Patient-Reported Outcomes Measurement Information System (PROMIS): Efficient, standardized tools to measure self-reported health and quality of life

Margaret Bevans, PhD, RN[a,], Alyson Ross, PhD, RN[b], David Cella, PhD[c,d]*

[a] Nursing Research and Translational Science, Nursing Department, National Institutes of Health
Clinical Center, Bethesda, MD
[b] Nursing Research and Translational Science, Nursing Department, NIH Clinical Center, Bethesda, MD
[c] Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL
[d] Center for Patient-Centered Outcomes, Institute for Public Health and Medicine, Chicago, IL

ARTICLE INFO

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 article was to summarize the development and validation of the PROMIS measures and to describe its current functionality as it relates to nursing science.

Cite this article: Bevans, M., Ross, A., & Cella, D. (2014, OCTOBER). Patient-Reported Outcomes Measurement Information System (PROMIS): Efficient, standardized tools to measure self-reported health and quality of life. Nursing Outlook, 62(5), 339-345. http://dx.doi.org/10.1016/j.outlook.2014.05.009.

Introduction

All nurses, whether clinicians, researchers, or academicians, are interested in the effects of diseases and treatments on individuals. Although physiological data provide valuable information about what is occurring within one's body, it is difficult to capture accurately subjective attitudes, values, and experiences. Exploring the individual's perceptions is important because 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; Food and Drug Administration, 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: (a) the multiple questionnaires measuring the same concept, (b) the length and difficulty of many of the measures, (c) the variation in psychometric quality, (d) the difficulty in comparing or combining data across different studies and populations, and (e) 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 PROs and addressing associated problematic issues. Since its launch, PROMIIS 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 article 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 physical, mental, and social framework of health (World Health Organization, 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 on 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