NIH Toolbox

Technical Manual

Jerry Slotkin, PhD, Michael Kallen, PhD, James Griffith, PhD, Susan Magasi, PhD, John Salsman, PhD, Cindy Nowinski, MD, PhD, and Richard Gershon, PhD

September 27, 2012

NIH Toolbox Technical Manual

Domain:

SENSATION

Subdomain:

VISION

Measure:

NIH Toolbox Vision-Related Quality of Life Survey

(Supplemental Measure)

Expert Contributors-VISION

Eileen Birch, PhD (Retina Foundation of the Southwest), Karen Cruickshanks, PhD (University of Wisconsin), Ron Hays, PhD (University of California-Los Angeles), Chris Johnson, PhD (University of Iowa), Paul Lee, MD, JD (University of Michigan), Maureen Maguire, PhD (University of Pennsylvania), Robert Massof, PhD (The Johns Hopkins University), Roberta McKean-Cowdin, PhD (University of Southern California), Cynthia Owsley, PhD (University of Alabama-Birmingham), Michael Repka, MD (The Johns Hopkins University), Jerry Slotkin, PhD (Northwestern University), Rohit Varma, MD (University of Southern California), Susan Vitale, PhD (National Eye Institute, NIH), Michael Wall, MD (University of Iowa)

Member Institutes, Centers, and Offices

National Center for Complementary and Alternative Medicine (NCCAM)

National Eye Institute (NEI)

National Institute of Biomedical Imaging and Bioengineering (NIBIB)

Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)

National Institute of Dental and Craniofacial Research (NIDCR)

National Institute of Environmental Health Sciences (NIEHS)

National Institute of General Medical Sciences (NIGMS)

National Institute of Mental Health (NIMH)

National Institute of Neurological Disorders and Stroke (NINDS)

National Institute of Nursing Research (NINR)

National Institute on Aging (NIA)

National Institute on Alcohol Abuse and Alcoholism (NIAAA)

National Institute on Deafness and Other Communication Disorders (NIDCD)

National Institute on Drug Abuse (NIDA)

Office of Behavioral and Social Sciences Research (OBSSR)

This Technical Manual contains the following informational sections:

Section 1:   **Introduction to NIH Toolbox**

Section 2:   **Validation**

Section 3:   **Norming**

Section 4:   **NIH Toolbox and the National Children's Study (NCS)**

Section 5:   **Domain Definition**

Section 6:   **Subdomain Definition**

Section 7:   **Measure Description**

Section 8:   **Post-Validation/Post-Norming Changes to the Measure**

Section 9:   **The Measure's Scoring Model**

Section 10:   **Measure Norms**

Section 1:   Introduction to NIH Toolbox

NIH Toolbox is a multidimensional set of brief measures assessing cognitive, emotional, motor, and sensory function from ages 3-85. This suite of on-line and royalty-free measures can be administered to study participants 3 to 85 years of age in two hours or less, across diverse study designs and settings.

What is the NIH Toolbox?

The NIH Toolbox provides a standard set of royalty-free, brief, and comprehensive assessment tools that can be used by researchers and clinicians in a variety of settings, with a particular emphasis on measuring outcomes in longitudinal epidemiologic studies and prevention or intervention trials across the lifespan (ages 3-85). The battery ensures that assessment methods and results can be used for comparisons across existing and future studies and provides a "common currency" for the study of neurological research that promotes economies of scale and enhanced efficiency in measurement. The NIH Toolbox can be used to monitor neurological and behavioral function over time and measure key constructs across developmental stages. This facilitates the study of functional changes across the lifespan, including evaluating intervention and treatment effectiveness.

The NIH Toolbox Batteries

The basic NIH Toolbox can be administered within two hours and divides tests into four domain batteries: Cognition, Emotion, Motor, and Sensation. In addition, within some domains, there are supplemental measures that are available to be administered.

Selection of the NIH Toolbox Domains and Subdomains

Four domains were selected for the NIH Toolbox: Cognition, Emotion, Motor, and Sensation. Subdomain selection was based upon literature reviews, expert interviews, and multiple formal Requests for Information (RFI) of NIH-funded researchers. Initial literature and database reviews and an RFI identified the subdomains for inclusion in the NIH Toolbox, existing measures relevant to the project goals, and criteria for instrument selection. NIH Project Team members, external content experts, and contract scientists met at a follow-up consensus meeting to discuss potential subdomains along with the criteria affecting instrument selection, creation, and norming. Additional expert interviews were undertaken to gather more detailed information from clinical and scientific experts to help further refine the list of possible subdomains. A second consensus group meeting was held and results directed the decision for the final NIH Toolbox to assess four core domain areas (cognitive, emotional, motor, and sensory health and function).

Selection of Measures for the NIH Toolbox

More than 1,400 existing measures were identified and evaluated for potential inclusion in the NIH Toolbox. The selection criteria included a measure's applicability across the life span,

psychometric soundness, brevity, ease of use, applicability in diverse settings and with different groups, and lack of intellectual property constraints. There was also a preference for instruments that were already validated and normed for use with individuals between 3 and 85 years old. Results of the instrument selection process yielded draft development plans established for the NIH Toolbox measures.

Early Childhood Use

NIH Toolbox measure development focused special attention on assessing young children, to ensure that all tests given are developmentally appropriate for ages 3-7. A special team of early childhood assessment consultants was engaged to provide testing guidelines for the very young, to offer input on measure development, and to review all NIH Toolbox measures to ensure they fit the needs of young children. Advanced statistical methods were used to emphasize continuity of measurement, allowing Toolbox users to confidently conduct longitudinal measurement from age 3 through the life span while assessing the same domain constructs.

Section 2:   Validation

Validation studies were conducted for all NIH Toolbox Sensation domain measures, to assure that these important tools for research met rigorous psychometric standards. Studies were

conducted across the entire age range and were statistically compared against "gold standard" measures wherever available.

For specifics regarding Sensation domain measure validation, see: Coldwell et al., Gustation Assessment Using the NIH Toolbox, *Neurology*, in press; Cook et al., Pain Assessment Using the NIH Toolbox, *Neurology*, in press; Dalton et al., Olfaction Assessment using the NIH Toolbox, *Neurology*, in press; Dunn et al., Somatosensation Assessment Using the NIH Toolbox, *Neurology*, in press; Paz et al., Development of a Vision-Targeted Health-Related Quality of Life Item Bank, *manuscript submitted for publication*; Rine et al., Vestibular Function Assessment Using the NIH Toolbox, *Neurology,* in press; Varma et al., Vision Assessment Using the NIH Toolbox, *Neurology*, in press; Zecker et al., Audition Assessment Using the NIH Toolbox, *Neurology*, in press. These manuscripts describe measure development studies undertaken (e.g., expert panels for content development and validation; cognitive interviews; small and large-scale pilot testing) and psychometric characteristics (e.g., internal consistency and test-retest reliability; convergent and divergent validity).

Section 3:   Norming

NIH Toolbox conducted a large national standardization study in both English and Spanish languages to allow for normative comparisons on each assessment. A sample of 4,859 participants, ages 3-85 – representative of the U.S. population based on gender, ethnicity, race,

and socioeconomic status – was administered all of the NIH Toolbox measures at sites around the country (n = 2,917 English-speaking children, ages 3-17; n = 496 Spanish-speaking children, ages 3-7; n = 1,038 English-speaking adults, ages 18-85; n = 408 Spanish-speaking adults, ages 18-85). NIH Toolbox normative scores are now available for each year of age from 3 through 17, as well as for age ranges 18-29, 30-39, 40-49, 50-59, 60-69, and 70-85, allowing for targeted and accurate comparisons to the U.S. population.

Specifics regarding NIH Toolbox norming sampling methods (e.g., stratification by age, gender, and language preference; sampling a minimum of 25-100 individuals per targeted demographic and language subgroup) and norming analytic methods (e.g., post-stratification adjustment using iterative proportional fitting, i.e., "raking") can be found in the following publication: Beaumont et al., Norming Plans for the NIH Toolbox, *Neurology*, in press.

Section 4: NIH Toolbox and the National Children's Study (NCS)

In collaboration with NIH Toolbox scientists, NCS investigators selected measures from PROMIS and NIH Toolbox for a Maternal Health Profile, the Maternal Self-Reported Health Battery. This profile assesses Physical Health (Physical function, Fatigue, Sleep disturbance, Sleep-related impairment), Mental Health (Anger, Anxiety, Depression, Positive affect, Perceived stress, Self-efficacy), and Social Health (Social support and companionship, Social isolation). The Maternal

Self-Reported Health Battery was field tested in fall 2011, using an online sample of 1000 women (200 pre-conception, 150 pregnant women (50 per trimester), and 650 mothers with a child between 0-36 months of age). In addition, NIH Toolbox norming was jointly sponsored by the NCS and included: 3,413 children in single-year age bands (from 3-17 years); 1,446 adults in seven age bands, including the mothers of children also being tested; and 105 pregnant women. The NIH Toolbox sampling plan matched distributions of race/ethnicity and level of education for each age band.

Section 5:   Domain Definition

Domain: SENSATION

Sensation refers to the biochemical and neurologic process of detecting incoming nerve impulses as nervous system activity. Sensory processes are vital to one's level of independence, in relationships with others, in academic and occupational endeavors, and for activities of daily living. Objective measures of Sensation can systematically examine and determine if participants have intact sensory functioning. There is also fundamental overlap of certain sensory processes with cognitive and motor functioning. Measurement of sensory health and function is important to epidemiologic and longitudinal studies whether or not Sensation is the primary focus of the study. Given the changes in sensory functioning across the lifespan, there is value in characterizing age-related sensory improvement and decline. The Sensation domain includes measures of:

AUDITION

Measured by:

        NIH Toolbox Words-in-Noise Test

        NIH Toolbox Hearing Threshold Test (Supplemental Measure)

        NIH Toolbox Hearing Handicap Inventory (Supplemental Measure)

OLFACTION

Measured by:

        NIH Toolbox Odor Identification Test

PAIN

Measured by:

        NIH Toolbox Pain Intensity Survey

        NIH Toolbox Pain Interference Survey

TASTE

Measured by:

        NIH Toolbox Taste Intensity Test

VESTIBULAR

Measured by:

        NIH Toolbox Dynamic Visual Acuity Test

        NIH Toolbox Standing Balance Test (contained within the NIH Toolbox Motor

        battery)

VISION

Measured by:

NIH Toolbox Visual Acuity Test

NIH Toolbox Vision-Related Quality of Life Survey (Supplemental Measure)

SENSATION Batteries

The NIH Toolbox Sensation Battery for ages 3-5 includes Visual Acuity, Dynamic Visual Acuity, and Odor Identification tests. For ages 6-11, the Words-in-Noise Test is added to the battery. For ages 12-17, the Taste Intensity Test is included with the other four; and for ages 18-85, the two Pain surveys are added. There are individual scores provided for each measure, as described below, but no composite scores.

Section 6:   Subdomain Definition

Subdomain: VISION

Vision is a complex sensation that provides us with a personal, conscious representation of our surrounding environment. Loss of vision or blindness may limit a person's ability to complete normal, daily activities and decrease overall quality of life. The key aspect of vision prioritized for testing in the NIH Toolbox is visual acuity. Visual acuity tests are used to measure impairments in visual resolution that can be caused by blurring of the retinal image, neural

processing disorders, or damage to neurons in the retina or other parts of the visual pathway.

In NIH Toolbox, Vision is measured by:

NIH Toolbox Visual Acuity Test

NIH Toolbox Vision-Related Quality of Life Survey (Supplemental Measure)

Section 7: Measure Description

VISION Supplemental Measure

The NIH Toolbox Vision-Related Quality of Life Survey assesses an individual's self-reported quality of life related to visual function in six different areas: color vision, distance vision, near vision, ocular symptoms, psycho-social, and role performance. The participant responds to 53 questions, most of which start with, "How much," "To what extent," or "How much of a problem," choosing from a list of answer options ranging from "Not difficult at all," to "Very difficult." This survey is recommended for ages 18-85 and takes approximately 10 minutes to complete.

Section 8: Post-Validation/Post-Norming Changes to the Measure

In addition to those changes previously reported on during the measure's development and validation phases (Paz et al., Development of a Vision-Targeted Health-Related Quality of Life Item Bank, *manuscript submitted for publication*; Varma et al., Vision Assessment Using the NIH Toolbox, *Neurology*, in press), the following changes have been made to this measure:

Test scoring changes: Six individual area scores will be presented, one per quality of life area measured (i.e., Color Vision, Distance Vision, Near Vision, Ocular Symptoms, Psychosocial, and Role Performance).

Section 9: The Measure's Scoring Model

Measurement theory applied for scoring:

Rasch/Item Response Theory (IRT)

Rasch/IRT model employed:

Graded Response Model

Measure length:

Fixed (53 items)

Response data:

Ordinal

Scores computed/available*:

Rasch/IRT Color Vision Theta Score

Rasch/IRT Distance Vision Theta Score

Rasch/IRT Near Vision Theta Score

Rasch/IRT Ocular Symptoms Theta Score

Rasch/IRT Psycho-Social Theta Score

Rasch/IRT Role Performance Theta Score

Color Vision T-Score (converted from Color Vision Theta Score based on the measure's

validation sample; mean=50, standard deviation=10)

Distance Vision T-Score (converted from Distance Vision Theta Score based on the

measure's validation sample; mean=50, standard deviation=10)

Near Vision T-Score (converted from Near Vision Theta Score based on the measure's

validation sample; mean=50, standard deviation=10)

Ocular Symptoms T-Score (converted from Ocular Symptoms Theta Score based on the

measure's validation sample; mean=50, standard deviation=10)

Psycho-Social T-Score (converted from Psycho-Social Theta Score based on the

measure's validation sample; mean=50, standard deviation=10)

Role Performance T-Score (converted from Role Performance Theta Score based on the

measure's validation sample; mean=50, standard deviation=10)

*Details on these scores and their interpretations are available in the NIH Toolbox Scoring and

Interpretation Guide.

Section 10:   Measure Norms

NIH Toolbox normative data are not available for this measure.