



Published in final edited form as:

Nurs Outlook. 2014 ; 62(5): 339–345. doi:10.1016/j.outlook.2014.05.009.

Patient-Reported Outcomes Measurement Information System (PROMIS®): Efficient, Standardized Tools to Measure Self-Reported Health and Quality of Life

Margaret Bevens, RN, PhD^{a,*}, Alyson Ross, RN, PhD^b, and David Cella, PhD^c

^aClinical Nurse Scientist and Program Director, Scientific Resources, Nursing Research and Translational Science, Nursing Department, National Institutes of Health Clinical Center, Bethesda, MD

^bPostdoctoral Intramural Research Fellow, Nursing Research and Translational Science, Nursing Department, NIH Clinical Center, Bethesda, MD

^cChair, Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, and Director, Center for Patient-Centered Outcomes - Institute for Public Health and Medicine, Chicago, IL

Abstract

All nurses are interested in the effects of diseases and treatments on individuals. Patient reported outcome (PRO) measures are used to obtain self-reported information about symptoms, function, perceptions, and experiences. However, there are challenges to their use, including multiple measures of the same concept, widely varying quality, excessive length and complexity, and difficulty comparing findings across studies and conditions. To address these challenges, the National Institutes of Health funded the Patient-Reported Outcomes Measurement Information System (PROMIS®), a web-based repository of valid and reliable PRO measures of health concepts relevant to clinician and researchers. Through the PROMIS Assessment Center, clinicians and researchers can access PRO measures, administer computerized adaptive tests, collect self-report data, and report instant health assessments. The purpose of this paper is to summarize the development and validation of the PROMIS measures and to describe its current functionality as it relates to nursing science.

Introduction

All nurses, whether clinicians, researchers or academicians, are interested in the effects of diseases and treatments on individuals. While physiological data provide valuable information about what is occurring within one's body, it is difficult to capture accurately

*Corresponding author: Margaret Bevens, RN, PhD, Clinical Nurse Scientist and Program Director, Scientific Resources, Nursing Research and Translational Science, Nursing Department, National Institutes of Health Clinical Center, 10 Center Drive, Room 2B13 MSC 1151, Bethesda, MD 20892., mbevans@cc.nih.gov (M. Bevens).

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.

subjective attitudes, values and experiences. Exploring the individual's perceptions is important, as clinicians' perceptions often differ from those of their patients (Fromme, Eilers, Mori, Hsieh, & Beer, 2004), with clinicians underestimating symptom severity and overestimating function (Hendriks & Schouten, 2002; Laugsand et al., 2010). Patient self-reports long have been the backbone of nursing research and practice, and nurses value the rigorous process of identifying, operationalizing, and measuring these less tangible concepts (Waltz, Strickland, & Lenz, 2010). Only more recently regulatory agencies also have recognized the patient's perspective as essential for comprehensive quality care (Acquadro et al., 2003; FDA, 2006).

Patient reported outcome (PRO) measures are used to obtain self-reported information about an individual's function, such as physical, cognitive, and sexual function; symptoms such as sleep and fatigue; and perceptions such as social support and health-related quality of life (HRQL). PROs may stand alone as the sole measure of a concept, or they may complement clinician assessments and/or performance-based measures. In either case, PRO data enhance the making of treatment decisions and the determining of treatment effectiveness (Guyatt et al., 2007). These data can improve the accuracy of symptom and function assessment as well as HRQL reporting (Hendriks & Schouten, 2002), improve clinician-patient communication (Detmar, Muller, Schornagel, Wever, & Aaronson, 2002), and serve to validate patients while reinforcing patient autonomy (Lohr & Zebrack, 2009). Evidence suggests that the inclusion of routine PRO collection in patient care improves quality of care (Chen, Ou, & Hollis, 2013).

Historically, pencil and paper-based questionnaires have been the primary method for collecting self-reported data. Researchers across multiple disciplines have developed numerous questionnaires to measure generic concepts such as HRQL, as well as specific symptoms including anxiety and depression (McHorney, 1997). In the late 1980s, the Outcomes Management movement put a strong emphasis on "routinely and systematically measur[ing] the functioning and well-being of patients, along with disease specific clinical outcomes..." (Ellwood, 1988). Concurrently, interest in collecting PROs in clinical research was rising in many specialty areas such as oncology, rheumatology and cardiology. The demands of outcomes management and clinical research produced a variety of concerns including: 1) the multiple questionnaires measuring the same concept, 2) the length and difficulty of many of the measures, 3) the variation in psychometric quality, 4) the difficulty in comparing or combining data across different studies and populations, and 5) the difficulty in incorporating the measures in clinical practice. Addressing these concerns and enhancing measurement in health outcomes research became a priority (Reeve et al., 2007).

A method to systematically address the challenges faced in the area of PROs emerged from a National Institutes of Health (NIH) initiative titled, "the NIH Roadmap." In 2003, the NIH launched its Roadmap to focus on numerous challenges facing the scientific community and the roadblocks that were impeding these challenges, with a focus on issues that go beyond the interests of a single institute or center (Zerhouni, 2003). In 2004, under the specific objective to re-engineer the clinical research enterprise, the NIH funded the Patient-Reported Outcomes Measurement Information System (PROMIS[®]) project (nihpromis.gov). This multicenter cooperative group included six primary research sites in addition to a

statistical coordinating center, with the goal of centralizing the development and collection of PRO's and addressing associated problematic issues. Since its launch, PROMIS has implemented more than 40 studies involving more than 50,000 subjects. This has resulted in the development of multiple adult and pediatric patient-reported outcome measures, which are available in multiple languages (PROMIS, 2013). The purpose of this paper is to summarize the development and validation of the PROMIS measures and describe its current functionality as it relates to nursing science.

Development and Validation

The PROMIS initiative was led by a multicenter cooperative group with representatives from multiple disciplines. The cooperative group included a steering committee, an 11 member expert scientific advisory board, and an advisory panel for health outcomes that consisted of 22 experts and clinical trialists (Cella et al., 2007). Several of these content experts were the authors of the existing questionnaires or legacy measures that were often considered the “gold standard” for a particular concept or domain.

The organizing conceptual framework for PROMIS was grounded in the World Health Organization's (WHO) physical, mental, and social framework of health (WHO, 1946). This tripartite framework guided the identification of domains, specific feelings, functions, or perceptions important to patients including concepts such as pain and fatigue. This domain framework is detailed in Figure 1. Once the domains were established, candidate items were selected for an initial set of questions, otherwise known as an item bank, within each domain. Thousands of items measuring each domain were collected, many from existing legacy measures, and a rigorous qualitative item review process was initiated. Content experts, often including the authors of legacy measures, reviewed the items, along with healthy people as well as those with chronic health conditions. Subjects participated in focus groups and cognitive interviews to help identify gaps in previous measures and to ensure that items were understandable, even for those with low literacy (DeWalt, Rothrock, Yount, & Stone, 2007; Magasi et al., 2012).

Based upon the qualitative review, items were categorized, unified, and rewritten, a process referred to as binning and winnowing, to produce a set of relevant item pools (large collection of items) that could be calibrated using modern statistical techniques referred to as item response theory (IRT). In turn, IRT-calibrated item banks enable efficient, “smart testing” by a computer programmed to ask the most informative questions remaining in the item bank with a given individual respondent. This is referred to as computerized adaptive testing (CAT). Item pools within the individual domains have been extensively tested in both paper and various electronic forms for application and relevance in adults and pediatric (including parental proxy) populations as well as affected populations across multiple disease or illness categories including approximately 1000 patients with cancer and approximately 500 subjects per condition, including heart disease, rheumatoid arthritis, osteoarthritis, psychiatric conditions, spinal cord injury and COPD (Reeve, et al., 2007). Extensive testing was completed in the U.S. general population (n>7500) for normative data, with adequate representation with respect to gender, age, ethnicity and education. These data were analyzed and used to calibrate the item sets and build the PROMIS item banks.

PROMIS items and measures have undergone rigorous review and testing to ensure they are precise and accurate. Because of the extensive review process involved in the development of item pools, only the best items from existing legacy measures are included in PROMIS item pools. Problematic items, identified through expert review, cognitive interviews, confirmatory factor analysis and IRT calibration, were removed, leaving PROMIS with the ability to capture the information of the legacy measures, only with fewer questions and better precision. Pilkonis, Choi, Reise et al (2011) provide an excellent detailed example of the meticulous steps involved in the development and testing of PROMIS item banks measuring emotional distress including depression, anxiety and anger (Pilkonis et al., 2011). Following these rigorous steps ensured the reliability of PROMIS measures, and concurrent validity with legacy measures is strong across all domains (Cella et al., 2010).

Historically, most disciplines, including nursing, utilized classical-test theory (CTT) in the development of questionnaires (van Alphen, Halfens, Hasman, & Imbos, 1994). Questionnaires developed in the CTT tradition typically provide a total sum score of a concept with a fixed set of items. All items are considered parallel and equally-contributing subsets of the concept and are summed to provide a score that reflects the measured level of the respondent. The limitations of CTT may be tolerated in nursing and other disciplines because of a perception regarding the complexity of the primary alternative, Rasch Modeling and Item-Response Theory (IRT).

IRT is a family of measurement models that place items and people on a continuum from low to high levels of whatever underlying “trait” is being measured, such as depression, fatigue, physical or social function (Thomas, 2011). Each item can be modeled as a single measure of that trait, with more or less ability to discriminate people from one another at various points along the measurement continuum. As opposed to CTT that focuses on the reliability of the total measure, IRT focuses on the reliability of items as they vary along that continuum for a given person taking the test. That is, it varies depending on the respondent. Response-centered scaling increases the accuracy of measurement, providing an estimate of the concept, as opposed to the total score. This increased precision of measurement translates into fewer items that are more specific to the person, based on his or her response to the previous item (Lerdal et al., 2011).

IRT addresses many of the limitations of CTT. The advantages of IRT include: 1) fewer items and thus shorter questionnaires 2) increased item precision can reduce error and sample size requirements, and 3) calculating error at the individual level enables practical individual assessment (Thomas, 2011). Item responsiveness and reduction in floor and ceiling effects are two major advantages in the use of PROMIS measures (Fries, Rose, & Krishnan, 2011). IRT allows for computer adaptive testing (CAT) administration which addresses additional concerns with traditional pencil and paper questionnaires including subject burden and application to practice.

Application and Administration of PROMIS

The cooperative group built the Assessment Center (www.assessmentcenter.net), a repository for the item banks that serves as an electronic Web-based resource to administer

computerized adaptive tests, collect self-report data, and report instant health assessments (Gershon et al., 2010). This system, open to the public and intended for both research and clinical practice, was designed to ensure scientific excellence by rigorously and repeatedly testing and adapting the system for new populations and new domains and items.

In Assessment Center one can register as a user and have access to the instrument library that contains calibrated item banks and a number of instruments including domain-specific four to ten item short forms, profiles (collections of short forms), and CAT (Figure 2). In addition, one can access instruments from Neuro-QOL and the NIH toolbox, “sister” NIH initiatives that provide brief, reliable and valid measures of quality of life (Neuro-Qol) as well as cognitive, emotional, sensory, and motor function (NIH Toolbox) (Gershon, et al., 2010; Perez et al., 2007). The item banks are domain-specific as opposed to disease-specific, allowing for the opportunity to make comparisons across populations and conditions. PROMIS Item banks and measures are also available in a number of different languages.

Assessment Center offers a great deal of flexibility. Users can utilize the calibrated item banks and forms available in Assessment Center through PROMIS, Neuro-Qol, and the NIH toolbox, or they can customize data collection by adding their own demographic and non-PROMIS measures or adapting PROMIS forms for their specific population. Users can download and print instruments or use Assessment Center to collect and store secure data on study-specific websites. Researchers and clinicians can feel confident collecting PROMIS data using whatever method best fits the needs of their population, as all item banks have been tested with no significant variability across methods of administration including interactive voice response technology, traditional paper questionnaire, personal digital assistant, or personal computer on the Internet (Bjorner et al., 2013). Whatever method of delivery is chosen, online and telephone assistance is available from Assessment Center, and all of the PROMIS services are free to the public.

One of the advantages to using Assessment Center for data collection is the ability to use CAT. Because CAT uses subjects’ previous answers to choose the next question, each question zeroes in more closely to an individualized score. This improves reliability and allows for much shorter surveys, as typically only 3–7 items are needed to measure a concept. All PROMIS items are very brief, further reducing response burden.

Scoring and interpretation of PROMIS measures is made easy by applying a standardized metric that is based in the normative data of the general population ($M=50$, $SD=10$). Perhaps one of the greatest benefits of PROMIS is that the standardized metric not only allows for comparisons to the general U.S. population, but also allows for comparisons across studies and disease-specific populations. One of the biggest problems with past PRO research was the difficulty comparing study results using different measures of the same concept. In an attempt to address this issue and to further enhance the comparison potential of PROMIS, PROsetta Stone[®] was created (www.Prosettastone.org). PROsetta Stone[®] was developed to link PROMIS with related legacy measures (e.g. CES-D, SF-36, MASQ, Brief Pain Inventory, FACIT-Fatigue, etc...), providing equivalent scores for different scales measuring the same health outcome (Choi, Schalet, Cook, & Cella, In press). A cross-walk table matches each possible score on a “legacy” instrument to a T-score of the comparable

PROMIS instrument, establishing a relationship between scores on the different measures. A list of available cross-walk tables linking PROMIS domains to legacy measures is shown in Table 1.

The Future of PROMIS

With the passage of the Affordable Care Act, meaningful use currently is a hot topic. Payers are interested in creating a common data warehouse and in using PRO's to evaluate clinical care across providers and across populations; PROMIS could provide a level playing field for collecting such data. Clinicians are interested in incorporating PRO collection into electronic medical records (EMR), and PROMIS CAT are SNOMED and LOINC compliant and can thereby be integrated into almost any existing EMR with minimal programming to enhance data collection and symptom management. The reliability and real-time scoring and interpretation provided by PROMIS allow for true and meaningful translation into clinical practice; patients in clinical practice easily can complete the brief online surveys on their clinician's secure website, then have them immediately scored and interpreted in comparison to the general U.S. population or to their own scores from previous visits. Minimally important differences were estimated for PROMIS-Cancer scales in five domains (anxiety, depression, fatigue, pain, and physical functioning), and confirmation studies are underway in non-cancer populations (Yost, Eton, Garcia, & Cella, 2011).

In addition to clinical practice, PROMIS is being incorporated into a number of large-scale research initiatives. The PROMIS global health scale is being used to support the indicators of the Healthy People 2020 initiative (Barile et al., 2013). The PROMIS system is recommended for collecting PRO data for inclusion in clinical comparative effectiveness research (Basch et al., 2012) and as an effective system for use in multi-center clinical trials (Eisenstein, Diener, Nahm, & Weinfurt, 2011).

Conclusion & Recommendations

PROMIS provides a "universal language" for measuring health concepts of relevance to clinicians and researchers. Having a common metric allows for comparisons across populations, conditions, research studies and clinical practices. The advent of PROsetta stone to link PROMIS measures to legacy measures, allows researchers whose work is based on those measures to share a common metric with PROMIS or to integrate PROMIS measures into their research. The rigorous conceptual and scientific methodology incorporated into the development of PROMIS ensures the measures are not only valid and reliable, but are convenient, flexible, and inclusive regardless of language, literacy, and physical function.

PROMIS is not perfect. PROMIS is most precise when implemented using CAT, and the technology needed to implement CAT may not be available to all populations. However, the static (non-CAT) versions of PROMIS measures have been found to be at least as good as comparison ("legacy") measures. Painstaking care was taken in using a universal approach to translating PROMIS items into different languages so that there was one translation per language versus one translation per country or region. Likewise, nearly all PROMIS items are very simple and brief (3–5 words), to accommodate low levels of literacy. However,

problems with language and literacy can arise, just as they would with legacy measures. Should researchers or clinicians discover problems with understandability or interpretability in their populations, they should report these to Assessment Center.

PRO collection long has been a cornerstone of nursing research. The recent emphasis on the importance of PRO collection, highlighted by the enormous expenditure of resources devoted to initiatives such as PROMIS, the NIH toolbox, Neuro-QOL, as well as the Executive Abilities Measures and Instruments for Neurobehavioral Evaluation and Research (Examiner) and PhenX (consensus measures for phenotypes and exposures), represents an opportunity for nurse scientists. Because there is now an expectation for inclusion of PRO data into randomized clinical trials (RCTs), and this expectation is being promoted in the CONSORT guidelines (Calvert et al., 2013), opportunities exist for nurse scientist who specialize in PRO collection to partner with scientists conducting RCTs to add PROs to RCTs, strengthening the overall science while promoting interdisciplinary collaboration. Through the use of CAT, which takes an average of three to seven questions to reliably assess a domain such as anxiety, depression, physical function, pain, and fatigue, nurse scientists who conduct physiological research can collect symptom data on research subjects in multiple domains without excessively increasing subject burden. Because single momentary assessments are not often reliable (Stone, Broderick, & Kaell, 2010), the brevity of PROMIS measures makes them particularly useful for reducing subject burden in studies involving multiple assessments.

Perhaps the major obstacle to incorporating PROMIS into nursing science is the ability to let go of the attachment to CTT, legacy measures, and to the belief that measures have to be disease or population specific. This is no easy task. According to *Nursing Research* Editor, Susan Henly, “Clinging to traditional measurement practices is creating a quaint patchwork of results in nursing research journals and is blocking the accumulation of knowledge we need” (Henly, 2010). More than a decade ago the value of IRT was highlighted, noting that nurse researchers “are not reaping the benefits of IRT in the development of affective instruments” (Beck & Gable, 2001), yet many nurse scientists still know very little about IRT, and few nurses, whether in research or clinical practice, are well-versed in IRT-derived tools such as PROMIS. PROMIS is a tool that can be useful to nurses at all levels of research and practice, so nursing education programs should provide at least an introduction to item response theory as part of research methods coursework. It is imperative that all doctoral programs devote significant time to understanding measurement theory that leads to decision making about measures, so that as health science becomes more interdisciplinary in scope, everyone participates in selecting the appropriate measures for their clinical research.

Acknowledgments

Funding: Drs. Bevans and Ross: NIH Clinical Center Intramural Research Program, Dr. Cella: U54AR057951 PROMIS Statistical Center.

The authors would like to thank Stephen Klagholz, BS, for his support in manuscript formatting and submission.

References

- Acquadro C, Berzon R, Dubois D, Leidy NK, Marquis P, Revicki D, et al. Incorporating the patient's perspective into drug development and communication: An ad hoc task force report of the patient-reported outcomes (PRO) harmonization group meeting at the Food and Drug Administration, February 16, 2001. *Value in Health*. 2003; 6(5):522–531. [PubMed: 14627058]
- Barile JP, Reeve BB, Smith AW, Zack MM, Mitchell SA, Kobau R, et al. Monitoring population health for Healthy People 2020: evaluation of the NIH PROMIS(R) Global Health, CDC Healthy Days, and satisfaction with life instruments. *Quality of Life Research*. 2013; 22(6):1201–1211. [PubMed: 23404737]
- Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, et al. Recommendations for Incorporating Patient-Reported Outcomes Into Clinical Comparative Effectiveness Research in Adult Oncology. *Journal of Clinical Oncology*. 2012; 30(34):4249–4255. [PubMed: 23071244]
- Beck CT, Gable RK. Item response theory in affective instrument development: an illustration. *Journal of Nursing Measurement*. 2001; 9(1):5–22. [PubMed: 11469142]
- Bjorner JB, Rose M, Gandek B, Stone AA, Junghaenel DU, Ware JE Jr. Difference in method of administration did not significantly impact item response: an IRT-based analysis from the Patient-Reported Outcomes Measurement Information System (PROMIS) initiative. *Quality of Life Research*. 2013
- Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA*. 2013; 309(8):814–822. [PubMed: 23443445]
- Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, 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]
- Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, 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–S11. [PubMed: 17443116]
- Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research*. 2013; 13:211. [PubMed: 23758898]
- Choi SW, Schalet BD, Cook KF, Cella D. Establishing a common metric for depressive symptoms: Linking BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychological Assessment*. (In press).
- Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. *JAMA*. 2002; 288(23):3027–3034. [PubMed: 12479768]
- DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. *Medical Care*. 2007; 45(5 Suppl 1):S12–21. [PubMed: 17443114]
- Eisenstein EL, Diener LW, Nahm M, Weinfurt KP. Impact of the Patient-Reported Outcomes Management Information System (PROMIS) upon the Design and Operation of Multi-center Clinical Trials: a Qualitative Research Study. *Journal of Medical Systems*. 2011; 35(6):1521–1530. [PubMed: 20703765]
- Ellwood PM. Shattuck Lecture - Outcomes Management - a Technology of Patient Experience. *New England Journal of Medicine*. 1988; 318(23):1549–1556. [PubMed: 3367968]
- FDA; Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health and Quality of Life Outcomes*. 2006; 4:79. [PubMed: 17034633]
- Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. *Journal of Rheumatology*. 2011; 38(8):1759–1764. [PubMed: 21807798]
- Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-

- Life Questionnaire C30. *Journal of Clinical Oncology*. 2004; 22(17):3485–3490. [PubMed: 15337796]
- Gershon R, Rothrock NE, Hanrahan RT, Jansky LJ, Harniss M, Riley W. The development of a clinical outcomes survey research application: Assessment Center. *Quality of Life Research*. 2010; 19(5):677–685. [PubMed: 20306332]
- Guyatt GH, Ferrans CE, Halyard MY, Revicki DA, Symonds TL, Varricchio CG, et al. Exploration of the value of health-related quality-of-life information from clinical research and into clinical practice. *Mayo Clinic Proceedings*. 2007; 82(10):1229–1239. [PubMed: 17908529]
- Hendriks MG, Schouten HC. Quality of life after stem cell transplantation: a patient, partner and physician perspective. *Eur J Intern Med*. 2002; 13(1):52–56. [PubMed: 11836083]
- Henly SJ. The promise of PROMIS. *Nursing Research*. 2010; 59(2):77. [PubMed: 20216009]
- Laugsand EA, Sprangers MA, Bjordal K, Skorpen F, Kaasa S, Klepstad P. Health care providers underestimate symptom intensities of cancer patients: a multicenter European study. *Health and Quality of Life Outcomes*. 2010; 8:104. [PubMed: 20858248]
- Lerdal A, Kottorp A, Gay C, Aouizerat BE, Portillo CJ, Lee KA. A 7-item version of the fatigue severity scale has better psychometric properties among HIV-infected adults: an application of a Rasch model. *Quality of Life Research*. 2011; 20(9):1447–1456. [PubMed: 21547359]
- Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: Challenges and opportunities. *Quality of Life Research*. 2009; 18(1):99–107. [PubMed: 19034690]
- Magasi S, Ryan G, Revicki D, Lenderking W, Hays RD, Brod M, et al. Content validity of patient-reported outcome measures: perspectives from a PROMIS meeting. *Quality of Life Research*. 2012; 21(21866374):739–746. [PubMed: 21866374]
- McHorney CA. Generic health measurement: past accomplishments and a measurement paradigm for the 21st century. *Annals of Internal Medicine*. 1997; 127(8 Pt 2):743–750. [PubMed: 9382391]
- Perez L, Huang J, Jansky L, Nowinski C, Victorson D, Peterman A, et al. Using focus groups to inform the Neuro-QOL measurement tool: exploring patient-centered, health-related quality of life concepts across neurological conditions. *Journal of Neuroscience Nursing*. 2007; 39(6):342–353. [PubMed: 18186419]
- Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. *Assessment*. 2011; 18(21697139):263–283. [PubMed: 21697139]
- PROMIS. Instruments Available for Use in Assessment Center. Department of Medical Social Sciences, Northwestern University; 2013.
- Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, 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(17443115):22–31.
- Stone AA, Broderick JE, Kaell AT. Single momentary assessments are not reliable outcomes for clinical trials. *Contemporary Clinical Trials*. 2010; 31(5):466–472. [PubMed: 20580945]
- Thomas ML. The value of item response theory in clinical assessment: a review. *Assessment*. 2011; 18(3):291–307. [PubMed: 20644081]
- van Alphen A, Halfens R, Hasman A, Imbos T. Likert or Rasch? Nothing is more applicable than good theory. *Journal of Advanced Nursing*. 1994; 20(1):196–201. [PubMed: 7930122]
- Waltz, CF.; Strickland, OL.; Lenz, ER. *Measurement in Nursing and Health Research*. 4. Springer Publishing; 2010.
- Official Records of the World Health Organization; Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference; New York. 19–22 June, 1946; 1946. p. 100 signed on 22 July 1946 by the representatives of 61 States and entered into force on 7 April 1948
- Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System–Cancer scales in advanced-stage cancer patients. *Journal of Clinical Epidemiology*. 2011; 64(21447427):507–516. [PubMed: 21447427]
- Zerhouni E. *Medicine*. The NIH Roadmap. *Science*. 2003; 302(5642):63–72. [PubMed: 14526066]

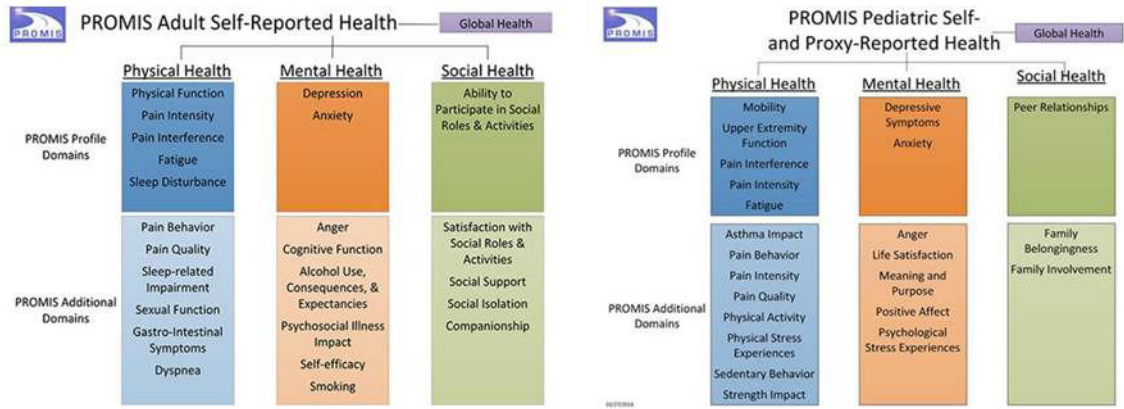


Figure 1.

Assessment CenterSM

[Register New User](#)
[Forgot Password](#)

Resources

- [Instrument Library](#)
- [Scoring Manuals](#)
- [Workshops](#)
- [Publications](#)
- [Presentations](#)
- [User Manuals](#)
- [Glossary](#)
- [Terms and Conditions](#)
- [FAQ](#)
- [Release History](#)
- [About Us](#)
- [Video Tutorials](#)

Contact Us: [Email](#)
 Help Line: 877-283-0596

© 2007-2013
 David Cella, Ph.D.,
 Richard Gershon, Ph.D.,
 Michael Bass, M.S. &
 Nan Rothrock, Ph.D.

WHAT IS ASSESSMENT CENTERSM

Assessment CenterSM is a free, online data collection tool that enables researchers to create study-specific websites for capturing participant data securely online.

Studies can include measures within the Assessment Center library as well as custom instruments entered by the researcher. The instrument library includes self- and proxy-report short forms, computerized adaptive tests (CATs), and batteries or profiles from:


- Patient-Reported Outcomes Measurement Information System (PROMIS)
- Quality of Life in Neurological Diseases (Neuro-QOL)
- NIH Toolbox

Among other important features, Assessment Center also enables:

- Downloading library instruments for administration on paper
- Customization of items or instruments (e.g., format, randomization, skip patterns)
- Real-time scoring of CATs and short forms
- Storage of protected health information (PHI) in a separate, secure database
- Automated accrual reports
- Real-time data export
- Ability to capture endorsement of online consent forms

Demonstration CAT

Try out a PROMIS Computerized Adaptive Test and get an instant score report. CATs include Emotional, Physical, and Social Health. [Click here to begin.](#)



Request PDFs of PROMIS Short Forms

Sponsors

- [PROMIS](#)
- [Neuro-QOL](#)
- [NIH Toolbox](#)

Studies of Interest

- [PROsetta Stone](#)

Assessment Center reserves the right to have regular scheduled maintenance windows. Weekly maintenance may be performed on Sunday mornings from 7am to 9am CST. During this time, Assessment Center may be unavailable. We apologize for any inconvenience this may have. If you have any questions, please [contact us](#).

Figure 2.

Table 1

Legacy measures linked to PROMIS Domains available in PROsetta Stone

Domain	Legacy Measure
PROMIS Anger	BPAQ
PROMIS Anxiety	MASQ
	SF-36/Mental Health
	GAD-7
	K6
PROMIS Depression	Neuro-QOL Anxiety
	CES-D
	SF-36/Mental Health
	PHQ-9
PROMIS Fatigue	Neuro-QOL Depression
	FACIT-Fatigue
	SF-36/Vitality
PROMIS Pain	BPI Severity
	BPI Interference
PROMIS Physical Function	HAQ-Disability Index
	SF-36/Physical Function
	Neuro-QOL Mobility
	Neuro-QOL Upper Extremity

Note: BPAQ – Buss-Perry Aggression Questionnaire; MASQ – Mood and Anxiety Symptom Questionnaire; SF-36 – Short Form (36); GAD-7 – General Anxiety Disorder 7-item; K6 – Kessler 6 Mental Health Scale; Neuro-QOL Anxiety – Quality of Life in Neurological Disorders; CES-D – Center for Epidemiologic Studies-Depression Scale; PHQ-9 – Patient Health Questionnaire; FACIT – Functional Assessment of Chronic Illness Therapy; BPI – Brief Pain Inventory; HAQ – Health Assessment Questionnaire.

Adapted from <http://www.prosettaStone.org/LinkingTables/Pages/default.aspx>.