HealthMeasures

TRANSFORMING HOW HEALTH IS MEASURED

HealthMeasures: A National Person-Centered Assessment Resource

www.HealthMeasures.net

PROMIS®
Self-reported and parent-reported measures of global, physical, mental, and social health for adults and children in the general population and those living with a chronic condition

Neuro-QoL
Self-reported and proxy-reported measures of physical, mental, and social health for adults and children living with a neurological condition

ASCQ-Me℠
Self-reported measures of physical, mental, and social health for adults living with sickle cell disease

NIH Toolbox®
Performance tests of cognitive, motor, and sensory function and self-reported measures of emotional function for adults and children in the general population and those living with a chronic condition

HEALTHMEASURES:
A NATIONAL PERSON-CENTERED ASSESSMENT RESOURCE

HealthMeasures is a novel set of four measurement systems that assess physical, mental, and social health, symptoms and wellbeing, along with physical function, social function, sensory function, and cognitive function across the human lifespan. These tools allow researchers and clinicians to compare individuals and groups across diseases and conditions.

Introduction to HealthMeasures

HealthMeasures consists of four comprehensive health outcome measurement systems including PROMIS® (Patient-Reported Outcomes Measurement Information System®), Neuro-QoL (Quality of Life in Neurological Disorders), ASCQ-Me℠ (Adult Sickle Cell Quality of Life Measurement Information System℠), and NIH Toolbox® (NIH Toolbox for Assessment of Neurological and Behavioral Function®) that assess wellbeing and function. HealthMeasures aims to assess domains that are meaningful to people in their daily lives in both clinical and research settings. The assessment tools are brief and easy to understand, psychometrically sound and interpretable, and scores can be expressed on a common scale. They cover individuals from early childhood to older adulthood and can be used across various diseases, languages, literacy levels, and ethnic groups. HealthMeasures can provide valid, clinically relevant data for use in clinical care, research, or health policy development.

Flexible Administration

The methodology underpinning the development of HealthMeasures is Item Response Theory (IRT). IRT enables the creation of item banks (large groups of questions measuring a single domain, such as fatigue). This approach allows for assessment using brief, fixed length forms or more flexible computer adaptive tests (CATs). CATs allow for a tailored, computer-assisted assessment in which questions are chosen intelligently, based on responses already given. CATs produce brief and accurate scores with little burden to the person being tested.

A Common Metric

A major benefit conferred with the use of HealthMeasures is the reporting of results on a common metric, or common scale. Because different fixed length short forms and CATs from the same item bank produce scores on the same scale, HealthMeasures enables comparing your results to those of others. In addition, measures developed outside of the HealthMeasures family can often be linked to a standardized HealthMeasures metric with "PROsetta Stone®" (www.ProsettaStone.org). This enables one to 'cross talk' between measures by placing them on the same metric.

Applications of HealthMeasures

The applications of HealthMeasures are wide ranging. Healthcare practitioners can use these systems to monitor the symptoms and functioning of their patients over time, as well as evaluate the quality of the care they provide. These measurement tools can be used in clinical, comparative effectiveness, epidemiological, or health services research. Within an educational setting, HealthMeasures may also be used for screening, determining eligibility for services such as special education, or monitoring progress.

To find more information, download copies of the measures, and see demonstrations of the NIH Toolbox, go to http://www.HealthMeasures.net/search-view-measures.

PROMIS®: PATIENT-REPORTED OUTCOMES MEASUREMENT INFORMATION SYSTEM®

PROMIS is a unique measurement system developed by researchers at multiple academic institutions and the National Institutes of Health to assess health and wellbeing in adults and children from the general population as well as those living with chronic conditions.

Person-Centered Measurement Systems

PROMIS (Patient-Reported Outcomes Measurement Information System) is a comprehensive, accurate, flexible, and accessible set of tools to measure self-reported physical, mental, and social health, including symptoms, functioning, and general perceptions of health and wellbeing, in people ages 5–90. PROMIS is a result of great efforts to advance the science of measuring patient-reported health status and has been developed and validated using state of the art methods to be psychometrically sound and relevant across diverse conditions. PROMIS measures have greater precision (with the administration of fewer items) and a larger range of measurement than the majority of conventional measures. This decreases respondent burden, lowers floor and ceiling effects, and increases power in studies without increasing sample size. Featuring more than 200 measures for adults and almost 100 measures for pediatric populations, PROMIS can be used to assess symptoms and functions across lifespan, disease state, and demographic groups.

PROMIS Measures

PROMIS measures are available in short forms (made up of a fixed set of two to ten questions per domain), profiles (fixed collections of short forms from multiple domains), and as computer adaptive tests (CATs; in which questions are dynamically selected from the item bank based on the respondent's previous answers). Measures include physical health (e.g., physical function, pain, fatigue, sleep, sexual function, symptoms), mental health (e.g., depression, anxiety, anger, positive affect, cognitive function, substance use, self-efficacy for management of chronic conditions, stress experiences), and social health (e.g., ability to participate in social roles and activities, satisfaction with participation in social roles and activities, social support, isolation, family relationships). Learn more about available PROMIS measures at www.HealthMeasures.net/PROMIS.

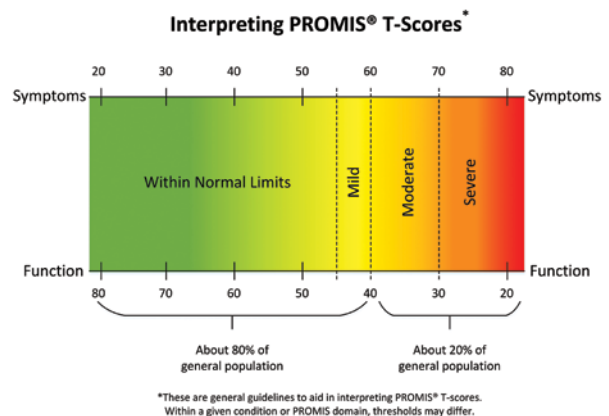### Development and Validation of Comprehensive Measures

PROMIS was constructed and validated using rigorous methods with more than 50 research protocols and over 60,000 people contributing data. Measurement development methods included a comprehensive literature review of

existing measures, focus groups testing existing and de novo questions with diverse patient groups, cognitive interviews, and measurement expert review. Psychometric testing confirmed the factor structure of each health domain and allowed for analysis at the item and bank level. Substantial qualitative and quantitative evidence supports the validity of PROMIS measures.

### Interpretable Scores and a Common Metric

One of the major benefits of PROMIS is the interpretability of scores and the ability to translate other health outcome measures to the PROMIS metric so comparisons can be made (www.ProsettaStone.org). PROMIS uses standardized scores known as T-scores which can be evaluated against a reference population (which is usually the U.S. general population). Using this metric, a score of 50 is the average of the reference population with a standard deviation of 10. Therefore, a score of 60 means the individual is one standard deviation above the reference population. Depending on the domain, this could be desirable or undesirable as high scores indicate more of the concept being measured. For example, a high fatigue score indicates high levels of fatigue whereas a high physical function score indicates better physical function.

Interpreting PROMIS® T-Scores*

*These are general guidelines to aid in interpreting PROMIS® T-scores. Within a given condition or PROMIS domain, thresholds may differ.

NEURO-QoL: QUALITY OF LIFE IN NEUROLOGICAL DISORDERS

Sponsored by the National Institute for Neurological Disorders and Stroke, Neuro-QoL is a measurement system designed to evaluate and monitor the physical, mental, and social effects experienced by adults and children living with a range of neurological conditions.

Interpreting Neuro-QoL T-Scores*
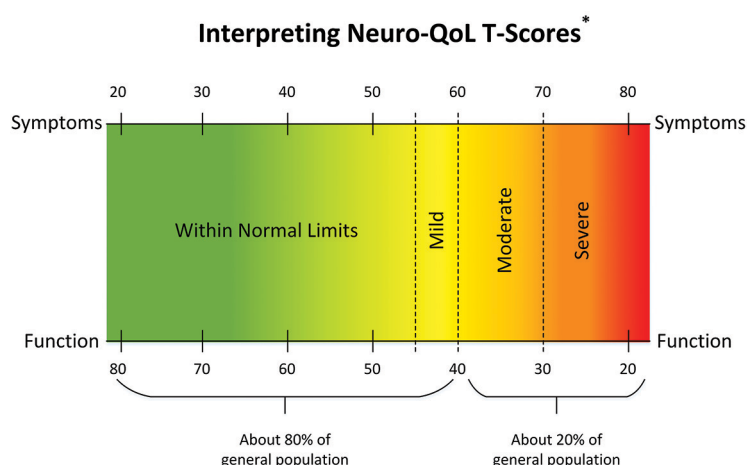
*These are general guidelines to aid in interpreting Neuro-QoL T-scores. Within a given condition or Neuro-QoL domain, thresholds may differ.

Evaluating Function and Wellbeing Across Conditions

Neuro-QoL (Quality of Life in Neurological Disorders) is a psychometrically sound measurement system developed and validated for neurological disorders using state of the science methods by researchers at Northwestern University and other academic institutions. Neuro-QoL is a set of measures, many of which are linked to PROMIS (Patient-Reported Outcomes Measurement Information System), that assess and report on symptoms, function, and wellbeing across a range of common neurological conditions, including multiple sclerosis, stroke, Parkinson's disease, epilepsy, amyotrophic lateral sclerosis (ALS), and muscular dystrophies in people ages 8 to 90. Neuro-QoL measures can be administered via paper, web, and app-based versions and are designed to minimize the burden on the patient.

Adult and pediatric versions are available and can be completed by a proxy responder when necessary. The measures can be used alongside PROMIS or NIH Toolbox instruments that assess other aspects of health and function. Learn more about Neuro-QoL at www.HealthMeasures.net/Neuro-QoL.

Holistic Person-Centered Assessment

Because living with a neurological disorder can affect numerous facets of an individual's life, Neuro-QoL evaluates aspects of physical, mental, and social health.

Physical health measures include lower extremity function and mobility, upper extremity function and fine motor skills, fatigue, and sleep disturbance. Mental health measures assess anxiety, depression, positive affect and wellbeing, emotional and behavioral dyscontrol, cognitive function, communication, and stigma. Social measures include the ability to participate in social roles and activities and satisfaction with social roles and activities. This includes one's degree of involvement and fulfillment in usual roles, responsibilities, and activities within society.

Development and Validation of a Standardized Measurement System

The Neuro-QoL development process was a collaborative multi-site research initiative. After an extensive literature review, researchers undertook two phases of in-depth expert interviews, as well as patient and caregiver focus groups and interviews. Through an iterative process, 17 content areas for adults and 11 content areas for children were selected. Measures were calibrated against the U.S. general population and clinical samples using Item Response Theory and validated in adult and pediatric clinical samples.

Interpretable Scores

Neuro-QoL uses standardized scores known as T-scores that can be evaluated against a reference population. A score of 50 is the average of the reference population with a standard deviation of 10. Therefore, a score of 60 means the individual is one standard deviation above the reference population. High scores indicate more of the concept being measured which can be desirable (e.g., ability to participate in social roles and activities) or undesirable (e.g., anxiety).

ASCQ-ME℠: ADULT SICKLE CELL QUALITY OF LIFE MEASUREMENT INFORMATION SYSTEM℠

ASCQ-Me is a patient-reported outcome measurement system developed to evaluate the physical, mental, and social wellbeing of adults with sickle cell disease. Sponsored by the National Heart, Lung, and Blood Institute, ASCQ-Me can be used by physicians and researchers to support care for affected patients.

Monitoring Effects of Sickle Cell Disease

According to the Centers for Disease Control and Prevention, sickle cell disease (SCD) affects approximately 100,000 individuals in the U.S. and 300,000 individuals worldwide. The Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-Me) is a psychometrically sound and content valid measurement system for monitoring symptoms and function in those affected by the disease. ASCQ-Me measures are appropriate for individuals ages 18 to 90 living with SCD.

As well as documenting the effects of SCD on adult functioning and wellbeing, ASCQ-Me can support treatment planning, clinical and health services research, evaluate the effectiveness of therapies, and inform programs to improve delivery of healthcare. It provides a comprehensive view of health and function when used with PROMIS or the NIH Toolbox.

Flexible Administration

ASCQ-Me measures include short forms (fixed set of five items per domain), checklists (fixed sets of descriptive items) and computer adaptive tests (CATs, in which items are dynamically selected based on the respondent's previous answers). The measures are freely available and can be delivered on paper or by computer.

Holistic Person-Centered Assessment

As well as assessing the severity of pain, ASCQ-Me measures the impact of the severity and unpredictability of pain on the person's functioning and the number and frequency of pain episodes. It also evaluates the effect of SCD on getting to sleep and staying asleep, and the severity of stiffness a person experiences. A SCD Medical History Checklist is available to record symptoms, treatments, organ and bone damage, and the quality of healthcare the individual has received. Mental and social health are taken into account through measures of the effects of SCD on social functioning and the emotional impact of SCD in terms of worry, loneliness, and depression.

Development and Validation of ASCQ-Me

ASCQ-Me was developed and validated by scientists with state of the art methods to measure wellbeing and to be complementary with physiological measures of disease severity. Development included a comprehensive literature review to identify important aspects for assessment, focus groups to create a suitable conceptual framework, and field testing of the items through geographically diverse U.S. clinics. Psychometric testing confirmed the factor structure and allowed for analysis at the item and bank level.

The validity of ASCQ-Me is based on substantial qualitative and quantitative evidence. Content validity is supported by structured interviews with patients and experts and is grounded in the life experiences documented in participant interviews.

Learn more about ASCQ-Me at www. HealthMeasures.net/ASCQ-Me.

NIH TOOLBOX® FOR THE ASSESSMENT OF NEUROLOGICAL AND BEHAVIORAL FUNCTION®
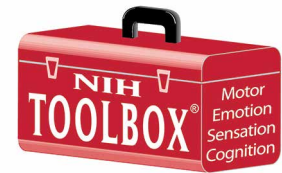
NIH Toolbox is a diverse set of measures which can be used to assess cognitive, sensory, motor, and emotional function across various study designs, settings, and conditions. It is available as an iOS app for an iPad.

Comprehensive Measures Straight to your iPad

The NIH Toolbox is a multidimensional set of brief measures that assess health and function across four domains: cognition, emotion, motor, and sensation. Each domain can be assessed using a 30-minute battery or any of the NIH Toolbox's over 100 measures can be used alone. Because it is non-disease specific and normed for ages 3 to 85, the NIH Toolbox can be used within numerous study designs with a particular emphasis on measuring outcomes in longitudinal epidemiologic studies and prevention or intervention trials. This suite of measures is cutting edge, psychometrically sound, cost efficient, and can be administered with minimal training. Measures can be utilized in a variety of settings, among individuals in the general population, and those living with chronic conditions.

Learn more at www.HealthMeasures. net/NIHToolbox.

Wide Array of Assessment

Through the utilization of Item Response Theory and computer adaptive testing, measures can be brief while remaining precise and valid. The NIH Toolbox contains two types of measures: performance based tests of function and self- or proxy-report measures.

The cognition battery assesses the mental processes involved in gaining knowledge, comprehending, communicating, problem solving and planning, and executing complex behaviors. Individual measures include attention, episodic memory, working memory, language, executive function, and processing speed. The cognition battery also includes summary scores, namely a cognitive function composite score, a fluid cognition composite score, a crystallized cognition composite score, and an early childhood composite score.

The emotion battery evaluates four domains: psychological wellbeing, stress and self-efficacy, social relationships, and negative affect. The self-report battery is recommended for ages 8 and over and there is a parent proxy version for parents of children between the ages of 3 and 12. The constructs assessed by the emotion battery include anger, fear, sadness, positive affect, general life satisfaction, meaning and purpose, perceived stress, self-efficacy, social support, companionship, and social distress. The parent proxy battery includes measures of positive affect, general life satisfaction, positive peer interaction, social withdrawal, peer rejection, empathic behaviors, self-efficacy, perceived stress, sadness, anger, and fear.

The motor battery measures an individual's ability to use and control muscles and movements. The battery appraises dexterity, strength, balance, endurance, and locomotion. The early childhood battery assesses dexterity, grip strength, standing balance, and endurance.

Finally, the sensation battery assesses sensory input and interpretation. It measures audition, vision, olfaction, pain, and taste. The early childhood sensation battery measures visual acuity and olfaction.

Development and Validation of NIH Toolbox

NIH Toolbox was developed by more than 250 contributing scientists at 80 institutions whose aim was to enhance data collection in large cohort studies and advance neurobehavioral research. More than 16,000 participants were involved in field testing, calibration, and validation. Studies were conducted across age ranges and were statistically compared to gold standard measures. Thanks to a large national standardization study, normative scores are available for targeted, accurate comparison throughout life stages.

ABOUT HEALTHMEASURES

About the Team

Many researchers, experts, and participants (including patients) contributed to the development of each of the four measurement systems included in HealthMeasures. From 2014–2018, a multi-institutional HealthMeasures team is supported by a cooperative agreement with Northwestern University (# U2C CA186878). This trans-NIH collaboration includes scientific and program support and input from 12 NIH Institutes, Centers, and Offices:

National Heart, Lung, and Blood Institute (NHLBI)
National Institute on Aging (NIA)
National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
National Institute on Deafness and Other Communication Disorders (NIDCD)
National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
National Institute on Drug Abuse (NIDA)
National Institute of Mental Health (NIMH)
National Institute of Neurological Disorders and Stroke (NINDS)
National Institute of Nursing Research (NINR)
National Center for Complementary and Integrative Health (NCCIH)
Office of Behavioral and Social Sciences Research (OBSSR)
Office of Research on Women's Health (ORWH)

The HealthMeasures team is made up of psychometricians and measurement science experts from Northwestern University, University of California Los Angeles, University of California San Diego, American Institutes for Research, University of Pittsburgh, University of North Carolina, and the National Institutes of Health. After 2018, HealthMeasures will operate as a non-profit, self-sustaining enterprise.

Measurement Expertise

The team's expertise lies in measure development, validation, translation, and expansion as well as the application of measures across a range of settings, diseases, and conditions. Members of the HealthMeasures team are available as consultants and collaborators through contracts and as co-investigators on grants conducting psychometric analyses, consulting on research design, constructing linking tables using PROsetta Stone methodologies, conducting cross-cultural validation and translation of measures, evaluating and interpreting the effects of interventions, applying HealthMeasures in clinical practice, and developing custom assessment technology in clinical, educational, and research settings. HealthMeasures experts routinely provide training and workshops. HealthMeasures assessment platforms are developed and supported by individuals with expertise in informatics, software development, quality assurance testing, and business analysis. This expertise was integral to the development and support of assessment technologies for research and clinical applications including the NIH Toolbox iPad app, the PROMIS iPad app, Assessment Center^{SM}, and the Assessment Center Application Programming Interface (API).

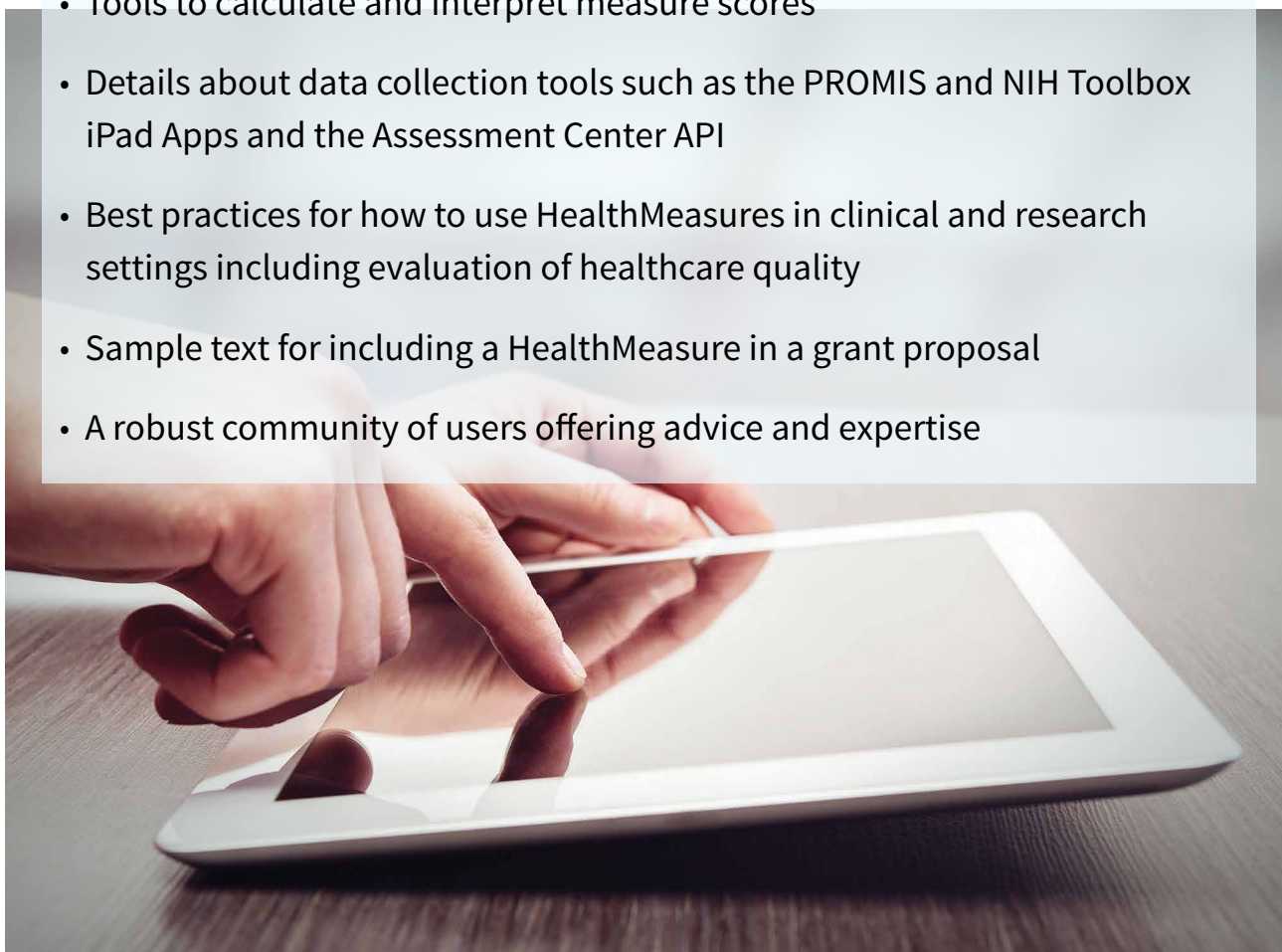To learn more about collaborating with a member of the HealthMeasures team, contact help@HealthMeasures.net.

HealthMeasures
TRANSFORMING HOW HEALTH IS MEASURED

PROMIS®, Neuro-QoL, ASCQ-Me^SM and NIH Toolbox® measures and more at HealthMeasures.net

- Free access to hundreds of administration-ready measures

- Guidance on selecting a measure for your purpose

- Tools to calculate and interpret measure scores

- Details about data collection tools such as the PROMIS and NIH Toolbox iPad Apps and the Assessment Center API

- Best practices for how to use HealthMeasures in clinical and research settings including evaluation of healthcare quality

- Sample text for including a HealthMeasure in a grant proposal

- A robust community of users offering advice and expertise