UCLA
UCLA Previously Published Works

Title

A Patient Reported Outcome Ontology: Conceptual Issues and Challenges Addressed by the Patient-Reported Outcomes Measurement Information System® (PROMIS®)

Permalink

https://escholarship.org/uc/item/3m22g11f

Authors

Cella, David
Hays, Ron D

Publication Date

2022

DOI

10.2147/prom.s371882

Peer reviewed

A Patient Reported Outcome Ontology: Conceptual Issues and Challenges Addressed by the Patient-Reported Outcomes Measurement Information System® (PROMIS®)

David Cella [ID][1], Ron D Hays [ID][2]

[1]Department of Medical Social Sciences, Northwestern University, Chicago, IL, USA; [2]Department of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Correspondence: David Cella, Email d-cella@northwestern.edu

Abstract: We briefly review the history of measuring perceptions of health and quality of life, followed by an examination of conceptual issues related to terminology that have led to potentially conflicting ontologies. Then, we discuss challenges posed by the lack of consensus on common meaning and the proliferation of measures. Next, we suggest a solution grounded in an ontology adopted by the National Institutes of Health (NIH) funded Patient-Reported Outcomes Measurement Information System (PROMIS) project. We conclude by discussing issues associated with mapping the PROMIS domain framework onto other familiar ontologies and recommend a way forward for PROMIS to provide a sustainable ontological structure to enable coherent common measurement.

Keywords: patient-reported outcomes, quality of life, health-related quality of life, ontology, PROMIS®, patient reported outcomes measurement information system

Introduction

Patient reports are a critically important aspect of documenting the impact of illness and health care interventions on individual's daily functioning and well-being. Medicine has a long history of emphasis on the "objective" or observable aspects of disease where the "patient" was considered a passive host. However, there is now recognition of the important need for a patient-centered approach to monitor disease and assess health care outcomes. Patient-*centered* medicine requires patient-*reported measures* that provide information essential to characterizing outcome heterogeneity for individuals with the same condition, and outcome comparability for individuals across different conditions.

Although it took nearly a century to be implemented, the modern era of patient-centeredness was initiated by Sir William Osler in the 19th century: "It is much more important to know what sort of a patient has a disease than what sort of a disease a patient has".[1] As we consider the ontology of patient-reported outcomes (PROs), we reflect on the history and evolution of the terminology and its aspects, including how they aggregate into higher-order components reflecting health perceptions. An examination of ontology leads to an appreciation of the value of consensus on the existence of these higher-order components and measurement of their many aspects. Indeed, generating consensus is critical if such an ontology is to be adopted to elevate the patient perspective to serious consideration of treatment outcomes.

In this paper, we briefly review the history of measuring perceptions of health and quality of life, followed by an examination of conceptual issues regarding terminology that have led to potentially conflicting ontologies. Then, we discuss challenges posed by the lack of consensus on common meaning and the proliferation of measures. Next, we suggest a solution grounded in an ontology adopted by the National Institutes of Health (NIH) funded Patient-Reported Outcomes Measurement Information System (PROMIS) project.[2] We conclude by discussing issues associated with

mapping the PROMIS domain framework onto other familiar ontologies and recommend a way forward for PROMIS to provide a sustainable ontological structure to enable coherent common measurement.

Derivation of the term "Patient-Reported Outcome" (PRO).

Relative to the work done in this area, the term PRO is new. It was originally proposed by the US Food and Drug Administration (FDA) as a neutral term to reflect the full range of treatment effects best reported directly from patients. The FDA definition of a PRO as "any report coming from patients about a health condition and its treatment" has been interpreted to include health behaviors such as adherence to doctor recommendations, smoking, and exercise. It also includes treatment preferences, perceptions of health care (including perceptions of benefit and harm), and participation in health care decisions. While a wide range of patient reports are important, this discussion of PROs is confined to a subset consistent with the precursor terms of "quality of life" (QoL) and "health-related quality of life (HRQoL). QoL is a broader concept than HRQoL by virtue of including a full range of environmental and personal factors that contribute to a sense of value in living. This includes health, certainly, but also climate, air quality, crime, wealth, food availability, and other contributions to quality of life. HRQoL as a concept emphasizes life quality as influenced by health. Although it is impossible to fully ignore other influences on QoL, HRQoL assessment aims to focus on the more salient health influences. This includes symptoms such as pain, fatigue, depression, and anxiety, side effects of treatment, functional level, and general perceptions of health and well-being. Thus, PROs are restricted in this proposed ontology to HRQoL and do not include patient reports about health behaviors such as medication adherence, substance use, and treatment preferences.[3]

Why is an ontology needed for patient reported outcomes?

PROs derive from a decades-long legacy of research examining the HRQoL associated with various diseases and their treatments. The myriad of HRQoL studies have deployed a wide range of survey instruments that ask different questions, often tapping different concepts from one another, but all drawing conclusions about HRQol or QoL. This has resulted in confusion, and at times disagreement, about what is HRQoL. The absence of a shared ontology for the concept has rendered it less useful in considerations of clinical benefit and policy impact. By contrast, a shared ontology, including an understanding of its components, would render HRQoL a more potent consideration in evaluating health care than it has been to date. In an era where patient-centeredness has become a priority, the need for a common language for PROs and HRQoL has become paramount.

Relevant Historical Ontologies

In 1948, the World Health Organization (WHO) publicly released a definition of health as "a state of complete physical, mental, and social well-being and not merely the absence of disease and infirmity".[4] This definition opened the door to move away from the notion that health could be defined simply as the absence of disease, a notion that had dominated medicine prior to World War II. It emphasizes the full spectrum of quality of life, with three primary components (physical, mental, social). Several groups, working independently, have confirmed the factorial structure of physical and mental health as separate-but-related concepts, but evidence for a distinct social health component is mixed.[5–8] Indeed, the social health aspect of the WHO has been viewed with skepticism. Larson[8] noted that "The RAND Health Insurance Experiment based its measures of health on the WHO definition but was critical of the social component of the definition. Their data did not confirm the existence of social well-being as an independent dimension of health". He argues for psychosocial health that combines mental and social health. However, the failure to identify a separate social health domain may be due to insufficient questions asking about social health and may require more appropriate measures.[9]

In the decades that followed the 1948 WHO definition, it was noted that it continued to imply that the absence of disease or disability connoted better health. Recognizing the prevalence of chronic illness and disability in the world, and in the interest of placing all people regardless of disease status on level ground, WHO began to revise its thinking about health. By 1984, its definition of health evolved to "the extent to which an individual or group is able to realize aspirations and satisfy needs and to change or cope with the environment".[4] A series of health promotion conferences, the first being in Ottawa, Canada (http://www.who.int/healthpromotion/conferences/previous/ottawa/en/), ensued, with the goal of enabling people to improve their lives by taking increased control over their health. To reach that "state of complete physical, mental and social well-being", people must be able to identify their aspirations, satisfy their needs,

and adapt or cope with life demands and the environment. The concept of health was thereby transformed into a positive concept emphasizing the balancing of social and personal resources with physical capacity.

Parallel to this shift in conceptualizing health as a positive aspiration, WHO developed its International Classification of Functioning, Disability, and Health (ICF). Designed as a complement to the more familiar and widely used International Classification of Diseases (ICD) series, ICF is a tool that provides internationally comparable information about the experience of health and disability. As indicated in Figure 1, ICF disability and functioning are viewed as outcomes of interactions between health conditions (diseases, injuries, etc.) and contextual factors, including environmental, physical, social, attitudinal, and personal influences.[4] Body functions are physiological and psychological functions of body systems. Body structures refer to the physical anatomy. Impairments are problems in body structure or function deviant from what is typical. Activity refers to one's execution of a task or intended action. Activity limitation can result from any of the model components. Finally, participation refers to one's involvement in a life situation. Participation restriction may also occur due to any of the model components. The ICF depicts limitations associated with disabilities to be largely a socially created problem.[10]

The ICF framework has been useful at the individual, institutional, and societal level, and it has helped in health policy development and planning, economic analysis, and research.[11] However, it is not well suited as an ontology for PROs for two reasons. First, its scope is much broader, as it includes physiological and anatomical aspects of a person, and second, because it includes environmental factors, also outside the scope of PROs or HRQoL.

Wilson and Cleary[12] proposed a model of HRQoL that depicts a causal chain with influential environmental and individual characteristics (Figure 2). The causal chain begins with biological and physiological influences from disease and treatment, which cause symptoms that may or may not impair functional status. Decrements in functional status, in turn, can produce decline in general or overall health status. Together, these components of the model influence a person's global evaluation of life quality as it relates to one's health. Each of these components (symptoms, function, general health, and quality of life) can be influenced by individual characteristics (personality, stoicism, amplification, values, etc.) and by environmental characteristics (social and psychological support, etc.). While the Wilson and Cleary model has been well received by the HRQoL community, it does not detail the ontology of symptoms, function, general health perceptions, and quality of life.

Valderas and Alonso[13] proposed a classification system for PROs that combined the ICF framework and Wilson and Cleary Model (Figure 3).

Using a common set of concepts consistently identified in the two classification systems, Valderas and Alonso[13] developed a cross-classification system and mapped several commonly used PRO questionnaires to it. While this illustrated how various PRO questionnaires could be mapped onto a descriptive system, it did not solve the fundamental problem of specifying the content of the subordinate domains within the classification system. For
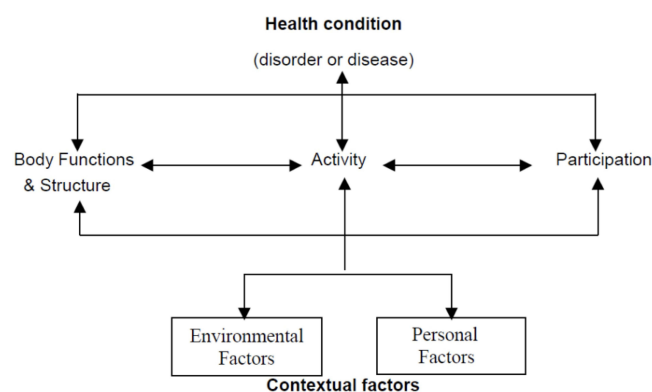
**Figure 1** The International Classification of Functioning, Disability, and Health (WHO, 2002).
Note: Reprinted from the World Health Organization. Towards a Common Language for Functioning, Disability and Health ICF, Geneva, Switzerland; Copyright 2002. Available from: https://cdn.who.int/media/docs/default-source/classification/icf/icfbeginnersguide.pdf.[4]
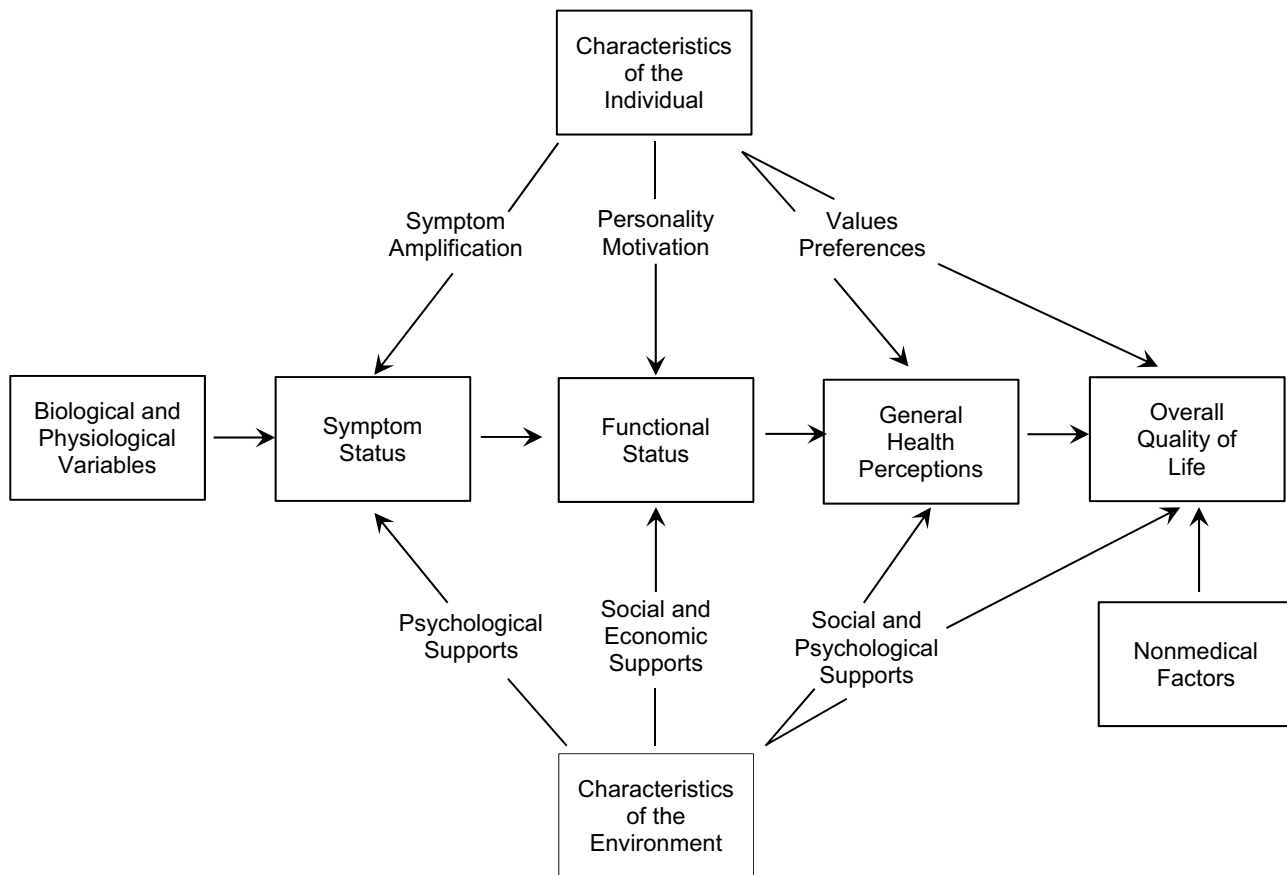
Figure 2 Wilson and Cleary (1995) model of health-related quality of life.
Note: Reproduced with permission from *JAMA*, 1995;273(1):59-65. Copyright© (1995) Wilson IB, Cleary PD. Linking Clinical Variables With Health-Related Quality of Life: A Conceptual Model of Patient Outcomes. *American Medical Association*. All rights reserved.[12]

example, it did not articulate common symptoms or functional domains in terms of their nature, boundaries, and relationship to one another. This articulation is needed to unify measurement into a common language at the subordinate domain level.

In summary, there is general agreement that there are three primary spheres of life (physical; mental; social), and that there is often a replicable causal path from disease and treatment physiology to symptoms, then to functioning, and then to broader perceptions of health and quality of life (albeit influenced by various individual and environmental character-istics). However, these general classification categories have not been sufficiently detailed about their content and subordinate components to enable researchers and policy-makers to coalesce around a common ontology. In fact, the opposite has been the case. Over the past 50 years, a vast number of PRO questionnaires have been developed and published to measure physical, mental, and social health. The PROQOLID database identifies over 4000 patient-centered clinical outcome assessments, the majority of which are PRO measures.[14] These assessments, and others not included in that database, utilize overlapping terminology that is often not reflected in the content of the questions. For example, among the many measures of pain, some reflect pain intensity, others reflect pain frequency or duration, and still others reflect the extent to which pain interferes with function or HRQoL. In some cases, these questions are separate, while in others they may be bundled together into a single "pain" score. As a result, pain is measured in different ways despite adoption of the same terminology. This makes it very difficult to evaluate pain as an endpoint in systematic reviews and meta-analyses within and across therapeutic areas. The same is true for other symptoms, such as fatigue, depression, anxiety, and dyspnea, to name a few. Similarly, there is a wide range of ability and limitation captured in physical function including mobility (ambulation; self-transport), upper extremity function (range of motion; dexterity), flexibility,
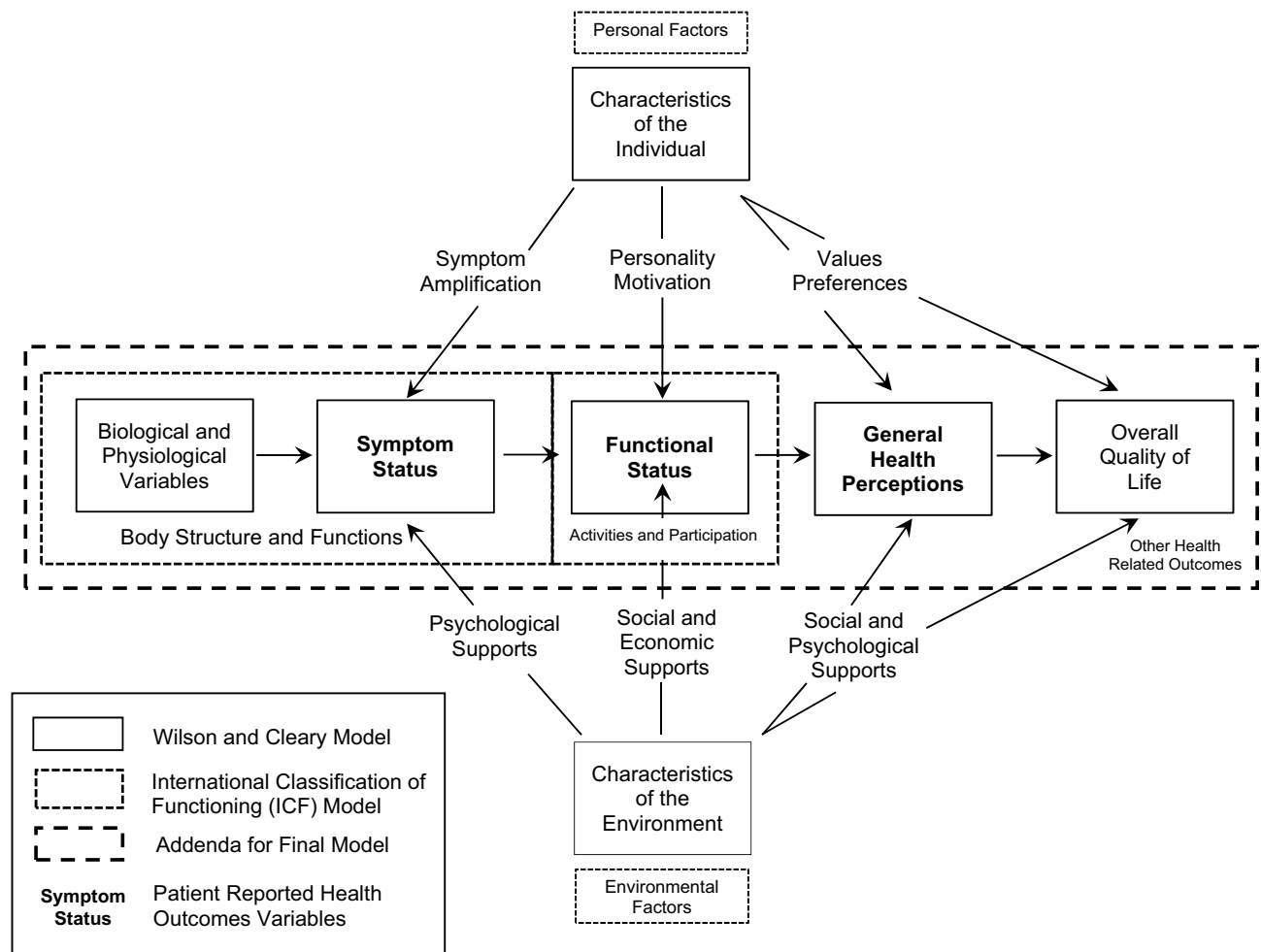
Figure 3 Valderas and Alonso (2008) proposed classification system.
Note: Reprinted by permission from *Springer: Quality of Life Research,* Patient reported outcome measures: a model-based classification system for research and clinical practice. Valderas JM, Alonso J. 2008;17:1125-1135.[13]

and activity level. Therefore, when one reports on improved physical function, it is necessary to identify the specific questions asked to identify the nature of that improvement.

There is a family of preference-based measures that summarize the important domains using a health utility score anchored by 0 (dead or a health state as bad as dead) and 1 (a health state equivalent to perfect health).[15,16] Each of these systems is based on a health classification system that weighs the life domains according to local community or societal values, to produce that utility value, which in turn can be used to modify survival time for its quality (eg, quality adjusted life year). While useful for economic and policy review, none of these systems provides the ontology for HRQoL that is the focus of this manuscript.

The implicit approach to PRO development has been one of "letting a thousand flowers bloom", justified as it encouraged innovation and improvement over existing assessment options. As a result, we now have a plethora of PRO instruments available, including general (so-called "generic") questionnaires and measures targeted to specific diseases and treatments. Unfortunately, there has been a significant downside. As a result of the organic evolution of PRO measures developed over the past 50 years, an evolution lacking an accepted, common terminology, we have different concepts being measured but using the same terms, similar concepts measured using different terms, and a resulting "Tower of Babel".[17] We are now at a point where it would be a major advance in the field of PRO measurement if the PRO developers and end-users were to adopt an ontology with a common language and associated common metrics.

The PROMIS Framework as an Ontology for PROs

In 2004, under its Roadmap (now Common Fund) initiative, the NIH launched a cooperative group of PRO measurement scientists to address the issues identified above. This initiative, PROMIS, was charged with improving assessment of self-reported symptoms and other domains of health across a wide range of chronic diseases. PROMIS was situated to be a translational and ideally transformational arm of the larger NIH effort: "Re-engineering the Clinical Research Enterprise." Aiming to standardize assessment of common self-reported symptoms and function across the full range of chronic disease, the use of item response theory (IRT) was proposed as a solution to the conflicting array of measurement within and across disease.[18–22] IRT allows one to build banks of questions, each of which is a representation of the underlying domain (trait). Once the assumptions underlying IRT are met, and the fit of the model is satisfactory, the items in the bank can be used interchangeably.[23] This provides great flexibility in the selection of specific items for specific settings while at the same time providing a score on a single, common metric. In addition, IRT enables computerized adaptive testing which customizes assessment to the level of the individual responding to the questions, thereby enhancing measurement precision (reliability). Relevant to the focus of this manuscript, the building of these item banks, and the testing of their measurement properties, requires careful definition of terms and even more careful articulation of the specific domains that are included in the PROMIS framework.

In the early months of the 10-year PROMIS project, armed with literature review and their measurement experience, investigators discussed what would be the optimal framework under which to articulate self-reported health domains and items. They considered both WHO definitions, the ICF, the Wilson and Cleary Model, and other options, and decided that the original 1948 WHO health definition was best for its purposes. This was because, as mentioned, the ICF framework and the Wilson and Cleary model included components outside the scope of our charge, which was focused on self-reported health outcomes. By contrast, the WHO definition of physical, mental, and social well-being provided a useful, open starting point.[2] From the beginning, PROMIS investigators recognized that these three components of health are not mutually exclusive. Pain, situated under physical health, has mental and even social components. Sleep is both physical and mental. Sexual function includes physical, mental, and social aspects. Nevertheless, they partitioned symptoms, function, and more general health perceptions into one or another of these broad components and proceeded to define and measure each domain.[23,24]

PROMIS domain teams were established and deployed across physical, mental, and social health, with each embarking on the process of adopting, modifying, and writing items according to standard protocols for mixed (qualitative and quantitative) methods.[2,25–29] Over the course of the first two years of PROMIS funding, several meetings of steering committee members, NIH program officers, and PROMIS investigators were held to develop consensus that was based on literature reviews conducted by the domain teams and emerging qualitative data. Discussions included optimal terminology that would be suitable for the clinical research community, review of evidence in support of and in contradiction to prevailing scientific opinions, and input from patient interviews regarding candidate domain language. A "domain framework" committee gathered the emerging evidence and built out an articulated array of important health domains organized under the broad headings of physical, mental, and social health. A guiding principle through this process was the consideration of important self-reported health concepts that could be considered sufficiently unidimensional for measurement purposes. Common terminology with domain definitions under this domain framework were then tested and adopted with supporting evidence, using responses obtained from large-scale testing. The result was a PROMIS domain framework that now includes over 100 domains, each with a specific definition and item set that has been demonstrated to reliably assess a single underlying concept. Anticipating the interest of researchers to focus on a core set of domains from this longer list, the PROMIS steering committee identified seven domains to place into a self-reported health profile: Pain, fatigue, depressive symptoms, anxiety, sleep disturbance, social function, and physical function. Supplementary Table 1, appended to the manuscript, provides a list of the PROMIS domains and their definitions as of July, 2022, each consisting of a set of calibrated items. This provides a possible ontology for adoption and use by others seeking a more standardized reporting and communication. Supplementary Table 1, from left to right, reflects the respondent (adult; child; proxy), the category, or sphere of health (physical, mental, social), and the second-level sub-categories of each broad category. Physical health is divided into physical symptoms and physical function.

Mental health includes affect, behaviors, and cognition, and social health includes social function and social relationships (family, friends, etc). Domains and sub-domains are narrower components of the sub-categories, typically the basis from which unidimensional item banks are derived. Each of these domains and sub-domains is defined in the far-right column of Supplementary Table 1, providing working definitions for the proposed PRO ontology.

Standardized Measurement Development and Evaluation

A statistical plan was developed to guide analyses of the psychometric properties of the PROMIS measures. Item response theory (IRT) is a family of measurement models that help improve the accuracy of tests by modeling the relationship between latent (unobservable) traits (such as depression or pain) and the responses made to questions or items meant to tap into that latent trait. Analyses of response data collected by the PROMIS group evaluated the assumptions of IRT (monotonicity, sufficient unidimensionality and local dependence), IRT model fit, item fit, and differential item functioning.[30] Results of analyses were provided on an iterative basis to domain teams leading the development of the measures to finalize item banks. Analysts and content experts discussed each item, and decisions were made as to whether an individual item should be calibrated and included in item banks; not calibrated but retained for possible future calibration; or excluded from further consideration. The final selected item banks were calibrated using the IRT graded response model.[30,31] To identify the metric for the PROMIS item parameter estimates, the scale for person parameters was fixed relative to the United States general population with a mean of 50 and SD of 10.[32] The reliability (scale information) and validity (including responsiveness to change) of the PROMIS measures have been extensively evaluated.[19–29,32–35]

Linking to Legacy Measures

If PROMIS is to become a unifying ontology for HRQoL assessment, it will be important for those who prefer other measure alternatives to suspend that preference and consider either switching to PROMIS or demonstrating that scores on their preferred questionnaire can be expressed on the common PROMIS metric. The latter option can be achieved through various linking procedures. When two measures are sufficiently correlated with one another and define a common under-lying continuum; they can be empirically linked on a common metric.[19–21,23] For example, the PROMIS global physical and mental health scores were empirically linked to the legacy Veterans RAND 12-item general health survey. The link makes it possible to estimate the legacy scores (RAND physical and mental health scores) directly from the PROMIS scores (global health) and vice versa.[21] Much of this work has been done for commonly used legacy measures in physical, mental, and social health, with more than 60 linking tables and reports posted at www.prosettastone.org.

Indeed, agreeing on an ontology is not necessarily tantamount to agreeing on a specific assessment choice. PROMIS is available for use in this regard, but so too are several other measures of domains in the PROMIS ontology. Some important advantages are conferred by using IRT or other methods to link a legacy measure (or any other preferred measure, for that matter) to a PROMIS item bank. These include the advantages of comparability and common language, both of which can accelerate scientific advances from a clearer, well-defined, and shared understanding of HRQoL.

Reconciling Ontologies

The ontology offered by the PROMIS framework has been mapped to the WHO ICF model.[36,37] At the time this was done, all 1006 PROMIS items could be mapped to ICF concepts. PROMIS item banks mapped to multiple ICF codes indicating one-to-one, one-to-many and many-to-one mappings between PROMIS and ICF second-level classification codes.[37] Most major concepts in the ICF Body Functions (b) and Activity & Participation (d) components were found to be supported by PROMIS. Custom instruments to measure ICF codes can therefore be developed. Given PROMIS focuses on PROs, ICF concepts related to body structures and many of the ICF concepts related to body functions and environmental factors have limited coverage in PROMIS. As a result of this mapping, users interested in measuring ICF content that is captured by self-report can use PROMIS to do so. Similarly, the PRO elements captured in the Wilson and Cleary[12] model can be measured with PROMIS instruments, or legacy instruments that have been linked to the PROMIS metric.

Sustaining and Updating a PROMIS Ontology Over Time

After supporting the development and initial validation of PROMIS for 10 years (2004–2014), the NIH supported a 4-year "Person Centered Assessment Resource." The explicit purpose of this resource was to transition PROMIS and three other NIH-supported measurement systems [Quality of Life in Neurological Disorders (Neuro-QoL), NIH Toolbox for Assessment of Neurological and Behavioral Function (NIH Toolbox), and Adult Sickle Cell Quality of Life Measurement System (ASCQ-Me)] into a sustaining resource for wide use. Independence from government support was deemed critical to sustainability. Northwestern University received the 4-year award to build this sustaining infrastructure which would make these tools as freely available as possible while supporting the necessary human resources to curate, distribute, and update with improvements over time. With the ending of this award in 2018, Northwestern has assumed management and distribution responsibility for the PROMIS item banks through the "HealthMeasures" website (www.healthmeasures.net). From this site, anyone can download PROMIS fixed short-form assessments in English and Spanish. A free scoring service will generate and return IRT-based scores. Computerized Adaptive Testing capability is provided through an application–programming interface (API), which is licensed to include technology-supporting fees either for individual use or for distributors. Languages other than English are provided with distribution fees that can be waived for academic use. This model has sustained the costs of curating and distributing PROMIS in the years since public funding has ceased. In addition to this, the 18 principal investigators of the PROMIS Common Fund effort started a charitable [501(c)(3)] foundation with the purpose of educating people about PROMIS and relieving the burden of government. This foundation, The PROMIS Health Organization (PHO), now has several hundred members and conducts educational workshops and an annual international meeting. Together, the partnership between Northwestern University and the PHO provides a stable base to secure the future of PROMIS.

Acknowledgments

Disclosure

Dr David Cella reports grants from the National Institutes of Health, during the conduct of the study, and he is an uncompensated board member, PROMIS Health Organization (501(c)(3) charitable organization). The authors report no other conflicts of interest in this work.

References

1. Osler W. *The Principles and Practice of Medicine: Designed for the Use of Practitioners and Students of Medicine.* New York: D. Appleton & Co; 1892.
2. 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;45(5 Suppl 1):S3–S11. doi:10.1097/01.mlr.0000258615.42478.55
3. Fung CH, Hays RD. Prospects and challenges in using patient-reported outcomes in clinical practice. *Qual Life Res*. 2008;17:297–302. doi:10.1007/s11136-008-9379-5
4. World Health Organization. Towards a common language for functioning, disability and health ICF, Geneva, Switzerland; copyright; 2002. Available from: https://cdn.who.int/media/docs/default-source/classification/icf/icfbeginnersguide.pdf. Accessed April 22, 2022.
5. Carle AC, Riley W, Hays RD, Cella D. Confirmatory factor analysis of the Patient Reported Outcomes Measurement Information System (PROMIS) adult domain framework using item response theory scores. *Med Care*. 2015;53(10):894–900. doi:10.1097/MLR.0000000000000413
6. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS®-29 v2.0 profile physical and mental health summary scores. *Qual Life Res*. 2018;27(7):1885–1891. doi:10.1007/s11136-018-1842-3
7. Hays RD, Stewart AL. The structure of self-reported health in chronic disease patients. *Psychol Assess*. 1990;2:22–30. doi:10.1037/1040-3590.2.1.22
8. Larson JS. The World Health Organization's definition of health: social versus spiritual health. *Soc Indic Res*. 1996;381:181–192. doi:10.1007/BF00300458
9. Carlson JA, Sarkin AJ, Levack AE, et al. Evaluating a measure of social health derived from two mental health recovery measures: the California Quality of Life (CA-QOL) and Mental Health Statistics Improvement Program Consumer Survey (MHSIP). *Community Ment Health J*. 2011;47(4):454–462. doi:10.1007/s10597-010-9347-8
10. World Health Organization. *Towards a Common Language for Functioning, Disability and Health ICF.* Geneva, Switzerland: World Health Organization; 2002.
11. Post MW, de Witte LP, Schrijvers AJ. Quality of life and the ICIDH: towards an integrated conceptual model for rehabilitation outcomes research. *Clin Rehabil*. 1999;13(1):5–15. doi:10.1191/026921599701532072

12. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. *JAMA*. 1995;273 (1):59–65. doi:10.1001/jama.1995.03520250075037

13. Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res*. 2008;17:1125–1135. doi:10.1007/s11136-008-9396-4

14. PROQOLID. Available from: https://eprovide.mapi-trust.org/. Accessed November 4, 2021.

15. Kaplan RM, Tally S, Hays RD, et al. Five preference-based indexes in cataract and heart failure patients were not equally responsive to change. *J Clin Epidemiol*. 2011;64(5):497–506. doi:10.1016/j.jclinepi.2010.04.010

16. Kaplan RM, Hays RD. Health-related quality of life measurement in public health. *Annu Rev Public Health*. 2022;43:355–373. doi:10.1146/annurev-publhealth-052120-012811

17. Mor V, Guadagnoli E. Quality of life measurement: a psychometric tower of Babel. *J Clin Epidemiol*. 1988;41(11):1055–1058. doi:10.1016/0895-4356(88)90074-1

18. Bjorner JB. Solving the tower of babel problem for patient-reported outcome measures: comments on: linking scores with patient-reported health outcome instruments: a validation study and comparison of three linking methods. *Psychometrika*. 2021;86(3):747–753. doi:10.1007/s11336-021-09778-x

19. Cella D, Choi S, Schalet B, Lai JS, Kallen M, Cook K. PROsetta Stone®: a method and common metric to link pro measures for comparative effectiveness research (CER). *Qual Life Res*. 2013;22(S1):32.

20. Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychol Assess*. 2014;26(2):513–527. doi:10.1037/a0035768

21. Schalet BD, Rothrock NE, Hays RD, et al. Linking physical and mental health summary scores from the veterans RAND 12-item health survey (VR-12) to the PROMIS(®) global health scale. *J Gen Intern Med*. 2015;30(10):1524–1530. doi:10.1007/s11606-015-3453-9

22. Schalet BD, Hays RD, Jensen SE, Beaumont JL, Fries JF, Cella D. Validity of PROMIS physical function measured in diverse clinical samples. *J Clin Epidemiol*. 2016;73:112–118. doi:10.1016/j.jclinepi.2015.08.039

23. Hays RD, Schalet BD, Spritzer KL, Cella D. Two-item PROMIS® global physical and mental health scales. *J Patient Rep Outcomes*. 2017;1(1):2. doi:10.1186/s41687-017-0003-8

24. Schalet BD, Lim S, Cella D, Choi SW. Linking scores with patient-reported health outcome instruments: a validation study and comparison of three linking methods. *Psychometrika*. 2021;86(3):717–746. doi:10.1007/s11336-021-09776-z

25. Segawa E, Schalet B, Cella D. A comparison of computer adaptive tests (CATs) and short forms in terms of accuracy and number of items administered using PROMIS profile. *Qual Life Res*. 2020;29(1):213–221. doi:10.1007/s11136-019-02312-8

26. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D; PROMIS Cooperative Group. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. *Assessment*. 2011;18(3):263–283. doi:10.1177/1073191111411667

27. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol*. 2010;63(11):1179–1194. doi:10.1016/j.jclinepi.2010.04.011

28. Cella D, Choi SW, Condon DM, et al. PROMIS® adult health profiles: efficient short-form measures of seven health domains. *Value Health*. 2019;22(5):537–544. doi:10.1016/j.jval.2019.02.004

29. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care*. 2007;45(5):S12e21. doi:10.1097/01.mlr.0000254567.79743.e2

30. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 Suppl 1):S22–31. doi:10.1097/01.mlr.0000250483.85507.04

31. Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika*. 1969;34:1–97. doi:10.1007/BF03372160

32. Liu H, Cella D, Gershon R, et al. Representativeness of the patient-reported outcomes measurement information system internet panel. *J Clin Epidemiol*. 2010;63(11):1169–1178. doi:10.1016/j.jclinepi.2009.11.021

33. Hays RD, Spritzer KL, Fries JF, Krishnan E. Responsiveness and minimally important difference for the patient-reported outcomes measurement information system (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. *Ann Rheum Dis*. 2015;74(1):104–107. doi:10.1136/annrheumdis-2013-204053

34. Irwin DE, Atwood CA Jr, Hays RD, et al. Correlation of PROMIS scales and clinical measures among chronic obstructive pulmonary disease patients with and without exacerbations. *Qual Life Res*. 2015;24(4):999–1009. doi:10.1007/s11136-014-0818-1

35. Jensen RE, Moinpour CM, Potosky AL, et al. Responsiveness of 8 Patient-Reported Outcomes Measurement Information System (PROMIS) measures in a large, community-based cancer study cohort. *Cancer*. 2017;123(2):327–335. doi:10.1002/cncr.30354

36. Tucker CA, Cieza A, Riley AW, et al. Concept analysis of the patient reported outcomes measurement information system (PROMIS(®)) and the international classification of functioning, disability and health (ICF). *Qual Life Res*. 2014;23(6):1677–1686. doi:10.1007/s11136-014-0622-y

37. Tucker CA, Escorpizo R, Cieza A, et al. Mapping the content of the Patient-Reported Outcomes Measurement Information System (PROMIS®) using the international classification of functioning, health and disability. *Qual Life Res*. 2014;23(9):2431–2438. doi:10.1007/s11136-014-0691-y

38. Zhang J, Dewitt B, Tang E, et al. Evaluation of PROMIS Preference Scoring System (PROPr) in patients undergoing hemodialysis or kidney transplant. *Clin J Am Soc Nephrol*. 2021;16(9):1328–1336. doi:10.2215/CJN.01880221