



Published in final edited form as:

Arch Phys Med Rehabil. 2011 October ; 92(10 0): S12–S19. doi:10.1016/j.apmr.2011.04.025.

The PROMIS Initiative: Examples of Applications in Rehabilitation

Dagmar Amtmann, PhD¹, Karon F. Cook, PhD¹, Kurt L. Johnson, Ph.D.¹, and David Cella, PhD²

¹Department of Rehabilitation Medicine, University of Washington, Seattle, WA

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

Abstract

In order to fully capture the impact of a disease or condition on the lives of individuals, patient-reported outcomes (PROs) are considered a necessary component of health measurement in rehabilitation. This article provides an overview of the involvement of rehabilitation stakeholders in the development of sound measurement tools for the Patient-Reported Outcomes Measurement Information System (PROMIS), a National Institutes of Health-funded initiative. PROMIS is a multi-site study that included many different populations. Here we focus on the involvement of people with several chronic conditions, including multiple sclerosis, spinal cord injury, and arthritis in the development of PROMIS measures. We describe both qualitative and quantitative methods used, including expert panels, focus groups, cognitive interviews and item response theory modeling, which resulted in enhanced utility of PROMIS measures in rehabilitation. The measures include a set of global health items and twelve item banks representing six domains. Scores are reported in the T-score metric (Mean = 50, SD=10) and centered on means from the U.S. general population. The PROMIS item banks measure quality of life and symptoms of people with chronic conditions and have the potential to enhance research and clinical practice by facilitating comparisons of scores across domains and populations.

Keywords

Outcome Assessment (Health Care); rehabilitation; Disabled Persons

© 2011 The American Congress of Rehabilitation Medicine. Published by Elsevier Inc. All rights reserved.

Corresponding author: Dagmar Amtmann, University of Washington, Department of Rehabilitation Medicine, Box 354237, Seattle, WA 98195-4237, (206) 543-4741 V, (206) 685-9224 Fax, dagmara@u.washington.edu.

Address reprint requests to: Dagmar Amtmann, University of Washington, Department of Rehabilitation Medicine, Box 354237, Seattle, WA 98195-4237, (206) 543-4741 V, (206) 685-9224 Fax, dagmara@u.washington.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Some of the results of focus groups and cognitive interviews mentioned in this manuscript were presented at the PROMIS conferences held in Bethesda, MD on 9/11-12/2006 and 3/3-4/2008

We certify that no party having a direct interest in the results of the research supporting this article has or will confer a benefit on us or on any organization with which we are associated and we certify that all financial and material support for this research and work are clearly identified below.

The manuscript submitted does not contain information about medical device(s).

A patient-reported outcome (PRO) is defined as “a measurement of any aspect of a patient's health status that comes directly from the patient (i.e., without the interpretation of the patient's responses by a physician or anyone else).”¹ In recognition of the important role that PROs play in health research, the Patient-Reported Outcomes Measurement Information System (PROMIS) was funded by the National Institutes of Health in 2004 to develop a psychometrically sound and clinically meaningful measurement system of PROs. PROMIS is based on the World Health Organization's (WHO) definition of health: “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.”² In emerging synergy with the WHO's International Classification of Functioning, Disability and Health (ICF), which provides a classification taxonomy to organize components of function, disability and health, PROMIS instruments provide a means to measure health related concepts with a focus on the individual's perceived health experiences.³ Using state-of the art qualitative and quantitative methodology, twelve item banks covering six domains of perceived health were developed in the first PROMIS funding cycle between 2004 and 2009.⁴ Data on all items in the banks have been collected in large samples from both community and clinical populations, including samples of people with chronic conditions. For each domain, computerized adaptive testing (CAT) and 6- to 10-item short form (SF) formats are available. The work by the PROMIS network has been documented extensively on the website (www.nihpromis.org) and in numerous publications.⁵⁻¹⁷

The critical role outcomes research plays in rehabilitation has been acknowledged. As stated by Johnston and colleagues,¹⁸ “By definition, outcomes are central to medical rehabilitation. Nothing, then, can be more essential to research in medical rehabilitation than outcomes research,” (p.1). This article describes the involvement of various rehabilitation stakeholders in the development of PROMIS item banks and provides examples of how PROMIS instruments can be effectively used in rehabilitation research. The PROMIS project is complex and involves many facets. We highlight aspects of PROMIS development and application that relate to a variety of stakeholder groups and research methods. Sources for the detailed methodologies are referenced. First, we outline the role of PROs in rehabilitation research, and then we describe how perspectives of clinicians and people with chronic conditions were included in the development and testing of item banks. Lastly, we offer a preview of current and future efforts aimed at ensuring that the PROMIS tools are relevant to rehabilitation populations.

THE ROLE OF PATIENT REPORT IN REHABILITATION RESEARCH

Prior to the introduction of the concept of PROs, health measures predominantly targeted pathophysiologic disturbances focusing on death, disease, and disability.¹⁹ While such measures of impairments of body functions and structures are useful to describe and categorize the course of a disease or condition, they fail to fully capture the impact of a disease or condition on the lives of individuals. The dimensions of health that PROs measure, including quality of life and aspects of social and mental health, tend to be experienced and perceived at the level of the individual. PROs are the best means to capture such critical information which augments existing measures of impaired body functions and structures,²⁰ thus facilitating patient-clinician communication and shared decision-making.²¹

In many fields PROs have been found to be as predictive, and sometimes more predictive of survival duration than traditional prognostic factors.²²⁻³⁰ In rehabilitation the relationship between symptoms and disability is often complex with many factors contributing to global outcomes such as employment and quality of life. Patient report measures ask respondents to report their perceptions or opinions without any interpretation or interference by a researcher

or a clinician. Patient report measures often are contrasted with performance-based measures. Performance-based measures ask participants to perform a task, such as walking 20 feet, and record the performance by, for instance measuring the time or the number of correct responses. Typically, patient report and performance measures have not been found to be highly correlated.³¹

Brouwer et al³² and Reneman et al³³ suggested the possibility that patient report and performance-based measures access different information. More important than the degree to which patient report and performance-based measures correlate with each other is whether they identify outcomes of interest. For example, PROs have been found to identify changes in health status overlooked by so-called “objective measures.”³⁴ An important consideration in clinical trials is whether the chosen measures are both sensitive and responsive to change in the outcome of interest.³⁵ Latham and colleagues (2008) found that performance-based and patient report measures were equally and adequately sensitive and responsive to change.³⁶ More longitudinal research is necessary to better understand predictive validity of performance-based measures as compared to PROs in order to select measures that are most predictive of future outcomes of interest to rehabilitation researchers.

INVOLVEMENT OF REHABILITATION STAKEHOLDERS AND VARIED RESEARCH METHODS IN PROMIS ITEM BANK DEVELOPMENT

The development of PROMIS item banks followed detailed protocols approved by the PROMIS network. The process involved numerous steps which have been described elsewhere.⁸ The University of Washington Center on Outcomes Research in Rehabilitation (UWCORR) in Seattle was the PROMIS research site that focused on measurement in rehabilitation populations. All focus groups and cognitive interviews conducted at UWCORR involved people with chronic conditions. Here we describe: (1) involvement of rehabilitation professionals in expert review; (2) focus groups with people with chronic conditions to evaluate the impact of pain, fatigue, and physical function limitations on life activities and experiences; (3) cognitive interviews with people with chronic conditions to evaluate the function of PROMIS pain items; and (4) work to adapt PROMIS physical function items for users of assistive technology.

Expert Panels

PROMIS convened panels of experts for each domain, called domain workgroups. The experts were nominated from each PROMIS research site and the coordinating center, and experts from outside of PROMIS were also invited to participate.⁸ As part of the pain item bank development, health care providers who work with people with chronic conditions were involved in: (1) defining constructs to be measured, (2) identifying relevant subdomains, (3) reviewing individual items, and (4) interpreting results from the psychometric analyses of archival data. The team of experts consulted in the development of the PROMIS pain bank included clinical psychologists, physicians, and other clinicians who specialize in working with rehabilitation populations. In addition, outcomes researchers with particular experience in developing patient report measures for rehabilitation populations reviewed items and assisted in the development of items.

Focus Groups

The initial list of PROMIS item banks covered domains of pain, fatigue, social function, emotional distress and physical function. In spring 2007, focus groups were conducted with people with chronic conditions living with pain, fatigue, and/or limitations in physical function in order to confirm domain definitions, identify common language used when

discussing each domain, and identify any important concepts not covered by the respective candidate PROMIS item banks.

Participants were 23 adults with a variety of chronic conditions who were recruited from the rehabilitation clinics at the University of Washington. Recruitment procedures and detailed methodology can be found elsewhere.³⁷ Five outcome-specific focus groups were conducted at the Seattle site—two each on pain and physical function and one on fatigue. Demographics of participants are reported in Table 1 and reflect the ethnic and educational background of patient populations at the recruiting site. Following the approved PROMIS study protocol, each focus group was moderated by a trained facilitator and recorded by a court-recorder. We used semistructured interviews with open-ended questions to encourage group participation and discussion. Examples of open-ended prompts included:

1. To what extent have you been able to predict your pain/fatigue?
2. In the past 30 days, how has pain/fatigue affected your ability to carry out your daily activities?
3. What has been the greatest impact of your pain/fatigue on your life?
4. We are interested in the best ways to ask questions about pain...

The facilitator focused the discussion on the topic at hand, ensured that all participants had an opportunity to contribute, and summarized participants' responses, but was otherwise non-directive, non-evaluative and supportive.

Transcripts of focus groups were analyzed using qualitative research methodology. Researchers carefully read the transcripts, identifying common themes and sub-themes related to the participants' experiences with pain, fatigue, or physical function limitations. A series of codes associated with quotes in the transcript were used to index the data. Overall, the focus groups found that the construct definitions were appropriate and the item banks covered most of the relevant concepts of living with pain, fatigue, and limitations in physical function. However, based on participant input the content of the candidate PROMIS item banks was modified to cover all important areas identified in focus groups.

Focus group participants reported that the most relevant aspect of their pain, fatigue, and physical limitations was interference with their abilities to function. They believed measuring the impact of pain, fatigue, and physical function limitation on valued activities and roles was more important than measuring the severity of these symptoms per se. Participants observed that it is difficult to meaningfully answer questions that ask them to separate the effects of different symptoms. For example, participants experiencing substantial fatigue and pain found it difficult to accurately quantify the contribution of each to social activities. These results revealed the tension between scientists' efforts to distinguish how each of the different constructs (e.g., pain or fatigue) contributes to patients' experiences and the global and cumulative impact of different symptoms on valued roles and activities. As a result, statements requiring respondents to make attributions, such as whether their fatigue was a result of pain (e.g., because of my pain I was tired), were removed.

Participants identified predictability (or unpredictability) as an important concept that was not well covered by the PROMIS item banks. For example, some reported that the unpredictability of their pain or fatigue made planning for social activities difficult, resulting in reduced social participation and the perception of others that they were "unreliable." Based on this feedback, new items were added to the candidate PROMIS item bank that asked about impact of unpredictability of pain on activities.

Once feedback from focus groups was incorporated into the item banks, the candidate items were ready for cognitive interviews. The Seattle site only conducted cognitive interviews with the pain items. Qualitative studies of the fatigue⁷ and physical function³⁸ item banks were handled at different PROMIS research centers and did not include the populations directly relevant to rehabilitation.

Cognitive Interviews to Evaluate Items Measuring Different Aspects of Pain

Cognitive interviewing is a method to assess whether items are meaningful to respondents and whether item content is perceived similarly across individuals.³⁹ Cognitive interviews were conducted in an effort to determine whether the experiences and perspectives of people with chronic conditions influence their interpretation of items. In these interviews researchers examined the processes participants used in responding to candidate items and whether the items were understood as intended. Cognitive interviewing protocols were based on the work of Willis.³⁹ The sample size was based on the requirement that a minimum of five people review each item. All candidate items were divided into sets of 30 items and each set of 30 was reviewed with 5 individuals. This sample size was considered sufficient because of the specific focus of the cognitive interviews on item functioning. In addition, most items for PROMIS banks were modifications of existing items rather than newly created items, and most items had been previously examined in cognitive interviews. Items that underwent major revisions based on the first round of cognitive interviews were subjected to 3 to 5 additional interviews. Each item was reviewed by at least 2 participants with less than 12 years of education and/or a measured reading level less than the ninth grade using the Wide Range Achievement Test-3 Reading subtest; and/or a diagnosis associated with cognitive impairment. For more information about the procedures, inclusion and exclusion criteria see Dewalt, et al (2007).⁸ All participants were screened during the consenting process by trained interviewers to assure that their cognitive functioning was adequate to understand questions and respond meaningfully. People with mild cognitive impairment were encouraged to participate.

Forty-four persons were recruited from rehabilitation clinics at the University of Washington to participate in cognitive interviews of candidate PROMIS pain items. The detailed outline of the development of the pain interference bank, including how the items were selected, tested, and the psychometric properties of the resulting item bank can be found in Amtmann et al (2010).⁴⁰ Here we only summarize the work as it relates to the involvement of people with chronic conditions. Demographics for participants are reported in Table 1. Cognitive interview participants reported on average “moderate” (on a scale from none to severe or very severe) levels of pain, fatigue and physical function limitations over the past seven days. Some participants without pain in the past seven days were recruited to evaluate the functioning of items for people across all levels of pain, including no pain. Many health conditions were reported by various participants as their primary diagnoses. The most common diagnoses were arthritis, spinal cord injury and brain injury (including stroke).

Based on expert reviews and feedback from the focus groups, a set of items measuring three pain domains (pain interference, pain quality and pain behaviors) was created. Participants reviewed approximately 30 items each, and were then asked what they were thinking when they answered each question. They were also encouraged to rephrase or give alternative wording for questions if appropriate. Based on participant feedback in the first round of cognitive interviews, 97 of the 167 items were found to require no or only minor changes. The remaining 70 items were rewritten and reviewed in a second round of cognitive interviews. In addition, 5 new items referencing how pain affects sleep, mood, and work were developed based on participant suggestions. Revised and new items were reviewed in additional rounds (up to seven) of cognitive interviews until the items were either considered to be acceptable or were eliminated. Final candidate item banks for pain subdomains (pain

interference, pain quality, and pain behaviors) were administered in the PROMIS large scale data collection.

Cognitive interviews identified problems with item clarity, timeframe, and response options. By far the most common problems with items reported by participants were issues with clarity. Because of difficulty understanding the meaning of specific terms and phrases, there was substantial variation in how people interpreted questions. Thus, people may respond to a vague item based on conceptually distinct understandings of what is being asked. Items asking people about different pain qualities (e.g., shooting, burning) were particularly subject to varied understandings. An example of this is the term “splitting” pain, which was variously perceived as referring to joint and muscle pain for one person, the stomach flu for another, and “a lot of pain” for another. Participants reported that many of these items were difficult to understand, and items were revised where possible.

Cognitive interview participants also expressed difficulty with the 7 day timeframe and the 11-point response scale used for PROMIS pain items. Although participants were explicitly asked to consider the past 7 days, many disregarded this instruction and substituted a timeframe they perceived as more meaningful or appropriate, such as “in general” or “since my diagnosis.” We also examined participant preferences for response options with numerical versus word descriptors. Although some participants strongly preferred the 0 to 10 response option, about 40% disliked it and reported that the numbers were meaningless to them. Overall, most participants preferred frequency responses (“never, rarely, sometimes, often, always”) to intensity response options (“not at all, a little bit, somewhat, quite a bit, very much”).

PROMIS Physical Function Bank and Users of Assistive Technology

One of the domains for which PROMIS developed an item bank was physical function. Our review of the completed bank suggested that not all items were appropriate for rehabilitation populations, especially those who use assistive technology (AT). For example, items about “climbing stairs” and “walking a block” were not appropriate for persons who use a wheelchair. Assistive technology is defined as any item, piece of equipment, or system, whether acquired commercially, modified, or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities.⁴¹

In 2008, the PROMIS research site in Seattle received supplemental funding to ensure that the PROMIS physical function bank was appropriate for users of wheelchairs, mobility aids and other AT. Three expert panels (one in Houston and two in Seattle) consisting of occupational therapists who specialize in AT, physical therapists (including one board-certified neurologic specialist), psychometricians, and two discussion leaders for each panel were convened to define physical function in the context of AT use. Convenience samples of health care providers were recruited from local health care facilities to form the expert panels. The main recruitment consideration was that the panelists be familiar with issues of measuring physical function in users of mobility aids and to ensure that most of the diverse perspectives important to the task of making the item bank appropriate for users of mobility aids were likely to be heard.

The three expert panels were unanimous and emphatic in their recommendation that physical function be measured as what people with chronic conditions can accomplish using the AT available to them, without regard for whether functional tasks were aided or not. In this context, “available AT” included any physical device (e.g., brace, cane, reacher) that assisted participants in independently accomplishing physical tasks and chores. Because of the project’s focus on AT, personal assistants were specifically excluded from this definition. The members of the expert panels also noted that the ability to distinguish

between what people could do with and without AT was important in some contexts (e.g., improvement indicated not by increased function with support, but by reduced need for AT). They recommended, however, that we first develop an item bank that can measure a person's performance with AT, and then explore options for measuring abilities both with and without AT, as well as with the help of another person.

Based on the definitions of physical function suggested by the expert panels, the research team reviewed all PROMIS physical function items and identified those that could be problematic for users of AT (e.g., items that assumed ability to walk or run). Potentially problematic items were modified (when possible), and new items were written to cover content relevant to measuring physical function in the context of AT. Items that could not be modified (e.g. items about stairs) were flagged. In future uses of PROMIS, these items will not be administered to persons for whom they would be inappropriate.

The modifications to the physical function item bank were not intended to distinguish between what people can accomplish with and without AT, but rather to make the existing bank appropriate for users of AT. New and modified items were tested in several rounds of cognitive interviews with users of AT. Based on feedback, a candidate item bank of supplemental items to the PROMIS physical function bank was finalized and administered to over 700 participants, divided approximately equally between multiple sclerosis (MS), spinal cord injury (SCI), and arthritis (the numbers vary slightly depending on the missing data for a specific instrument). Participant demographics are reported in Table 1. A subset of unchanged PROMIS items was also administered. These served as "anchor" items for calibrating the new items to the original PROMIS-Physical Function metric.

Item responses were calibrated using an item response theory (IRT) model. Subsequent analyses indicated that the added items substantially improved the precision of the bank, particularly in measuring individuals with low physical function. Because the new items do not specifically mention AT, they are appropriate both for people who do and do not use AT. The process of IRT as it relates to the development of CAT is described by Lai et al in this issue. Modified banks will also be available for CAT via the PROMIS Assessment Center, which is PROMIS's publicly available, online, dynamic research management tool. It contains features that promote instrument development, study administration, data management, and storage of statistical analysis results. Assessment Center also houses a library of instruments and items with an emphasis on health-related quality of life.⁴²

USING PROMIS SHORT FORMS TO ASSESS PEOPLE WITH MS AND SCI

At the University of Washington PROMIS research site 11 PROMIS SFs were administered to a sample of approximately 700 individuals with MS and SCI as a part of the longitudinal study of the natural course of disease. A total of 76 items were administered to measure 11 PROMIS domains: Anger, Anxiety, Depression, Fatigue, Pain-Interference, Physical Function, Satisfaction with Discretionary Social Activities, Satisfaction with Social Roles, Sleep Disturbance, and Sleep-Related Impairment. The PROMIS SFs were administered at the fifth time-point in a longitudinal study that measured participants on a variety of domains, including pain, fatigue, physical function, depression, anxiety, and sleep, every four months using paper-and-pencil surveys. Demographics for this sample are reported in Table 2. We note that the results come from a longitudinal study using a sample that was not specifically recruited to be representative of MS and SCI populations in general. We present the results here to demonstrate how PROMIS scores can be used to examine complex patterns of symptoms and QOL indicators and compare these patterns across populations. Table 3 lists scores on all PROMIS measures administered to the MS and SCI samples.

One of the advantages of the PROMIS item banks is that scores are reported in a T-score metric that is anchored to the average levels of the measured outcomes in the United States general population (Mean=50, standard deviation (SD)=10).⁴³ The general population sample used for norming was similar to the population characteristics in the 2000 U.S. Census. U.S. Census data were used as the standard to generate analytical weights to ensure that the inferences based on PROMIS estimates would be applicable to the general population. PROMIS specified target representation for the calibration sample in terms of gender (50% female), age (20% in each of the five age groups: 18–29, 30–44, 45–59, 60–74, 75+ years), race/ethnicity (12.5% each for black and Hispanic), and education (10% less than high school graduate). Equivalence testing showed similarity between PROMIS general population and national norms with regard to body mass index, EQ-5D health index (EuroQol group's descriptive system of health-related quality of life states consisting of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression),⁴⁴ and a self-rating of general health.⁴³ A mean PROMIS-Physical Function score of 40 indicates a physical function level one standard deviation below the population mean.

The results from the MS and SCI samples are compared to the U.S. general population sample in graphical form in Figure 1. The direction of PROMIS scores follows the name of the bank. For example, higher scores on physical function denote greater physical function, and higher scores on fatigue indicate greater fatigue. As the figure indicates, compared to the general population mean of 50, our samples scored lower in Physical Function, Social Roles (involvement and satisfaction with work and family responsibilities) and Discretionary Roles (involvement and satisfaction with leisure activities and friendships), and higher in Fatigue and Pain Interference. Mean scores on Depression, Anxiety, Sleep and Sleep-Related Impairment in this sample were similar to those of the general population. Also evident in the figure is the similarity in the score profiles of the MS and SCI samples. The most substantial difference was in physical function. The mean score for people with MS was approximately one SD below the general population mean, while the mean score for people with SCI was almost two SDs below the general population mean. With increased use of the PROMIS measures data will be collected on larger clinical samples and condition-specific norms will be developed. Such data will facilitate score comparisons to other clinical populations as well as to the general population.

CURRENT AND PROSPECTIVE APPLICATIONS OF PROMIS MEASURES TO REHABILITATION RESEARCH

The use of IRT-calibrated item banks for PRO measurement offers many advantages over traditional approaches (e.g., CAT, centering scores to the mean of the general population). However, some rehabilitation researchers have expressed concerns regarding the applicability and sensitivity of measures developed in generic (cross-population) samples to rehabilitation samples.⁴⁵ Use of item banks to develop disease or condition-specific SFs may be one of the best ways to address such concerns. We recently initiated such an effort to develop an MS-specific PROMIS SF to measure fatigue. Rehabilitation researchers believe that fatigue is qualitatively different in MS because of its severity, sometimes sudden onset and considerable impact on cognitive function. An MS-specific fatigue SF would provide more evidence that PROMIS tools are appropriate for use in MS-related clinical trials.

To develop this SF, we recruited 23 clinicians to review PROMIS fatigue items and identify the items they believed to be most relevant to fatigue in MS. We are currently collecting data on experiences with fatigue from persons who have MS. Participants are asked to rank items in order of helpfulness in describing their experiences with fatigue, and also asked if there are any questions about their experience with fatigue that are not listed. When completed, the clinician and patient rankings will be triangulated with psychometric results

using both community and clinical sample data to develop a PROMIS-MS Fatigue SF. Once the MS-targeted SF is developed, we will administer both the MS and general SFs to a sample of people with MS and compare the results. This study will help us quantify the advantage (if any) of the disease-specific versus generic measurement tools. We plan to repeat this process with additional item banks and publish the results as they become available.

In the future, validation studies involving all PROMIS domains and populations with different conditions will assess sensitivity and responsiveness to change of PROMIS tools as compared to the currently used tools. In these studies the estimation of minimal clinically important differences (MCID) will play an important role. An MCID is the smallest measurable difference in a particular domain that patients and physicians consider to be clinically important.⁴⁶ Instruments that are responsive enough to detect MCIDs will help to evaluate with greater precision the effects of rehabilitation interventions. One current validation study involves people with chronic back pain receiving epidural steroid injections. Participants are measured on most PROMIS domains before the treatment (i.e., before the injection) and are then monitored for 3 months. The study will examine psychometric properties, including sensitivity and responsiveness to change, of the PROMIS CAT scores compared to the legacy instruments in people with back pain. This is the first study to examine functioning of the PROMIS instruments in rehabilitation populations. Data collection finished in spring 2010 and data are being analyzed at the time of submission of this article.

PROMIS also developed pediatric patient report and proxy-report item banks.^{47–49} Children with chronic conditions and their parents or caregivers have been involved in the development of pediatric banks as well. A total of 196 youths with chronic conditions participated in development of pediatric PROMIS item banks at the Seattle research site. Participant demographics are reported in Table 4. The most common conditions reported were neuromuscular disease, spina bifida and cerebral palsy, although a considerable number of other conditions were represented. In addition, 125 parents or guardians participated in the development of proxy banks. In the second cycle of PROMIS the pediatric scores and adult scores will be linked, providing a way for measuring people across the lifespan using age-appropriate items and obtaining scores that are on the same metric. This will greatly facilitate longitudinal research.

CONCLUSION

PROMIS item banks have been developed to address the needs of rehabilitation researchers with contributions from people with chronic conditions, rehabilitation clinicians, and researchers. Patient report provides an important perspective on the disability experience that cannot be obtained in any other way. Publicly available, psychometrically sound instruments that can be customized for measuring specific populations or conditions have the potential to greatly improve measurement of PROs in rehabilitation. With brief, flexible, and sensitive instruments freely available as part of the PROMIS Assessment Center, clinical trialists can measure the constructs that matter the most to patients and clinicians, such as interference with valued activities. A considerable body of research was completed in the first cycle of PROMIS, much of it highly relevant to rehabilitation. Out of this body of research will come the published results of other validation studies, the development of disease specific norms for all domains, and the release of pediatric item banks.

Acknowledgments

NIH Science Officers on this project have included Deborah Ader, PhD, Susan Czajkowski, PhD, Lawrence Fine, MD, DrPH, Laura Lee Johnson, PhD, Louis Quatrano, PhD, Bryce Reeve, PhD, William Riley, PhD, Susana Serrate-Sztejn, MD, and James Witter, MD, PhD. See the web site at www.nihpromis.org for additional information on the PROMIS cooperative group.

Funding: The Patient-Reported Outcomes Measurement Information System (PROMIS) is a National Institutes of Health (NIH) Roadmap initiative to develop a computerized system measuring patient-reported outcomes in respondents with a wide range of chronic diseases and demographic characteristics. PROMIS was funded by cooperative agreements to a Statistical Coordinating Center (Northwestern University, PI: David Cella, PhD, U01AR52177) and six Primary Research Sites (Duke University, PI: Kevin Weinfurt, PhD, U01AR52186; University of North Carolina, PI: Darren DeWalt, MD, MPH, U01AR52181; University of Pittsburgh, PI: Paul A. Pilkonis, PhD, U01AR52155; Stanford University, PI: James Fries, MD, U01AR52158; Stony Brook University, PI: Arthur Stone, PhD, U01AR52170; and University of Washington, PI: Dagmar Amtmann, PhD, U01AR52171).

LIST OF ABBREVIATIONS

AT	assistive technology
CAT	computerized adaptive testing
ICF	International Classification of Functioning, Disability and Health
IRT	item response theory
MCID	minimal clinically important difference
MS	multiple sclerosis
PRO	patient-reported outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
SCI	spinal cord injury
SD	standard deviation
SF	short form
WHO	World Health Organization

REFERENCES

1. Administration USFaD. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006; 4:79. [PubMed: 17034633]
2. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference; 19 June – 22 July 1946; New York. signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948
3. Tucker, CA.; Riley, A.; Lai, JS.; Cella, D.; Riley, W.; Forrest, CB. Annual Meeting of the WHO-FIC Network. Toronto, Canada: 2010. Conceptual Frameworks and Synergies for Measurement: PROMIS and ICF.
4. Cella D, Gershon R, Lai J, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res*. 2007; 16(Suppl 1):133–141. [PubMed: 17401637]
5. 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. *Med Care*. 2007 May; 45 Suppl 1(5):S3–S11. [PubMed: 17443116]
6. Castel L, Williams K, Bosworth H, et al. Content validity in the PROMIS social-health domain: a qualitative analysis of focus-group data. *Qual Life Res*. 2008; 17:737–749. [PubMed: 18478368]

7. Christodoulou C, Junghaenel D, DeWalt D, Rothrock N, Stone A. Cognitive interviewing in the evaluation of fatigue items: results from the patient-reported outcomes measurement information system (PROMIS). *Qual Life Res.* 2008; 17:1239–1246. [PubMed: 18850327]
8. DeWalt D, Rothrock N, Yount S, Stone A. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care.* 2007; 45 Suppl 1(5):S12–S21. [PubMed: 17443114]
9. Fries J, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol.* 2005; 23 Suppl 39(5):S53–S57. [PubMed: 16273785]
10. Fries J, Cella D, Rose M, Krishnan E, Bruce B. Progress in Assessing Physical Function in Arthritis: PROMIS Short Forms and Computerized Adaptive Testing. *J Rheumatol.* 2009; 36:2061–2066. [PubMed: 19738214]
11. Fries J, Krishnan E, Bruce B. Items, Instruments, Crosswalks, and PROMIS. *J Rheumatol.* 2009; 36:1093–1095. [PubMed: 19509084]
12. Garcia S, Cella D, Clauser S, et al. Standardizing patient-reported outcomes assessment in cancer clinical trials: a patient-reported outcomes measurement information system initiative. *J Clin Oncol.* 2007; 25:5106–5112. [PubMed: 17991929]
13. Klem M, Saghabi E, Abromitis R, Stover A, Dew M, Pilkonis P. Building PROMIS item banks: librarians as co-investigators. *Qual Life Res.* 2009; 18:881–888. [PubMed: 19548118]
14. Reeve B, Hays R, Bjorner J, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care.* 2007 May; 45 Suppl 1(5):S22–S31. [PubMed: 17443115]
15. Revicki D, Chen W, Harnam N, et al. Development and psychometric analysis of the PROMIS pain behavior item bank. *Pain.* 2009 Aug 14.
16. Revicki D, Kawata A, Harnam N, Chen W, Hays R, Cella D. Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. *Qual Life Res.* 2009; 18:783–791. [PubMed: 19472072]
17. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol.* 2008; 61:17–33. [PubMed: 18083459]
18. Johnston, M.; Stineman, M.; Velozo, C. Outcomes research in medical rehabilitation: Foundations from the past and directions for the future. In: M, F., editor. *Assessing Medical Rehabilitation Practices : The Promise of Outcomes Research.* Baltimore, MD: Paul H. Brookes Publishing Co; 1997. p. 1-42.
19. Greenfield S, Nelson E. Recent developments and future issues in the use of health status assessment measures in clinical settings. *Med Care.* 1992; 30:MS23. [PubMed: 1583935]
20. Schwartz C, Sprangers M. An introduction to quality of life assessment in oncology: The value of measuring patient-reported outcomes. *Amer J Manag Care.* 2002; 9(18 Suppl):S550–S559. [PubMed: 12512979]
21. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *J Eval Clin Pract.* 1999; 5:401–416. [PubMed: 10579704]
22. Butow P, Coates A, Dunn S. Psychosocial predictors of survival in metastatic melanoma. *J Clin Oncol.* 1999; 17:2256–2263. [PubMed: 10561283]
23. Coates A, Gebiski V, Signorini D, et al. Prognostic value of quality-of-life scores during chemotherapy for advanced breast cancer. *J Clin Oncol.* 1992; 10:1833–1838. [PubMed: 1453197]
24. Earlam S, Glover C, Fordy C, Burke D, Allen-Mersh T. Relation between tumor size, quality of life, and survival in patients with colorectal liver metastases. *J Clin Oncol.* 1996; 14:171–175. [PubMed: 8558193]
25. Goss C, Quittner A. Patient-reported outcomes in cystic fibrosis. *Proc Am Thorac Soc.* 2007; 4:378–386. [PubMed: 17652505]
26. Heidenreich P, Spertus J, Jones P, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. *J Amer Coll Cardiol.* 2006; 47:752–756. [PubMed: 16487840]

27. Matza L, Boye K, Yurgin N. Validation of two generic patient-reported outcome measures in patients with type 2 diabetes. *Health Qual Life Outcomes*. 2007; 31:47. [PubMed: 17672906]
28. Soto G, Jones P, Weintraub W, Krumholz H, Spertus J. Prognostic value of health status in patients with heart failure after acute myocardial infarction. *Circulation*. 2004; 110:546–551. [PubMed: 15262843]
29. Stull D, Krouse J, Meltzer E, et al. Development and Validation of the congestion quantifier seven-item test (CQ7): A screening tool for nasal congestion. *Value Health*. 2007; 10:457–465. [PubMed: 17970928]
30. Dharma-Wardene M, Au H, Hanson J, Dupere D, Hewitt J, Feeny D. Baseline FACT-G score is a predictor of survival for advanced lung cancer. *Qual Life Res*. 2004; 13:1209–1216. [PubMed: 15473499]
31. Stratford P, Kennedy D, Pagura S, Gollish J. The relationship between self-report and performance-related measures: questioning the content validity of timed tests. *Arthritis Rheum*. 2003; 49:535–540. [PubMed: 12910560]
32. Brouwer S, Dijkstra P, Steward R, Goeken L, Groothoff J, Geertzen J. Comparing self-report, clinical examination and functional testing in the assessment of work related limitations in patients with chronic low back pain. *Disabil Rehabil*. 2005; 27:999–1005. [PubMed: 16096253]
33. Reneman M, Jorritsma W, Schellekens J, Goeken L. Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain. *J Occup Rehabil*. 2002; 12:119–129. [PubMed: 12228943]
34. Rothman M, Beltran P, Cappelleri J, Lipscomb J, Teschendorf B. Group MFP-ROCM. Patient reported outcomes: conceptual issues. *Value Health*. 2007 Nov-Dec; Suppl 2(10):S66–S75. [PubMed: 17995476]
35. Liang MH. Longitudinal construct validity: Establishment of clinical meaning in patient evaluative instruments. *Medical Care*. 2000; 38 Suppl II(9):II84–II90. [PubMed: 10982093]
36. Latham N, Mehta V, Nguyen A, et al. Performance-based or self-report measures of physical function: Which should be used in clinical trials of hip fracture patients? *Arch Phys Med Rehabil*. 2008; 89:2146–2155. [PubMed: 18996244]
37. Yorkston K, Johnson K, Boesflug E, Skala J, Amtmann D. Communicating about the experience of pain and fatigue in disability. *Qual Life Res*. 2010 Dec 24; 19:243–251. [Epub ahead of print]. [PubMed: 20033786]
38. Bruce B, Fries JF, Ambrosini D, et al. Better assessment of physical function: item improvement is neglected but essential. *Arthritis Res Ther*. 2009; 11(6):R191. [PubMed: 20015354]
39. Willis, GB. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage Publications; 2005.
40. Amtmann D, Cook K, Jensen M, et al. Development of a PROMIS item bank to measure pain interference. *Pain*. 2010; 150(1):173–182. [PubMed: 20554116]
41. Rehabilitation Act Amendments, 29 U.S.C. §794. 1998
42. Cella D. Assessment Center User Manual. 2009 <http://www.assessmentcenter.net/ac1/>.
43. Liu H, Cella D, Gershon R, et al. Representativeness of the Patient-Reported Outcomes Measurement System Internet Panel. *J Clin Epidemiol*. 2010; 11(63):1169–1178. [PubMed: 20688473]
44. Group E. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy*. 1990; 16:199–208. [PubMed: 10109801]
45. Flynn KEDC, DeWitt EM, Schulman KA, Weinfurt KP. Using item banks to construct measures of patient reported outcomes in clinical trials: investigator perceptions. *Clin Trials*. 2008; 5(6):575–586. [PubMed: 19029206]
46. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989 Dec; 10(4):407–415. [PubMed: 2691207]
47. Irwin DE, Varni JW, Yeatts K, Dewalt DA. Cognitive interviewing methodology in the development of a pediatric item bank: a patient reported outcomes measurement information system (PROMIS) study. *Health Qual Life Outcomes*. 2009; 7(3)

48. Varni JW, Stucky BD, Thissen D, et al. PROMIS Pediatric Pain Interference Scale: An item response theory analysis of the Pediatric Pain Item Bank. *Journal of Pain*. 2010; 11(11):1109–1119. [PubMed: 20627819]
49. Yeatts KB, Stucky BD, Thissen D, et al. Construction of the Pediatric Asthma Impact Scale (PAIS) for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Journal of Asthma*. 2010; 47(3):295–302. [PubMed: 20394514]

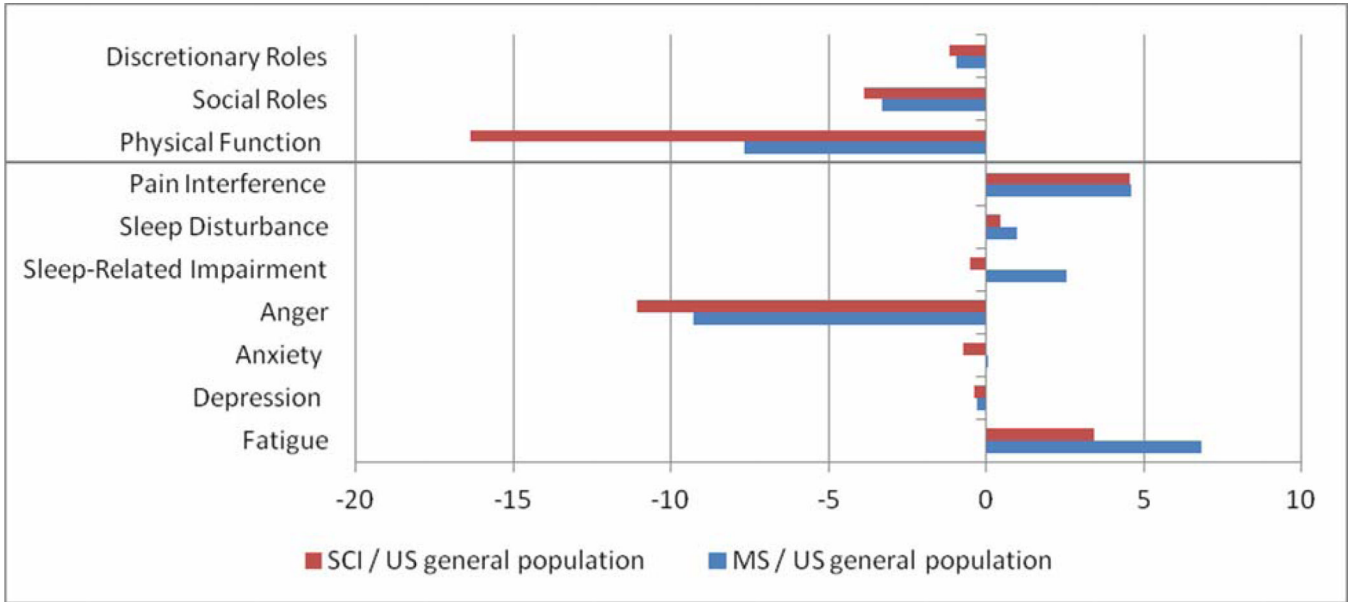


Figure 1. PROMIS T-scores for a sample of people with MS (N=461) and SCI (N=239) compared to PROMIS U.S. general population scores on PROMIS Short Forms. Above the line, higher scores = better (discretionary roles, social roles, and physical function) and below the line, higher scores = worse (pain interference, sleep disturbance, sleep-related impairment, anger, anxiety, depression, and fatigue).

Table 1

Participant Demographics for Focus Groups, Cognitive Interviews, and Physical Function Bank Studies

	Focus Groups (N=23)*	Cognitive Interviews (N=44)*	Physical Function Bank (N=758)
<i>Age</i>			
mean	45	46	54
SD	11	17	14
Female	18 (78%)	26 (59%)	467 (62%)
White	22 (96%)	37 (84%)	639 (84%)
High school or >	21 (91%)	38 (86%)	710 (94%)
Employed (part-time or full-time)	4 (17%)	18 (41%)	Unknown
<i>Primary Condition</i>			
Acute bronchitis	0	1 (2%)	0
Amputation/limb deficiency	1 (4%)	0	0
Arthritis	1 (4%)	10 (23%)	216 (29%)
Blindness/low vision	0	1 (2%)	0
Brain injury (including stroke)	1 (4%)	6 (14%)	0
Cerebral palsy	2 (9%)	1 (2%)	0
Chronic fatigue syndrome	0	1 (2%)	0
Crohn's disease	0	1 (2%)	0
Diabetes mellitus	0	1 (2%)	0
Hearing loss	0	1 (2%)	0
Hereditary nerve disorder	1 (9%)	1 (2%)	0
HIV	0	1 (2%)	0
Multiple chemical sensitivities	1 (4%)	0	0
Multiple sclerosis	5 (22%)	4 (9%)	274 (36%)
Neuromuscular disease	1 (9%)	1 (2%)	0
Post-polio syndrome	2 (9%)	1 (2%)	0
Spinal cord injury	7 (30%)	6 (14%)	268 (35%)
Systemic lupus	0	1 (2%)	0
Thalassemias (hemolytic anemia)	0	1 (2%)	0
Vestibular dysfunction	1 (4%)	0	0
Other	0	4 (9%)	0

* Two individuals participated in both focus groups and cognitive interviews.

Table 2

Demographic Variables for Participants with MS or SCI Who Responded to 11 PROMIS Short Forms

	MS (N=461)	SCI (N=239)
Age		
mean	51	46
SD	11	14
Women	83%	38%
White	97%	85%
High school or >		
Employed (part-time or full-time)	56%	33%
Married	69%	46%
Duration of condition, in years		
mean	15	13
SD	10	10

Table 3

Short Form Scores for Participants with MS and SCI

	MS (N=461)		SCI (N=239)	
	Mean	SD	Mean	SD
Physical Function	40.22	10.21	31.53	8.74
Fatigue	56.84	9.14	53.41	8.09
Depression	50.21	9.83	50.13	9.88
Anxiety	50.47	9.60	49.67	9.22
Anger	48.35	9.35	48.47	9.93
Sleep-Related Impairment	52.55	9.24	49.48	9.4
Sleep Disturbance	50.79	10.08	50.27	10.05
Pain Interference *	57.42	8.15	57.37	7.28
Social Roles	46.12	9.63	45.53	9.23
Discretionary Roles	48.36	9.23	48.15	9.17

* Pain Interference only measured in participants experiencing pain: MS N=369; SCI N=239

Table 4

Participant Demographics for Youth Studies

	Youth Participants (N=196)
<i>Age</i>	
mean	13.5
SD	3
Female	91 (46%)
White	141 (72%)
<i>Primary Condition</i>	
Amputation/limb deficiency	16 (8%)
Arthritis	2 (1%)
Arthrogryposis	3 (1.5%)
Asthma	2 (1%)
Bladder dysfunction	2 (1%)
Brain injury	15 (8%)
Cancer	1 (0.5%)
Cerebral palsy	42 (21%)
Congenital brain malformation	2 (1%)
Congenital disorder, other	3 (1.5%)
Developmental delay	3 (1.5%)
Diabetes mellitus	2 (1%)
Guillain-Barre Syndrome	1 (0.5%)
Hereditary nerve disorder	2 (1%)
Hereditary bone disorder	2 (1%)
Lesch-Nyhan Disease	1 (0.5%)
Neuromuscular disease	38 (19%)
Pre-term stroke	2 (1%)
Spina bifida or tethered spinal cord syndrome	32 (16%)
Spinal cord injury	4 (2%)
Tuberous sclerosis	1 (0.5%)
Other	12 (6%)
Unknown	8 (4%)