) **response**

options: response choices from which the respondent is asked to select; (4) **time frame**: (if stated) the period of time that the respondent was to consider in answering the question; and (5) **instrument of origin**. *See the initial item scope:* http://www.nihpromis.org/science/LiteratureReviewItems.

B. Archival data analysis

To maximize information already available about measuring PROs, the Statistical Coordinating Center (SCC) identified several extant data sets with self-report data on the five PROMIS core domains: physical functioning, fatigue, pain, emotional distress, and social role participation. The SCC identified and reviewed large data sets that included patient-reported outcomes and selected 11 for initial analysis. Working groups identified items representing the PROMIS domains. Items were then subjected to IRT analyses by PROMIS psychometricians. All results were reviewed collectively by the SCC analysis team and summaries were presented to the appropriate domain working group. Our primary goal was to use these archival data to better understand the dimensional structure of items that tap one of the five selected domains. Secondarily, we aimed to inform the revision of items in the item library, inform the identification of the most useful response sets, and guide new item construction in preparation for the first wave of PROMIS testing. These posted reports summarize the contribution of secondary (archival) data analyses to the production of PROMIS items for network testing. Each of the links below will connect with either the general statistical analysis plan for secondary data analyses, or one of the available domain reports. The physical function domain report is a published article (Hays et al, 2007).

See the Statistical Analysis Plan (see also Reeve et al, 2007):
http://www.nihpromis.org/documents/PROMIS%20Statistical%20Analysis%20Plan.pdf.

See the Social Health Archival Analysis Report, September, 2006:
http:www.nihpromis.org/documents/SocialHealthArchivalAnalysisReport.pdf.

See the Emotional Distress Archival Analysis Report, September 2006:
http://www.nihpromis.org/documents/EmotionalDistressArchivalAnalysisReport.pdf.

See the Pain Archival Analysis Report, September, 2006:
http://www.nihpromis.org/documents/PainArchivalAnalysisReport.pdf.

See the Fatigue Archival Analysis Report, September, 2006:
http://www.nihpromis.org/documents/FatigueArchivalAnalysisReport.pdf.

C. Item classification

Confronted with thousands of items, a method for sorting through the content and deciding on the most representative and informative items was needed. We called this method “binning and winnowing.” Items were placed into common “bins,” typically defined by the content of the stem. Next, items were “winnowed” out if they were deemed redundant or inferior to alternative items in the same bin.

Binning. The PROMIS domain workgroups first selected those items from the item library that they believed represented their domain. This process was done in teams so that at least 2 people reviewed each item for inclusion. Upon completion of domain identification, domain workgroups proceeded with the task of binning items. Binning refers to a systematic process for grouping items according to meaning and specific latent construct. For example, “walking” became a bin within the physical function domain. The final goal was to have a bin from which a small number of items could be chosen to adequately represent the bin. We did not predetermine the number of items that would adequately represent a bin. Rather, the goal for this process was to identify enough items to capture the meaning of the bin and to eliminate unnecessary redundancy in the item pool. By grouping items systematically, the domain workgroups could observe redundancy among items and identify the best potential items based on qualitative characteristics. PROMIS domain workgroups (including several investigators across the PROMIS research sites and the SCC) began by creating a set of bins based on a review of that domain’s literature, including previous factor analytic studies of domain items, and theory-based studies of the domain. This “top-down” approach began with a conceptual model of the facets of each domain. However, each domain workgroup approached the process with the flexibility to add or subtract bins based on the content of items themselves. By taking this approach, we retained the organizational structure put forth by the domain experts, but took advantage of new ideas as expressed by the items written in the clinical literature. This allowed for the most inclusive and open approach at this stage of item evaluation. The purpose of binning was to enable the identification of redundant and inferior items. Thus, what is important is that the final set of items emerging from this process adequately represents the facets of the latent trait.

Winnowing. The goal of winnowing was to reduce the large item pool down to a representative set of items. The process of winnowing helped to identify item characteristics that would include or exclude them from the PROMIS item banks based on domain definitions. Ultimately, this process was based on the judgment of reviewers and was accomplished by a consensus process of 2 or 3 reviewers for each domain. We adopted a set of criteria for excluding entire bins or items within bins because they were not applicable to the current domain activities for PROMIS. Many items excluded seem to measure important domains or subdomains that are not currently the focus of PROMIS item banks. PROMIS investigators used the following criteria to remove items from consideration: (1) item content was inconsistent with the domain definition; (2) an item was semantically redundant with a previous item; (3) the item content was too narrow to have universal applicability; (4) the stem of the item was disease specific, reducing general applicability of the item; and (5) the item was confusing. For example, items related to satisfaction with physical function were identified, binned and removed from the physical function item bank consideration, because satisfaction was not in the PROMIS definition of physical function. Across all domains, approximately 30% of the items were eliminated due to redundancy, and approximately 45% were eliminated because they did not fit within the domain definitions adopted by PROMIS investigators. Table 3 has examples of items that were eliminated and the reasons for doing so. By carefully analyzing each item and comparing them to other items within a given bin, domain workgroups were better able to apply the several criteria to each item. As with all other aspects of the QIR process, all decisions about item winnowing were reviewed by multiple members of the domain workgroup and members of the SCC to ensure a high level of consensus and to impose some standardization of processes across domain groups. The process of binning and winnowing yielded a smaller set of items that were then subjected to editing to match PROMIS stylistic conventions in a process of item revision.

D. Qualitative item review (text drawn heavily from DeWalt et al, 2007)

For detail on the qualitative item review methodology, see the *Qualitative Item Review Protocol* (http://www.nihpromis.org/Documents/PROMIS_QIR_Protocol.pdf).

Expert item review and revision was conducted by trained professionals who reviewed the wording of each item and revised as appropriate for readability, adhering to conventions adopted by the PROMIS network. **Focus groups** were used to inform domain definitions, and to identify new areas of item development for current and future PROMIS item banks. **Cognitive interviews** were used to examine individual items, specifically their comprehensibility. Items successfully screened through this process were sent to field testing and were subjected to innovative scale construction procedures (DeWalt et al, 2007).

There are more than 100 documents that summarize the qualitative research conducted in support of developing the item pool for the first wave of PROMIS testing. These documents describe the expert item reviews of content, readability, translatability, and status with regard to other comparable items in the same domain facet. They also summarize themes uncovered in the many focus groups held concurrent with item review, and results of individual cognitive interviews in which respondents were asked to "think aloud" regarding the content of each individual item selected for the initial item pool. These documents can be accessed by linking to the following: *see summary data by domain* (http://www.nihpromis.org/science/qualitativeitemreview).

Item Review Process. While being reviewed for fidelity to content, clarity, and readability, candidate items were reviewed by experts trained to ensure acceptability to respondents with low literacy, and for language translatability. Item surface characteristics (e.g., intensity/difficulty, frequency, point recall versus interval, specific event or experience, association between two or more experiences, interference of one experience upon other, multi-barreled or confusing wording) were assigned to each item. Each unit of input into the review process, both from respondents and from experts, was considered as a basis for revision to item wording to context, stem or response options. After items were revised as recommended from experts and respondents from selected focus groups, each proposed item was administered to a minimum of five (5) naïve individuals for under the cognitive interviewing component of the *QIR Protocol* (http://www.nihpromis.org/Documents/PROMIS_QIR_Protocol.pdf).

Results of this entire item selection, modification, testing and finalization are summarized in DeWalt et al (2007) and detailed in the 109 individual reports. (*See Summary data by domain:* http://www.nihpromis.org/science/qualitativeitemreview.)

E. PROMIS preliminary item pools

The Domain Mapping, Archival Analysis, and Qualitative Item Review activity collectively produced a 2006 version of PROMIS preliminary items for wave one testing. These 2006 Preliminary PROMIS items were included in Wave 1 PROMIS testing and are *not* the final calibrated item banks. PROMIS version 1.0 items, short forms, calibrated item banks, and CATs are now available through *Assessment Center* (http://www.assessmentcenter.net/ac1).

IV. Item Bank Testing

A. Mode of Testing Study to inform item administration format

In designing the testing platform there was uncertainty about (1) whether or not a participant should automatically be forwarded to the next item after providing an answer and (2) if a participant should be able to go back to the previous item. A study was conducted using one of the social role participation banks. The four test conditions included (1) automatic advance with a back button, (2) automatic advance without a back button, (3) use of a "next item" button after selecting an answer with a back button, and (4) use of a "next item" button after selecting an answer without a back button. Test data were analyzed to help decide on the final test platform, which auto advanced after response and enabled respondents to go back an item.

B. Wave I testing

A summary of Wave 1 testing can be found at *Wave I Testing Summary* (http://www.nihpromis.org/science/calibrationtesting).

Additional Chronic Pain Sample

In order to increase the number of participants experiencing pain in the calibration sample, individuals with chronic pain were recruited through the American Chronic Pain Association (ACPA). An invitation to complete the PROMIS pain survey was posted on the ACPA website. To be eligible, participants had to be 21 years of age or older and have at least one chronic pain condition for at least 3 months prior to participating in the survey. Those who met eligibility criteria were asked to complete an online informed consent form. After obtaining informed consent, participants immediately began the survey. The survey was posted on the website of the ACPA from September 2007 to March 2008. There were 967 participants who responded to 47 pain impact, 42 pain behavior, and 41 pain quality items, and one global average pain intensity item through online administration. The average age was 48.2 years (SD = 11.1). Eighty-one percent of the respondents were female, and 91% were Caucasian, 1.5% were African-American, and 5% were of Hispanic origin. Eighty-one percent of the participants had an education equal to or greater than high school. The data was combined with Wave 1 full-bank test data to calculate item calibrations.

Sleep and Wake Disturbance Sample

The sample for the Sleep and Wake Disturbance banks included administration of 128 Sleep Wake Functioning items to 1,993 individuals from YouGovPolimetrix (1,259 from general population, and 734 with self-identified sleep problem) and 259 individuals with sleep disorders from Pittsburgh clinical sites. The overall sample (n = 2, 252) was 43.8% female. The median age was 52 years old with 20.7% 65 and older and 79.3% less than 65 years old. Eighty-two percent were White, 12.6% Black, 2.7% Native American or Alaskan, 0.4% Native Hawaiian or Pacific Islander, and 5.9% other. Ten percent of the sample was Hispanic or Latino. Educational attainment ranged from high school or less (13.6%), some college (38.6%), college degree (27.9%) to advanced degree (19.9%). Item response data from the overall sample (2, 252 individuals) were used for item calibration.

C. Wave II testing

Incorporating feedback and recommendations from the PROMIS NIH science officers as well as the Scientific Advisory Board regarding direction of the next wave (Wave II) of PROMIS network research activity, the SC decided to focus on *validation of existing banks* for the five domains versus *development of new banks*. The SC established a strategy for deciding on Wave II activity, which included defining study design and validation criteria for Wave II network protocols, defining clinical populations for study, and initiating draft protocols in:

- Depression and low back pain with a community sample comparison
- Arthritis
- Congestive heart failure (CHF)
- Chronic obstructive pulmonary disease (COPD)
- Mode of Administration
- Cancer PROMIS Supplement

A summary of the ongoing Wave II testing protocols can be found at *Wave II Testing* (http://www.nihpromis.org/science/validitystudies.aspx), or at: http://www.nihpromis.org/documents/Protocol_Summaries_ClinicalTrials_gov[1].pdf.

V. Item Banks and Tools

PROMIS instruments derive primarily from calibrated item banks that enable computerized adaptive testing and multiple short forms of varying length. The following instruments are available as of April 1, 2008. All instruments can be accessed within Assessment Center[SM] (http://www.assessmentcenter.net/).

Instrument properties and a comment on the Wave I sample

PROMIS Version 1.0 instruments were developed based on data collected on an internet survey platform. As such, it can safely be considered valid for internet or personal computer-based applications with screen presentations of individual items. Comparability of results obtained using paper or telephone administration cannot be assumed until further testing is done. On average, respondents will answer 5 questions per minute, suggesting, for example, that a CAT administration of all 9 banks with an average of 5 items per bank will take about 10 minutes to complete.

A sample reflecting Year 2000 census demographics data was used for scale setting. Many participants with chronic disease (e.g., cancer, heart disease, arthritis, COPD, spinal cord injury, psychiatric conditions) were included.

A. Version 1.0 Item Banks

Domain	Item Bank # of Items	Short Form # of Items
Emotional Distress – Anger	29	8
Emotional Distress – Anxiety	29	7
Emotional Distress – Depression	28	8
Fatigue	95	7
Pain – Behavior	39	7
Pain – Impact	41	6
Physical Function	125	10
Satisfaction with Discretionary Social Activities	12	7
Satisfaction with Social Roles	14	7
Sleep Disturbance	27	8
Wake Disturbance	16	8
Global Health	—	10

Recall period. Most PROMIS item banks utilize a 7-day recall period. Some items were revised to accommodate this decision. This time frame was chosen because it is on the upper limits of ecological validity for specific events (especially for subjective symptoms), yet long enough to allow time for people to experience enough events. Additionally, many clinical trials look for response to therapy over a few weeks time, and a longer recall time would undermine the ability to detect benefits. “The past 7 days” is the reference period for all items in Anxiety, Anger, Depression, Fatigue, Pain Behavior, Pain Impact, Satisfaction with Discretionary Social Activities, and Satisfaction with Social Roles. Physical function items emphasize current capabilities and therefore do not employ a recall period. Item stems begin with phrases such as “Does your health now limit you” or “Are you able to.” Some global health items use a 7-day recall period while others do not employ a recall period and emphasize current status “in general.”

Response options. The term "response options" refers to the set of answers a respondent can choose from when responding to a question or statement. Virtually all PROMIS items employ five response options (e.g., 1=Not at all, 2=A little bit, 3=Somewhat, 4=Quite a bit, 5=Very much). Pain Behavior uses six response options to allow for respondents to endorse "no pain." The 10 PROMIS Global Health items each have 5 response choices, with the exception of the common 11-point pain intensity item ("How would you rate your pain on average" with 0=No pain and 10=Worst imaginable pain).

To the extent possible, the wording of response categories was kept consistent within banks, and a limited degree of variation in response options was used across banks. This was done to ease patient burden. An effort was made for items within a PROMIS domain to utilize the same or a small number of response options. Some flexibility in response choices within banks was considered important, however, to capture the range of patient experience in a domain (e.g., intensity, frequency, duration). Therefore, for example, most banks employed a common set of response options for intensity (i.e., "Not at all" to "Very much") and frequency (i.e., "Never" to "Always"). The final response categories were pre-tested with cognitive interviews to confirm patient comprehension, prior to field testing for item calibration. When possible, to ease responder burden, initial short forms developed from PROMIS item banks comprise items with a common set of response options. Exceptions are the 10 Global Health items which utilize 5 unique response options and the Physical Function short form which uses 2 response options. *Also see Instruments Available for Use:* http://www.nihpromis.org/Documents/Item_Bank_Tables_Feb_2011.pdf.

B. PROMIS User Manual Version 1.1.

The PROMIS User Manual provides information on the development of PROMIS instruments as well as instructions for their use. This includes (1) description of the development of the PROMIS instruments, (2) review of the properties of PROMIS instruments, (3) information about administration and scoring, and (4) discussion of validity and interpretation of PROMIS instruments. The PROMIS User Manual can be obtained through the Assessment Center.

C. Assessment Center (http://www.assessmentcenter.net/)

Assessment Center SM is a dynamic application that will allow researchers to centralize all research activities. Assessment Center SM includes features that promote instrument development, study administration, data management, and storage of statistical analysis results.

Assessment Center SM will aid in the selection, development and administration of research measures. Patient-reported outcome instruments developed through the PROMIS Network are the basis for the Assessment Center SM Item and Instrument Library.

Users will be able to review PROMIS instruments by previewing them as they would appear in data collection and/or with CAT details. Users may also download PDF versions of the PROMIS instruments. The PROMIS short forms and computer adaptive tests (CATs) may be administered to study participants via Assessment Center SM. Additionally, users can create their own PRO instruments through selecting or modifying existing items and/or creating new items.

Item development histories and Item/Instrument statistics are viewable for all PROMIS products. Users may define data collection elements in Assessment Center [SM] and create study-specific websites to administer PRO measures and consent forms. Users may manage data collection through accrual reports and participant registration screens. Assessment Center [SM] will enable users to export study data to external applications for analysis.

D. Assessment Center User Manual. (*See Assessment Center Manual:* http://www.assessmentcenter.net/ac1/AssessmentCenter_Manual.pdf.)

A comprehensive Assessment Center User Manual provides step by step instructions and information on registration, instruments, study set-up (basic and advanced), administration, copyright and terms of use, and security.

E. CAT Demonstration

A demonstration of a working computerized adaptive testing (CAT) for all calibrated PROMIS banks can be found at: *CAT demo* (https://www.assessmentcenter.net/ac1/Default.aspx?SID=DD7EBF2C-7F6D-4FFD-9953-94CBEA8169D3).

VI. Publications and Presentations

The PROMIS Publications Subcommittee was established in 2005 to oversee approval and dissemination of all publications derived from this project. The multi-institutional, multi-instrument, cooperative nature of this effort required careful consideration of the priority of a proposed publication vis-à-vis other proposals and available resources, as well as composition of authorship. The Publications Committee has been responsible for recommending general policy and procedure for review and approval of all PROMIS network presentations and publications, including any standardized way of representing PROMIS in the disseminated document. All co-investigators in PROMIS have been eligible to publish papers from the PROMIS effort, including its item bank and CAT activity. Individuals other than PROMIS investigators, such as consultants, were also eligible for authorship. All requests for access to data, proposed data analyses and draft manuscripts were sent in writing to the Publications Committee which reviewed each request in the context of the overall study goals, other requests, and the PROMIS mission, with attention to authorship guidelines and equity across collaborators. There are currently 38 published, 3 in-press, and 38 in-progress PROMIS-related manuscripts and over 200 PROMIS-related presentations. For a list of manuscripts go to *PROMIS publications* (http://www.nihpromis.org/science/publicationsyears); for a full list of presentations go to *PROMIS presentations* (http://www.nihpromis.org/science/presentations).

VII. References

Beels, C.C., Gutwirth, L., Berkeley, J., & Struening, E. (1984). Measurements of social support in schizophrenia. *Schizophrenia Bulletin*, 10(3):399-411.

Bjorner, J.B., Fayers, P.M., & Idler, E.L. (2005). Self-rated health. In P.M. Fayers. & R.D. Hays (eds), *Assessing Quality of Life* (pp. 309-323). Oxford: Oxford University Press.

Bowling, A., & Farquhar, M. (1991). Associations with social networks, social support, health status and psychiatric morbidity in three samples of elderly people. *Social Psychiatry and Psychiatric Epidemiology*, 26(3):115-26.

Brekke, J.S., Long, J.D., & Kay, D.D. (2002). The structure and invariance of a model of social functioning in schizophrenia. *Journal of Nervous and Mental Disorders*, 190(2):63-72.

Chang, H. (1999). Cancer pain management. *Medical Clinics of North America*, 83(3):711-736.

Clark, L.A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, 100, 316-336.

Cohen, S., & Syme, S.L. (1985). Issues in the study and application of social support. In S. Cohen & S.L. Syme (eds). *Social Support and Health* (pp. 3-22). Orlando, FL: Academic Press, Inc.

DeWalt, D.A., Rothrock, N., Yount, S., & Stone, A. on behalf of the PROMIS Cooperative Group (2007). Evaluation of item candidates: The PROMIS qualitative item review. *Medical Care,* 45*:* S12-S21.

Dijkers, M. P., Whiteneck, G., & El-Jaroudi, R. (2000). Measures of social outcomes in disability research. *Archives of Physical Medicine and Rehabilitation*, 81(12, suppl 2):S63-S80.

Glaus, A. (1998). *Fatigue in Patients With Cancer: Analysis and Assessment*. Heidelberg, Germany: Springer-Verlag Berlin.

Haley, S.M., Coster, W.J., & Binda-Sundberg, K. (1994). Measuring physical disablement: The contextual challenge. *Physical Therapy*, 74(5):443-451.

Haley, S.M., McHorney, C.A., & Ware, J.E., Jr. (1994). Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *Journal of Clinical Epidemiology*, 47(6):671-684.

Hays, R., Cella, D., Du, H. et al. (2007) Item response theory analyses of physical functioning items in the medical outcomes study. Medical Care, 45 (5), S32-S38.

Horowitz, L.M., Rosenberg, S.E., Baer, B.A., Ureno, G., & Villasenor, V.S. (1988). The inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56, 885-892.

Larson, J.S. (1993). The measurement of social well-being. *Social Indicators Research*, 28:285-296.

Lord, F. M. & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings." Applied Psychological Measurement, 8, 453-461.

Lorenz, K.A., Cunningham, W.E., Spritzer, K.L., & Hays, R.D. (2006). Changes in symptoms and health-related quality of life in a nationally representative sample of adults in treatment for HIV. *Quality of Life Research*, 15, 951-958.

McDowell, I., & Newell, C. (1996). *Measuring Health: A Guide to Rating Scales and Questionnaires, second edition*. New York, NY: Oxford University Press.

Merskey H., & Bogduk, N. (eds.) (1994). *Classification of Chronic Pain: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms, second edition*. Seattle, WA: IASP Press.

Meuser, T., Pietruck, C., Radbruch, L., Stute, P., Lehmann, K. A., & Grond, S. (2001). Symptoms during cancer pain treatment following WHO-guidelines: A longitudinal follow-up study of symptom prevalence, severity, and etiology. *Pain*, 93(3):247-257.

North American Nursing Diagnosis Association (1996). *Nursing Diagnoses: Definition and Classification, 1997-1998*. Philadelphia, PA: McGraw-Hill.

Reeve, B., Hays, R., Bjorner, J., Cook, K., Crane, P., Teresi, J., Thissen, D., Revicki, D., Weiss, D., Hambleton, R., Liu, H., Gershon, R., Reise, S., Cella, D. (2007) Psychometric evaluation and calibration of health-related quality of life items banks: Plans for the Patient-Reported Outcome Measurement Information System (PROMIS). Medical Care, 45 (5), S22-S31.

Sherbourne, C. D. (1992). Pain measures. In A. L. Stewart & J.E. Ware (eds.) *Measuring Functional Status and Well-Being: The Medical Outcomes Study Approach* (pp. 230-234). Durham, NC: Duke University Press.

Stewart, A. L., & Kamberg, C. (1992). Physical functioning. In A.L. Stewart & J.E. Ware (eds.) *Measuring Functional Status and Well-Being: The Medical Outcomes Study Approach* (pp. 86-142). Durham, NC: Duke University Press.

Stewart, A. L., Hays, R. D., & Ware, J. E. (1992). Health perceptions, energy/fatigue, and health distress measures. In A.L. Stewart & J.E. Ware (eds.) *Measuring Functional Status and Well-Being: The Medical Outcomes Study Approach* (pp. 143-172). Durham, NC: Duke University Press.

Wills, T.A. (1985). Supportive functions of interpersonal relationships. In S. Cohen & S.L. Syme (eds). *Social Support and Health* (pp. 61-82). Orlando, FL: Academic Press, Inc.

Wilson, I.B., & Cleary, P.D. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Journal of the American Medical Association*, 273(1):59-65.

Wortman, C.B., & Conway, T.L. (1985). The role of social support in adaptation and recovery from physical illness. In S. Cohen & S.L. Syme (eds). *Social Support and Health* (pp. 281-302). Orlando, FL: Academic Press, Inc.

Watson, D., & Clark, L.A. (1992). Affects separable and inseparable: On the hierarchical arrangement of the negative affects. *Journal of Personality and Social Psychology*, 62, 489-505.