Title

HDQLIFE: development and assessment of health-related quality of life in Huntington disease (HD)

Journal

Quality of Life Research, 25(10)

ISSN

0962-9343

Authors

Carlozzi, NE
Schilling, SG
Lai, J-S
et al.

Publication Date

2016-10-01

DOI

10.1007/s11136-016-1386-3

Peer reviewed

Qual Life Res (2016) 25:2441–2455
DOI 10.1007/s11136-016-1386-3

HDQLIFE: development and assessment of health-related quality of life in Huntington disease (HD)

N. E. Carlozzi[1] · S. G. Schilling[1,2] · J.-S. Lai[3] · J. S. Paulsen[4,5,6] · E. A. Hahn[3] · J. S. Perlmutter[7,8] · C. A. Ross[9] · N. R. Downing[10] · A. L. Kratz[1] · M. K. McCormack[11] · M. A. Nance[12,13] · K. A. Quaid[14] · J. C. Stout[15] · R. C. Gershon[3,16] · R. E. Ready[17] · J. A. Miner[1] · S. K. Barton[7] · S. L. Perlman[18] · S. M. Rao[19] · S. Frank[20] · I. Shoulson[21] · H. Marin[22] · M. D. Geschwind[23] · P. Dayalu[24] · S. M. Goodnight[1] · D. Cella[3]

Accepted: 3 August 2016 / Published online: 13 August 2016

Abstract

Purpose Huntington disease (HD) is a chronic, debilitating genetic disease that affects physical, emotional, cognitive, and social health. Existing patient-reported outcomes (PROs) of health-related quality of life (HRQOL) used in HD are neither comprehensive, nor do they adequately account for clinically meaningful changes in function. While new PROs examining HRQOL (i.e., Neuro-QoL—Quality of Life in Neurological Disorders and PROMIS—Patient-Reported Outcomes Measurement Information System) offer solutions to many of these shortcomings, they do not include HD-specific content, nor have they been validated in HD. HDQLIFE addresses this by validating 12 PROMIS/Neuro-QoL domains in individuals with HD and by using established PROMIS methodology to develop new, HD-specific content.

Methods New item pools were developed using cognitive debriefing with individuals with HD, and expert, literacy,

✉ N. E. Carlozzi
carlozzi@med.umich.edu

[1] Department of Physical Medicine and Rehabilitation, University of Michigan, North Campus Research Complex, 2800 Plymouth Road, Building NCRC B14, Room G216, Ann Arbor, MI 48109-2800, USA

[2] Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

[3] Department of Medical Social Sciences, Northwestern University, Chicago, IL, USA

[4] Department of Psychiatry, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA

[5] Department of Neurology, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA

[6] Department of Psychology, The University of Iowa, Iowa City, IA, USA

[7] Departments of Neurology, Radiology, and Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO, USA

[8] Program in Occupational Therapy and Program in Physical Therapy, Washington University School of Medicine, St. Louis, MO, USA

[9] Departments of Psychiatry, Neurology, Pharmacology and Neuroscience, Johns Hopkins University, Baltimore, MD, USA

[10] College of Nursing, The University of Iowa, Iowa City, IA, USA

[11] Department of Pathology, Rowan University, Piscataway, NJ, USA

[12] Struthers Parkinson's Center, Golden Valley, MN, USA

[13] Hennepin County Medical Center, Minneapolis, MN, USA

[14] Department of Medical and Molecular Genetics, Indiana University, Indianapolis, IN, USA

[15] School of Psychological Sciences, Monash University, Clayton, VIC, Australia

[16] Department of Preventative Medicine, Northwestern University, Evanston, IL, USA

[17] Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, MA, USA

[18] Department of Neurology, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA

[19] Department of Psychiatry and Psychology, Cleveland Clinic, Cleveland, OH, USA

[20] Beth Israel Deaconess Medical Center, Boston, MA, USA

[21] Department of Neurology, Georgetown University, Washington, DC, USA

[22] Department of Psychiatry, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA

and translatability reviews. Existing item banks and new item pools were field tested in 536 individuals with prodromal, early-, or late-stage HD.
Results Moderate to strong relationships between Neuro-QoL/PROMIS measures and generic self-report measures of HRQOL, and moderate relationships between Neuro-QoL/PROMIS and clinician-rated measures of similar constructs supported the validity of Neuro-QoL/PROMIS in individuals with HD. Exploratory and confirmatory factor analysis, item response theory, and differential item functioning analyses were utilized to develop new item banks for Chorea, Speech Difficulties, Swallowing Difficulties, and Concern with Death and Dying, with corresponding six-item short forms. A four-item short form was developed for Meaning and Purpose.
Conclusions HDQLIFE encompasses both validated Neuro-QoL/PROMIS measures, as well as five new scales in order to provide a comprehensive assessment of HRQOL in HD.

Keywords Neuro-QoL · PROMIS · Health-related quality of life · HDQLIFE · Huntington disease · Patient-reported outcome (PRO)

Introduction

Huntington disease (HD) is a hereditary neurodegenerative disorder caused by a CAG triplet repeat expansion in the gene *huntingtin*; HD affects approximately 1 in 10,000 individuals in populations of European descent [1–4]. Since HD is a dominantly inherited disease, a person whose parent carries the HD mutation gene has a 50 % chance inheriting it at the time of conception. About 150,000 individuals in the USA are "at risk" for HD. The age of onset of HD is inversely related to the length of the CAG repeat; for the most common expansion lengths, mutation signs and clinical diagnosis of HD (based on characteristic motor symptoms) typically occur between ages 30 and 50. Motor, cognitive, and psychiatric abnormalities may emerge gradually, more than a decade before diagnosis (prodromal HD), and worsen progressively [5]. Although the rate of clinical progression differs for each person, HD is generally fatal within 15–20 years of clinical diagnosis [6]. The fact that this progressive, fatal disease typically strikes individuals during the prime of their lives underscores the need for interventions that slow the disease progression and maximize health-related quality of life (HRQOL).

[23] Memory and Aging Center, UCSF, San Francisco, CA, USA

[24] Department of Neurology, University of Michigan, Ann Arbor, MI, USA

HRQOL is a multidimensional construct defined as the impact that a disease or disability has on different aspects of well-being [7]. This follows the World Health Organization (WHO) framework for HRQOL which includes physical, social, and emotional well-being [8]. HRQOL differs from general quality of life (QOL), which is a poorly defined concept that lacks a consensus definition, that may or may not be synonymous with HRQOL [9, 10]. Current measures of HRQOL are insufficient to capture the broad extent of functional and symptom distress in HD and are also insensitive to potential intervention effects in HD. Most HRQOL measures used in HD were developed for other clinical populations and are inadequate because of important differences in symptoms across the neurodegenerative diseases. For example, although Parkinson's disease (PD) and HD are both basal ganglia disorders characterized by motor abnormalities, these motor manifestations present differently; PD is typically characterized by tremor and bradykinesia, whereas HD is typically choreic (involuntary "dance-like" movements) and hyperkinetic [11]. Therefore, a measure of motor functioning developed for PD may not be meaningful for HD. Similarly, although cognitive dysfunction in HD overlaps with cognitive dysfunction in Alzheimer disease (AD), individuals with HD typically have "subcortical" deficits (in attention, processing speed, and executive dysfunction), whereas individuals with AD also have prominent "cortical" deficits (in memory, language, and executive dysfunction) [12]. Furthermore, generic measures of HRQOL cannot detect subtle differences in function for prodromal HD symptoms [13], and single-item ratings [14, 15] have inadequate sensitivity and reliability to detect change over time [15]. In addition, the only existing HD-specific measure of HRQOL, the HDQoL [16], has evidence to support reliability and construct validity, but did not meet minimally established sample size criteria for its developmental approach and takes ~22 min to complete [17]. Thus, there is a critical need for a well-developed, validated, brief HD-specific patient-reported outcome (PRO) measure of HRQOL. This is especially important given that the focus of clinical interventions for HD is not just to prolong life, but also to prolong quality living.

Recently, there has been an investment in the development of new, state-of the-art systems to better assess HRQOL for a variety of chronic health conditions. Specifically, the Quality of Life for Neurological Disorders (Neuro-QoL) measurement system [18, 19] and Patient-Reported Outcomes Measurement Information System (PROMIS) [20] were designed to create and disseminate reliable and valid standardized PROs that measure key symptoms and health concepts. PROMIS was developed for use in individuals with chronic conditions, and Neuro-QoL extended this work to neurological disorders (stroke,

PD, multiple sclerosis, child and adult epilepsy, amyotrophic lateral sclerosis, and muscular dystrophy). Neuro-QoL and PROMIS offer several advantages over more traditional measures. These systems allow for cross-disease comparison. In addition, PROMIS and Neuro-QoL contain many identical items that allow linkage between measures such that a score on one PROMIS measure can be used to estimate a score on a Neuro-QoL measure. Third, these PRO systems utilize computerized adaptive test (CAT) technology, a method whereby each individually administered item is selected based on the previous item response. CATs allow for the sensitive measurement of a broad range of symptomatology with the administration of a small subset of items (between 5 and 12 items) without losing the precision of a longer measure. This reduces response burden, which is particularly important in HD where motor, psychiatric, and cognitive symptoms may impair the ability to respond to long questionnaires. The exact subset of items administered in a CAT depends on upon item response theory (IRT) calibrations [21]. A calibrated item bank is a set of carefully crafted questions that develop, define, and quantify a common theme [22, 23]. The items can be arranged along a scale, e.g., from no symptoms to extreme symptoms. The dynamic nature of CAT allows for greater sensitivity across the disease spectrum than most traditional, static measures while still retaining the integrity of the full measure. This is especially relevant in HD, where many measures exhibit floor effects during the prodromal phase of the disease and ceiling effects for the later stages of the disease [24]. CAT also provides better precision and lower standard error than static measures, even when the number of items administered for each is identical [25]. This is true even when short forms target a specific end of a symptom trait (such as low-end or high-end fatigue) [25].

The purpose of the current study was to develop and validate a PRO measurement system that captures both the generic and more unique aspects of HRQOL in HD ("HDQLIFE"). Given the complexity of the multi-phase study to develop the HDQLIFE, this paper provides a broad introduction to the processes and aims of each phase of the study; further details on the methods and results of each phase are found in the companion articles [26–28]. Broadly, this study focused on validating existing measurement systems to capture generic, relevant aspects of HRQOL for individuals with HD (i.e., Neuro-QoL and PROMIS, described below), and developing additional content that would allow for disease-specific sensitivity utilizing a computer adaptive test framework (using PROMIS measurement development standards [29]). These results are complemented by the companion articles which include detailed presentations of exploratory and confirmatory factor analysis results, graded response model results, differential item functioning analysis results, as well as item-level calibration data and preliminary validation data generated using post hoc computer adaptive test simulations [26–28].

Methods

Literature reviews and a qualitative focus group study [30] were conducted to characterize HD-relevant HRQOL domains (including the identification of relevant Neuro-QoL/PROMIS measures) and develop items for domains that were not captured within these existing systems. Next, a quantitative study served to: validate the existing, relevant Neuro-QoL/PROMIS measures in individuals with HD, and create and validate new, HD-specific item banks (i.e., computer adaptive tests).

1. Item Development.

A qualitative focus group study and literature review were conducted to determine the domains, subdomains, and items that should be used to assess HRQOL in HD [30]. Focus groups were conducted with key HD stakeholders and included six groups with individuals at various stages of diagnosed, symptomatic HD, five groups with individuals either at risk for HD (i.e., have not been tested and were not diagnosed with HD yet but have a parent with HD) or with prodromal HD (i.e., have a positive gene test, but not diagnosed with manifest HD), three groups with non-clinical HD caregivers (e.g., family members), and two groups with HD clinicians (e.g., physicians, nurses). Participants discussed what the term "quality of life" meant to them, what they believed to be the most important aspects of HRQOL, and how HD affected their HRQOL. Focus group discussions were transcribed verbatim and analyzed according to a well-established frequency analysis approach [31]. Detailed qualitative findings have been described elsewhere [30]. Briefly, results showed that several PROMIS/Neuro-QoL measures were relevant in HD and that a number of HD-specific HRQOL issues were not captured by these PROs (see Fig. 1).

The next step of the development of the HDQLIFE measurement system was to create preliminary item pools examining chorea, speech and swallowing difficulties, and end of life issues. Each item pool went through several different iterations based on expert review, cognitive debriefing interviews with individuals with HD, literacy review, and translatability review (to enable future translation into different languages). Expert review included insight from measurement development experts and professionals with clinical expertise in HD; experts provided feedback with regard to item overlap, appropriateness of the content to each subdomain, wording suggestions/changes, and content coverage (i.e.,

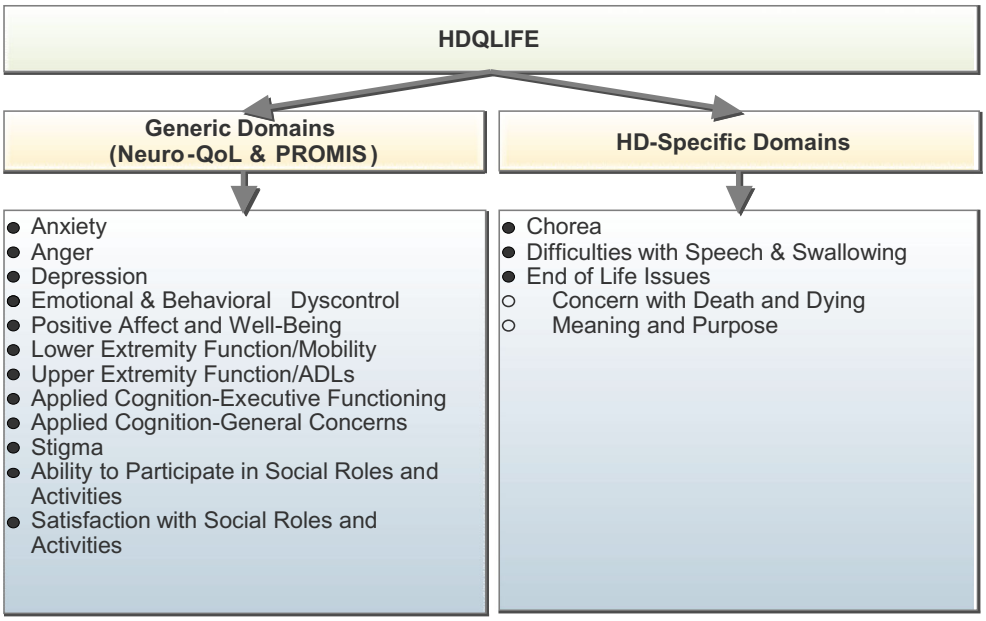

Fig. 1 Components of the HDQLIFE measurement system

that all aspects of the specified domain were represented). Additional items were developed in cases where content coverage was deemed inadequate. All new items were also reviewed by at least 5 individuals with prodromal or symptomatic HD (i.e., cognitive debriefing) to ascertain comprehension, processes used to arrive at a particular response (retrieving relevant information from memory, response selection including motivation and social desirability), and overall relevance of an item to the content being measured [32]. All new items also underwent a literacy review using the Lexile framework [33] to ensure that the items were written no higher than a fifth-grade reading level. Thus, we maximized the accessibility of this measure to participants, regardless of their level of education or cognitive impairment. Finally, a translatability review was conducted to maximize the potential for this measure to be translated into other languages in the future. We focused on Spanish translation for this review. Forward and backwards Spanish translations were conducted by a Spanish-speaking translation scientist to identify potential concerns, such as items that contained wording or concepts that would be difficult to translate.

HDQLIFE Chorea Item Pool Literature review and focus group data were used to create an initial item pool of 141 chorea items; 75 items were deleted and 3 items were revised based on expert review, 0 items were deleted and 9 items were revised based on translation review, and 2 items were deleted and 5 items were revised based on cognitive interview feedback. The final chorea item pool was comprised of 64 items.

HDQLIFE Speech and Swallowing Item Pool Literature review and focus group data were used to create an initial item pool of 102 speech and swallowing items; 49 items were deleted and 12 items were revised based on expert review, 1 item was deleted and 3 items were revised based on translation review, and 5 items were deleted and 25 items were revised based on cognitive interview feedback. The final speech and swallowing item pool was comprised of 47 items.

HDQLIFE End of Life Concerns Item Pool Literature review and focus group data were used to create an initial item pool of 69 items related to end of life concerns; 21 items were deleted and 0 items were revised based on expert review, 0 items were deleted and 39 items were revised based on translation review, and 3 items were deleted and 13 items were revised based on cognitive interview feedback. The final end of life concerns item pool was comprised of 45 items.

2. Quantitative Study.

Once the item pools were developed, all items were field tested in 536 individuals including those with prodromal HD and manifest HD to meet the standards established by PROMIS to develop new CATs [29].

Participants

Participants were 18 years old or older, able to read and understand English, had either a positive test for the HD gene mutation (CAG $\geq$ 36, but did not yet have an HD

clinical diagnosis, $n = 205$) and/or a clinical diagnosis of HD ($n = 331$), and had the ability to provide informed consent. The Total Functional Capacity (TFC) [34], as determined by clinician-rated administration, was used to classify participants with an HD diagnosis as either early-stage (sum scores of 7–13) or later-stage HD (sum scores of 0–6; described in more detail, below). Participants were recruited at several locations in the USA to ensure a geographically diverse sample. This included eight established HD clinics (Los Angeles, CA; Iowa City, IA; Indianapolis, IN; Baltimore, MD; Ann Arbor, MI; Golden Valley, MN; St. Louis, MO; Piscataway, NJ), the National Research Roster for Huntington's disease, online medical record data capture systems [35], and articles/advertisements in HD-specific newsletters and Websites. Participants were also recruited in conjunction with other ongoing research studies, such as Predict-HD (San Francisco, CA; Iowa City, IA; Indianapolis, IN; Baltimore, MD; St. Louis, MO; Cleveland, OH) [36–38], as well as in cooperation with HD support groups and HD specialized nursing home units (Phoenix, AZ; Tuscon, AZ; Denver, CO; Jacksonville, FL; Des Moines, IA; Louisville, KY; Lansing, MI; Robbinsdale, MN; Lakewood, NJ; Plainfield, NJ; New York City, NY; Dallas, TX; Seattle, WA). Participants received monetary compensation ($40) for participating in this study.

Measures

Participants were evaluated using the *Unified Huntington's Disease Rating Scales* (*UHDRS*) [39], a standardized clinical rating scale that assesses four components of HD: motor function, cognition, behavior, and functional abilities. Although the UHDRS has several documented shortcomings [24, 40–44], it is the most frequently used assessment measure in HD clinical trials [45] and is included in the common data element recommendations provided by the National Institute of Neurological Disorders and Stroke [46]. The reliability and internal consistency of the four components of the UHDRS have been well studied [39]. We examined Total Functional Capacity (TFC), Total Motor Score (TMS), Independence Scale, and two measures of Cognition (total score for Symbol Digit Modalities Test [SMDT] [47] and Stroop Interference [48, 49]). The TFC is a 5-item measure that provides an index of day-to-day functioning across the domains of occupation, finances, domestic chores, activities of daily living, and care level. Total score ranges from 0 to 13 with higher scores indicating better functioning. The TMS provides a composite measure of oculomotor function, dysarthria, chorea, dystonia, gait, and postural stability; higher scores indicate more motor dysfunction. The Independence Scale is rated from 0 to 100, with higher scores indicating better functioning/greater levels of independence. Executive function measures included the SDMT (processing speed) and Stroop tests (interference); higher scores indicate better performance. Participants also were administered the *Problem Behaviors Assessment Scale (PBA-s)* [50] which is a clinician-administered assessment of behavior. For the purposes of this study, we examined clinician-rated Apathy, Irritability, Aggression, Anxiety, and Depression.

Participants completed the three newly developed *HDQLIFE item pools* ($n = 64$ chorea items, $n = 47$ speech/swallowing difficulties items, and $n = 45$ end of life concerns items). Participants also completed CATs for 12 PRO item banks from the *Neuro-QoL and PROMIS* identified as relevant to HD (Anxiety, Anger, Stigma, Emotional and Behavioral Dyscontrol, Positive Affect and Well-Being, Depression, Ability to Participate in Social Roles and Activities, Satisfaction with Social Roles and Activities, Lower Extremity Function/Mobility, Upper Extremity Function/ADLs, Applied Cognition Executive Functioning, and Applied Cognition General Concerns). Finally, participants completed two generic measures of HRQOL, the 12-Item *World Health Organization Disability Assessment Schedule 2.0* (*WHODAS 2.0*) [51] and the *Euro-Qol-5D* (*EQ5D*) [52]. The WHODAS 2.0 is a 12-item standardized self-report measure of functioning and disability; higher scores indicate worse HRQOL. The EQ5D is a 5-item standardized measure of health status; higher scores indicate worse overall HRQOL.

Missing Data

Missing data rates were generally very low. The majority of our sample (99 %) had complete data for clinician-rated motor, functioning, and behavioral assessments (i.e., PBA-s and UHDRS Motor, Independence, and TFC measures); 99 % completed the EQ5D; 93–96 % completed the clinician-administered cognition measures (i.e., UHDRS SDMT and Stroop Interference); 93–95 % completed the HDQLIFE measures; 91–92 % completed the PROMIS/Neuro-QoL measures; and 89 % completed the WHODAS. Not surprisingly, rates of data loss were higher for the late-stage HD participants relative to both of the other HD groups for most of the measures (for prodromal vs. late-stage HD all X^2 $p < .05$ except for HDQLIFE Speech Difficulties; for early vs. late-stage HD all X^2 $p < .05$ except for the HDQLIFE measures and Neuro-QoL Positive Affect and Well-Being). IRT models (the primary method used in this paper) are designed to handle missing

data; the less missing data, the more stable parameter estimation. In general, less than 50 % of missing data for IRT models is considered acceptable [53, 54].

Data Capture

Study participants generally completed all measures (clinician-administered and PROs) within a 2-week time frame. PROs were completed through Assessment Center (https://www.assessmentcenter.net), either at a designated computer during the research visit (for individuals with restricted access to a computer or the internet), or on a personal or publically available computer with an Internet connection. Participants could opt to complete PROs independently, or with the assistance of local site staff or a family member; participants and caregivers were instructed that response selections should always be those of the participant. They were instructed that assistance should be limited to logging in to the online study, reading questions, and/or clicking response options, when appropriate; participants were provided the following written and verbal instructions, "IMPORTANT: It is okay if you ask a caregiver/friend/family member to help you complete this survey (use the mouse and keyboard or touchscreen), but we want to make sure that the answers reflect what you feel and believe. It is not okay for the caregiver/friend/family member to answer questions for you; each response should be based on what you believe and feel." Upon survey completion, participants were also asked to indicate whether they received help completing the survey: 65 % indicated completing the assessments independently; 15 % indicated receiving assistance from a caregiver/family member/friend; 10 % indicated receiving assistance from study staff; these data were missing for 9 % of participants. Participants indicating that they received assistance were also asked to indicate the type of assistance they received (participants could indicate more than one response): 89 participants indicated needing help using the computer/ipad (i.e., using mouse and/or keyboard or touchscreen); 67 participants indicated that his/her caregiver (family member, or friend) helped explain questions; 34 participants indicated that their caregiver (family member, or friend) answered questions by reminding them of important information; and 17 participants indicated that a caregiver (family member, or friend) helped by answering questions.

Psychometric Analysis Steps

Development of the new HRQOL item banks and CATs involved identifying unidimensional sets of items and conducting item response theory (IRT) [55] analyses to develop the calibration data needed to program the CAT. Each item pool (i.e., chorea, speech and swallowing difficulties, and end of life issues) was analyzed separately using factor analyses implemented in MPLUS (version 6.12) [56]; the sample was randomly divided into two separate datasets for these analyses. In the first dataset, exploratory factor analysis (EFA) was used to establish the number of unidimensional factors within each item pool as determined by: eigenvalues >1; scree plot review (i.e., number of factors before the break in the scree plot); and number of factors that explained >5 % of the variance. A promax rotation then was used to examine the association among factors by calculating their loadings (criterion > 0.4) and inter-factor correlations. Each unidimensional set of items (determined by EFA) was then subjected to confirmatory factor analysis (CFA) to assess model fit using the second randomly generated dataset. An iterative process including clinical input was taken into account to finalize item exclusion/inclusion [57–59].

Once unidimensional item sets were identified, an IRT graded response model (GRM) [60] was implemented in IRTPRO (version 2.1) [61]. To be retained, items had to demonstrate good psychometric properties. Items were also examined for differential item functioning (DIF) based on age, gender, and education; the LORDIF package within R (version 0.3–2) [62] was used to conduct these analyses [63]. DIF is an indication of unexpected behavior by an item on a test, such that an item performs differently for a subgroup of participants when it should not (e.g., men perform better than women). Items exhibiting DIF for age, gender, and/or education were excluded from the final item set.

Administration time for these new measures was recorded, and a univariate analysis was conducted to determine whether there were significant differences for the HD groups (prodromal vs. early-, vs. late-HD). An exploratory analysis examining Pearson correlations between CAG repeat number and the new HDQLIFE measures was also conducted.

Validation of PROMIS/Neuro-QoL

Pearson correlations between the PROMIS/Neuro-QoL measures and comparator measures were calculated to examine construct validity. Comparator measures included two generic self-report measures of HRQOL (WHODAS 2.0 [64] and EQ-5D [52]), as well as selected measures from two clinical rated measures: the UHDRS (TMS, Independence Scale, Symbol Digit Modalities Test [47], and Stroop Interference) [39] and the PBA-s (Apathy, Irritability, Aggression, Anxiety, and Depression) [50]. To demonstrate adequate construct validity, correlations between the new measures and generic measures should be moderate to large ($r = 0.5$–0.8) and correlations with clinician measures should be small to moderate ($r = 0.2$–0.4) [65].

Sample Size Considerations

Sample size consideration was determined based on the need for IRT analysis, the primary method in the current effort. While sample sizes of 200–1000 have been proposed when using graded response model (GRM), in which a larger sample size can produce more stable parameter estimation [60, 66], rules of thumb dictate that a minimum of 5–10 individuals are needed for every item within an item pool [67–69]. With an average of 50 items per item pool, 500 individuals were needed for reliable item response theory (IRT) calibration data. Additionally, differential item functioning (DIF) analyses (an indication of item bias) can be performed provided that there are at least 200 participants within each condition; sampling stratification considered age (≤40 vs. >40 and ≤50 vs. >50), gender (male vs. female), and education (≤high school vs. >high school]) [70].

Results

Participant demographics

Five hundred thirty-six individuals with HD (prodromal and manifest) participated in this study (Table 1); 205 individuals had prodromal HD, 202 had early-stage HD, and 125 had late-stage HD (4 participants did not have enough information to designate a classification). There were no significant group differences for sex, $X^2(2, 532) = 4.29$, $p = .12$, but there were small differences across groups for education; $F(2, 506) = 16.18$, $p < 0.0001$, with early-HD and late-HD groups having 1 to 1.5 fewer years of education than the prodromal HD group. As expected, since HD symptoms progress with age, analysis of age of the groups showed significant differences among the three groups, $F(2, 529) = 45.01$, $p < .0001$. The prodromal group ($M = 45.65$, SD = 11.99) was

Table 1 Demographic data for the HDQLIFE participants

Variable	Prodromal ($n = 205$)	Early ($n = 202$)	Late ($n = 125$)	All ($n = 536$)
Age (years)*				
M (SD)	42.56 (12.08)	51.42 (12.80)	54.69 (11.99)	48.74 (13.31)
Sex				
Female	64.4	54.5	57.6	59.0
Male	35.6	45.5	42.4	41.0
Ethnicity				
Not Hispanic or Latino	92.7	92.6	96.8	93.5
Hispanic or Latino	1.5	4.5	0.8	2.4
Not provided	5.9	3.0	2.4	4.1
Race (%)				
Caucasian	97.6	96.5	92.8	95.9
African American	0.0	2.0	6.4	2.2
Other	2.0	1.5	0.8	1.6
Unknown	0.5	0.0	0.0	0.4
Education (years)*				
M (SD)	15.91 (2.94)	14.71 (2.78)	14.20 (2.60)	15.05 (2.88)
Marital status				
Single, never married	15.6	14.4	11.2	14.0
Married	64.4	51.0	58.4	57.8
Separated/divorced	13.2	23.8	22.4	19.2
Widowed	0.0	3.0	3.2	1.9
Living with partner	2.9	4.0	0.0	2.6
Unknown	3.9	4.0	4.8	4.5
Years since diagnosis		($n = 159$)	($n = 77$)	($n = 238$)
M (SD)	–	3.07 (3.71)	5.88 (4.62)	3.97 (4.22)
CAG repeats	($n = 195$)	($n = 154$)	($n = 61$)	($n = 412$)
M (SD)	42.19 (2.90)	43.19 (2.90)	43.19 (3.92)	42.90 (4.09)

Entries in the table represent percentage of participants unless otherwise specified

* There were significant group differences for this variable

significantly younger than both manifest groups, and the early-HD group ($M = 51.42$, SD = 12.80) was significantly younger than the late-HD group ($M = 42.56$, SD = 12.08).

New HDQLIFE CAT Development

Across the 3 item pools, 156 items were field tested. For the chorea item pool, EFA and CFA supported 34 unidimensional items; the final Chorea item bank is comprised of 34 items (detailed analyses can be found in [27]). For the speech and swallowing difficulties item pool, EFA and CFA supported two separate unidimensional sets of items: difficulties with speech (27 unidimensional items) and difficulties with swallowing (15 unidimensional items). The final Speech Difficulties item bank is comprised of 27 items (no items were deleted based on IRT; detailed analyses can be found in [28]), and the final Swallowing Difficulties item bank is comprised of 16 items (1 item was deleted based on IRT; detailed analyses can be found in [28]). Finally, for the end of life item pool, EFA and CFA supported two unidimensional item sets: Concern with Death and Dying (12 unidimensional items) and Meaning and Purpose (7 items). The final Concern with Death and Dying item bank is comprised of 12 items (no items deleted based on IRT; detailed analyses can be found in [26]). There were not enough items retained to develop a CAT for Meaning and Purpose; thus, 4 items comprised the final Meaning and Purpose short form (3 items were deleted based on IRT; detailed analyses can be found in [26]). Four new CATs were developed: Chorea, Speech Difficulties, Swallowing Difficulties, and Concern with Death and Dying. Six-item short forms were selected by expert review for each of these measures; a 4-item short form was developed to assess Meaning and Purpose. The analysis results for the new HDQLIFE measures are shown in Table 2. Average administration time for each new HDQLIFE measure was less than 1 min; univariate analyses indicated significant differences among all three groups for all measures (in all cases, prodromal participants had the fastest completion times, early-HD completion times were in the middle, and late-HD participants had the slowest completion times; Table 3).

All newly developed HDQLIFE measures are scored on a *t* metric with a mean of 50 and standard deviation of 10, which is the same metric utilized for Neuro-QoL/PROMIS [71]. Thus, scores below 40 (1.0 SD below the mean) can be considered low and scores above 60 can be considered high. Note that the referent group (i.e., the group used to develop the algorithm for the CATs) for the new measures (i.e., Chorea, Speech Difficulties, Swallowing Difficulties, Concern with Death and Dying, and Meaning and Purpose) are individuals with HD, while the referent group for the

Table 2 New HDQLIFE measures statistics

Domain	Final items	Items in short form	Item(s) removed	CFI (>.90)	TLI	RMSEA (<.15)	Chronbach's alpha (>0.80)	Item-total correlation (>0.40)	Slope range	Threshold range	Administration time (s) *M* (SD)
Chorea	34	6	0	0.98	0.98	0.08	0.98	All > 0.4	2.64–6.21	−0.39 to 2.13	42.43 (31.31)
Speech	27	6	0	0.98	0.98	0.09	0.98	All > 0.4	2.91–6.32	−0.41 to 2.42	41.96 (30.85)
Swallowing	15	6	1	0.99	0.98	0.12	0.97	All > 0.4	3.06–7.11	0.11 to 3.00	43.53 (33.36)
Death and Dying	12	6	3	0.98	0.96	0.11	0.94	All > 0.4	1.48–4.57	−0.98 to 3.65	45.05 (27.32)
Meaning and Purpose	4	4	9	0.99	0.98	0.11	0.84	All > 0.4	2.26–4.75	−3.26 to 0.11	35.12 (21.07)

CFI Confirmatory Fit Index, *TLI* Tucker–Lewis Index, *RMSEA* root mean squared error

Table 3 Administration times (seconds) for prodromal, early-, and late-stage HD participants

	Prodromal HD *M* (SD)	Early-HD *M* (SD)	Late-HD *M* (SD)	*F*	*P*	Partial η^2
HDQLIFE Chorea	20.80 (15.10)	50.37 (26.94)	67.94 (34.44)	138.67	<0.0001	0.36
HDQLIFE Speech Difficulties	22.33 (16.49)	46.67 (27.00)	68.70 (33.53)	126.18	<0.0001	0.34
HDQLIFE Swallowing Difficulties	21.86 (16.13)	49.98 (29.45)	71.97 (37.16)	128.19	<0.0001	0.34
HDQLIFE Concern with Death and Dying	30.64 (14.30)	48.53 (26.37)	63.84 (32.29)	70.62	<0.0001	0.22
HDQLIFE Meaning and Purpose	24.45 (14.44)	37.70 (18.56)	50.75 (24.67)	70.76	<0.0001	0.23

Significant differences were found between all groups

Neuro-QoL/PROMIS measures is the general population. There were significant group differences among the three HD groups on all of the HDQLIFE measures except Concern with Death and Dying and Meaning and Purpose (Table 4); differences were in the expected direction. There were also statistically significant though very modest associations between CAG repeat number and all of the new HDQLIFE measures except Meaning and Purpose ($r = .21$, $p < .01$ for HDQLIFE Chorea; $r = .20$, $p < .01$ for HDQLIFE Speech Difficulties; $r = .23$, $p < .01$ for HDQLIFE Swallowing Difficulties; $r = .11$, $p < .05$ for HDQLIFE Concern with Death and Dying; and $r = -.07$, $p > .05$ for HDQLIFE Meaning and Purpose) providing preliminary support for construct validity of these new measures.

Preliminary support for Neuro-QoL/PROMIS Validity

Descriptive information regarding the Neuro-QoL/PROMIS measures, other generic self-reported measures of HRQOL, and clinician-rated assessments is provided in Table 4; for most measures, prodromal HD performed better than early-HD and late-HD, and early-HD performed better than late-HD. Neuro-QoL/PROMIS had moderate to strong relationships with generic self-report measures of HRQOL (*r*'s ranged from .34 to .74; Table 5). Neuro-QoL/PROMIS measures generally had moderate relationships with clinician-rated measures (*r*'s ranged from .35 to .70 with the majority between .42 and .49). Correlations tended to be highest between Neuro-QoL/PROMIS physical, social, and cognitive measures and corresponding self-report measures of these same constructs. Correlations were lowest among PROMIS emotion measures and corresponding measures, and highest among Neuro-QoL physical functioning measures and corresponding measures.

Discussion

Clinical trials aimed at slowing the progression of HD are underway. Unfortunately, these clinical studies employ few or no PROs [45]. When PROs are included, the specific measures utilized do not allow for cross-study or cross-disease comparison and may not be sufficiently sensitive to detect small but clinically meaningful changes in function [72]. Furthermore, existing PROs are often lengthy and time intensive. This is especially problematic given the regulatory and public interests to include PROs as meaningful endpoints in clinical trials [73]. The HDQLIFE measurement system uses state-of-the-art measurement techniques to help remedy these problems. HDQLIFE includes 12 validated Neuro-QoL/PROMIS measures in HD, as well as five new HD-specific measures: Chorea, Speech Difficulties, Swallowing Difficulties, Concern with Death and Dying, and Meaning and Purpose. HDQLIFE is unique in that it includes new HD-specific items based upon direct input from participants living with HD or the threat of HD, and consultation with experts who work with individuals with HD. HDQLIFE also includes "generic" items from PROMIS and Neuro-QoL to enable comparisons across different medical populations. Thus, HDQLIFE allows for *both* HD specificity and cross-disease comparisons, providing a significant advantage to more traditional measurement systems that require a trade-off between these two functions.

The HDQLIFE is also the first PRO assessment in HD to utilize item banking and CAT methodology. In CAT, each individual item is selected based on the response to the previous item. This "smart test" allows clinicians and researchers to ascertain a person's level of functioning using only a minimal number of items without losing the precision of a longer measure. CAT offers several advantages to traditional test administration, including specification of the minimum/maximum number of items, and/or maximum acceptable standard error. Further, most measures can be administered as fixed-length short forms (4–8 items), effectively reducing test length without sacrificing test sensitivity (i.e., administration time for each new HDQLIFE measure was less than 1 min). This is particularly important given time constraints inherent in clinical trials assessment and the need to limit participant burden during test administration, which is especially important during later-stage HD when cognition is compromised and processing speed is slowed [12]. CAT also has the

Table 4 Average scores for clinician-rated and self-report assessments for prodromal, early-, and late-stage HD participants

	Prodromal-HD	Early-HD	Late-HD	All
Self-report measures				
HDQLIFE—*M* (SD)				
Speech Difficulties SF[a,b,c]	43.11 (6.66)	50.34 (7.50)	54.70 (7.70)	48.52 (8.50)
Swallowing Difficulties SF[a,b,c]	50.16 (2.72)	53.83 (5.27)	57.31 (6.02)	53.18 (5.40)
Chorea SF[a,b,c]	43.44 (3.80)	51.64 (7.79)	57.05 (8.20)	49.63 (8.52)
Concern with Death and Dying SF	49.22 (8.31)	50.61 (8.63)	50.55 (10.62)	50.03 (9.00)
Meaning and Purpose SF	50.52 (9.06)	50.14 (9.62)	48.95 (8.26)	50.03 (9.12)
PROMIS—*M* (SD)				
Anger	46.16 (9.30)	47.95 (10.78)	47.26 (11.34)	47.17 (10.42)
Anxiety	52.12 (8.20)	52.93 (9.91)	54.02 (10.11)	52.89 (9.40)
Depression[b]	47.70 (8.88)	50.2 (10.04)	51.9 (10.03)	49.73 (9.75)
Neuro-QoL—*M* (SD)				
Positive Affect and Well-Being	55.05 (7.05)	54.36 (8.50)	54.44 (7.75)	54.66 (7.76)
Emotional Behavioral Dyscontrol[b,c]	44.41 (8.75)	47.59 (10.52)	47.56 (11.16)	46.30 (10.08)
Physical Functioning—Upper Extremity[a,b,c]	51.64 (5.14)	44.32 (9.30)	32.14 (8.83)	44.73 (10.63)
Physical Functioning—Lower Extremity[a,b,c]	55.81 (4.86)	48.04 (8.45)	39.01 (8.75)	49.27 (9.66)
Applied Cognition—Executive Functioning[a,b,c]	47.41 (9.34)	39.43 (9.15)	27.13 (8.19)	40.09 (11.72)
Applied Cognition—General Concerns[a,b,c]	45.21 (8.85)	40.00 (9.14)	35.59 (8.29)	41.14 (9.57)
Stigma[a,b,c]	43.43 (5.85)	50.15 (7.64)	52.34 (8.28)	47.93 (8.03)
Ability to Participate with Social Roles and Activities[a,b,c]	52.31 (7.93)	48.44 (8.23)	42.94 (7.63)	48.87 (8.69)
Satisfaction with Social Roles and Activities[a,b]	50.26 (6.51)	47.38 (6.18)	45.47 (6.74)	48.14 (6.68)
Generic HRQOL—*M* (SD)				
WHODAS[a,b,c]	3.74 (5.31)	9.43 (8.59)	19.84 (11.20)	9.14 (9.95)
EQ5D Health Scale[a,b,c]	84.30 (11.21)	78.45 (14.68)	73.35 (23.63)	79.50 (16.69)
EQ5D Index Value[a,b,c]	.89 (.12)	.80 (.14)	.71 (.17)	.81 (.15)
Clinician-rated measures				
Motor and Independence—*M* (SD)				
UHDRS Total Motor Score[a,b,c]	5.74 (6.28)	30.70 (14.68)	54.33 (21.28)	26.41 (23.49)
UHDRS Independence Scale[a,b,c]	97.71 (5.95)	85.02 (9.65)	61.40 (12.13)	84.30 (16.62)
Emotional functioning (% impaired)				
PBA-s Anger/Aggression	33.8	41.3	33.6	36.6
PBA-s Anxiety[a,c]	51.5	65.0	50.0	56.3
PBA-s Apathy[a,b]	29.4	51.1	54.6	43.2
PBA-s Depression[a,c]	42.7	61.0	45.1	50.2
PBA-s Irritability[a,c]	49.0	61.0	43.4	52.3
Cognition *M* (SD)				
SDMT Raw Score[a,b,c]	52.14 (10.93)	32.14 (11.59)	18.76 (10.10)	37.83 (16.93)
Stroop Interference Number Correct[a,b,c]	46.37 (15.22)	28.92 (10.00)	18.30 (8.95)	33.75 (16.43)

WHODAS World Health Organization Disability Assessment Schedule 2.0, *EQ-5D* Euro-Qol-5D, UHDRS Unified Huntington's Disease Rating Scales, PBA-s Problem Behaviors Assessment Scale, *SDMT* Symbol Digit Modalities Test, *SF* Short form, *HRQOL* Health-related quality of life, *PROMIS* Patient Reported Outcome Measurement Information System. HDQLIFE, PROMIS and Neuro-QoL scores are reported as *t* scores, all other scores reported are raw scores, except where noted

[a] Univariate analyses indicated significant group differences between prodromal and early-HD

[b] Univariate analyses indicates significant group differences between prodromal and late-HD

[c] Univariate analyses indicates significant group differences between early- and late-HD

advantage that new items can be evaluated for consistency with the original bank and then added at a later date, allowing for future expansion and adjustment.

There are several limitations to this study. First, this study sample was comprised of participants who were recruited through other research studies and through large,

Table 5 Pearson correlations for Neuro-QoL and PROMIS health-related quality of life (HRQOL) measures

	Neuro-QoL									PROMIS		
	Ability to Participate in Social Roles and Activities	Satisfaction with Social Roles and Activities	Applied Cognition—Executive Function	Applied Cognition—General Concerns	Lower Extremity Function—Mobility	Upper Extremity Function—Fine Motor ADL	Stigma	Emotional and Behavioral Dyscontrol	Positive Affect and Well-Being	Anger	Anxiety	Depression
Clinician-administered measures	.37	−.44	.49	.35	−.63	−.70	–	.49	–	.46	.42	.58
	UHDRS Ind. Scale	PBA-s Apathy	Stroop Interference	SDMT	UHDRS Motor	UHDRS Motor		PBA-s Irritability		PBA-s Aggression	PBA-s Anxiety	PBA-s Depression
WHODAS 2.0	−.60	−.66	−.73	−.62	−.74	−.72	.63	.42	−.35	.39	.46	.45
EQ-5D	.45	.50	.55	.48	.61	.58	−.48	−.41	.34	−.39	−.45	−.45

PROMIS Patient-Reported Outcomes Measurement Information System, *ADL* activities of daily living; *UHDRS* unified Huntington's Disease Rating Scales, *PBA-s* Problem Behaviors Assessment Scale, *WHODAS* World Health Organization Disability Assessment Schedule 2.0, *EQ-5D* Euro-QoL-5D

$p < .0001$ for all correlations

established HD clinics; this convenience sample may not represent the HD population at large. Specifically, while there is no evidence to suggest that there are gender differences in HD in the general population [74], our sample included slightly more females (59 %) than males. While consistent with other research studies in HD (females comprise 55–64 % of other large HD cohorts [75, 76]), and in research studies more generally [77–79], it is possible that our findings are not fully representative of males with HD. With regard to education, our prodromal participants were more highly educated than the manifest participants. This is not surprising given that individuals with greater education also have greater medical genetic knowledge [80] and are more likely to get medical testing [81] and that individuals with higher education are more likely to participate in HD research studies [36, 37, 82]. Rates for race/ethnicity were consistent with established prevalence rates [83–86] and other large HD research cohorts [75, 76, 87]. Second, participants completed the survey in multiple ways (online during research visits, online at home, by phone interview, or by in-person interview) and assistance was provided when appropriate (e.g., help logging into the online survey, help clicking the responses). Since a portion of the surveys were completed at home, it is possible that some participant answers were influenced by a person providing assistance with the survey. In addition, a small percentage of participants indicated that their caregiver (family member, or friend) answered questions for them or answered questions by reminding them of important information. A recent high-quality meta-analysis indicates that mode of PRO administration, including completing on paper versus electronically or independently versus with help, does not cause bias [88]; however, future work is warranted to better understand how this may have influenced responding. Third, survey completion allowed multiple sittings, as long as it was generally completed within two weeks of the clinic visit when the UHDRS and PBA-s were administered. Since the survey was not always completed at the same time as the in-clinic assessments, it is possible that the correlations between these measures were less robust than if they had been completed concurrently. Lastly, due to the effect of the disease on cognition, some of the HD participants, particularly those in the late stage of the disease, may not provide reliable self-reports of symptoms and concerns. Furthermore, we know that a small portion of our sample (largely our later-stage participants) were more likely to have incomplete survey data; some of this data loss was due to participant fatigue, while other data loss was due to practical limitations related to exceeding study visit lengths for reserved testing space (and an inability to complete the assessment outside of the clinic visit). Thus, additional work is needed to determine when self-report becomes unreliable [89].

The ultimate utility of the HDQLIFE will depend on its demonstration as a clinically meaningful outcome measure in controlled clinical trials of promising treatments for HD. Data from this study support the utility of the HDQLIFE as a standardized outcome instrument for efficiently capturing HRQOL in HD clinical and research settings. HDQLIFE will be available, free of charge, through www.assessmentcenter.net. Since HD is a relatively rare condition, the CAT platform of HDQLIFE should maximize the effectiveness of clinical trials by minimizing the number of participants needed to detect clinically meaningful changes in levels of function. The ability to conduct cross-disease comparisons may support advances in other neurodegenerative diseases. This should allow researchers to more effectively target interventions that are successful in diseases exhibiting symptom overlap with HD. The HDQLIFE offers a brief and more relevant alternative to current lengthier assessments of HRQOL. The HDQLIFE can also be used in the clinical setting, allowing patients to more effectively communicate symptoms of concern to treatment providers. This can also be accomplished from their home computers, tablets, and smart phones, facilitating better communication with HD specialists who may be geographically far from patients [90]. HDQLIFE provides the next generation of HRQOL measurement specific to HD, a disease that brings unique challenges and thus requires a validated assessment of the aspects of HRQOL that matter to HD patients and their caregivers.

Acknowledgments Work on this manuscript was supported by the National Institutes of Health (NIH), National Institute of Neurological Disorders and Stroke (R01NS077946), and the National Center for Advancing Translational Sciences (UL1TR000433). In addition, a portion of this study sample was collected in conjunction with the Predict-HD study. The Predict-HD study was supported by the NIH, National Institute of Neurological Disorders and Stroke (R01NS040068), the NIH, Center for Inherited Disease Research (provided supported for sample phenotyping), and the CHDI Foundation (award to the University of Iowa). We thank the University of Iowa, the Investigators and Coordinators of this study, the study participants, the National Research Roster for Huntington Disease Patients and Families, the Huntington Study Group, and the Huntington's Disease Society of America. We acknowledge the assistance of Jeffrey D. Long, Hans J. Johnson, Jeremy H. Bockholt, Roland Zschiegner, and Jane S. Paulsen. We also acknowledge Roger Albin, Kelvin Chou, and Henry Paulsen for the assistance with participant recruitment. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

HDQLIFE Site Investigators and Coordinators Noelle Carlozzi, Praveen Dayalu, Stephen Schilling, Amy Austin, Matthew Canter, Siera Goodnight, Jennifer Miner, Nicholas Migliore (University of Michigan, Ann Arbor, MI); Jane Paulsen, Nancy Downing, Isabella DeSoriano, Courtney Shadrick, Amanda Miller (University of Iowa, Iowa City, IA); Kimberly Quaid, Melissa Wesson (Indiana University, Indianapolis, IN); Christopher Ross, Gregory Churchill, Mary Jane Ong (Johns Hopkins University, Baltimore, MD); Susan Perlman, Brian Clemente, Aaron Fisher, Gloria Obialisi, Michael Rosco (University of California Los Angeles, Los Angeles, CA); Michael McCormack, Humberto Marin, Allison Dicke (Rutgers University, Piscataway, NJ); Joel Perlmutter, Stacey Barton, Shineeka Smith (Washington University, St. Louis, MO); Martha Nance, Pat Ede (Struthers Parkinson's Center); Stephen Rao, Anwar Ahmed, Michael Lengen, Lyla Mourany, Christine Reece, (Cleveland Clinic Foundation, Cleveland, OH); Michael Geschwind, Joseph Winer (University of California—San Francisco, San Francisco, CA), David Cella, Richard Gershon, Elizabeth Hahn, Jin-Shei Lai (Northwestern University, Chicago, IL).

Compliance with ethical standards

Conflict of interest *N.E. Carlozzi* currently has research grants from the NIH; she is also supported by grant funding from the NIH, NIDILRR, and CHDI; she declares no conflicts of interest; *S.G. Schilling* has a research grant from NSF. He also is supported by grant funding from NIH. He declares no conflicts of interest; *J.-S. Lai* currently has research grants from the NIH; she declares no conflicts of interest; *J.S. Paulsen* currently has research grants from the NIH; she is also supported by grant funding from NIH, NINDS, and CHDI; she declares no conflicts of interest; *E.A. Hahn* currently has research grants from the NIH; she is also supported by grant funding from the NIH and PCORI, and by research contracts from Merck and EMMES; she declares no conflicts of interest; *J.S. Perlmutter* currently has funding from the NIH, HDSA, CHDI, and APDA. He has received honoraria from the University of Rochester, American Academy of Neurology, Movement Disorders Society, Toronto Western Hospital, St. Luke's Hospital in St. Louis, Emory University, Penn State University, Alberta innovates, Indiana Neurological Society, Parkinson Disease Foundation, Columbia University, St. Louis University, Harvard University and the University of Michigan; *C.A. Ross* declares no conflicts of interest; *N.R. Downing* declares no conflicts of interest; *A.L. Kratz* currently has research grants from the NIH and the Craig H. Neilsen Foundation; she is also supported by grant funding from the NMSS; she declares no conflicts of interest; *M.K. McCormack* currently has grants from the NJ Department of Health; he declares no conflicts of interest. M.A. Nance declares no conflicts of interest; *K.A. Quaid* has research funding from NIH, NIA, NCRR, NINDS and CHDI. She also has funding from HDSA. She has no conflicts of interest to declare; *J.C. Stout* has received research funding in the past three years from the Australian National Health and Medical Research Council, University College London, the CHDI Foundation, Prana Biotechnology, and the University of California, Davis. She is a Director of Stout Neuropsych Pty Ltd, which has received funding from Omeros, Teva Pharmaceuticals, Vaccinex, and Isis. She has been a consultant to Prana Biotechnology and Roche. She receives compensation as a member of the Board of the Huntington's Study Group. *R.C. Gershon* receives research funds from numerous NIH institutes and the Department of Defense. He also receives consulting funds from AO Outcome Center, LLC (a for-profit arm of the nonprofit AO Foundation) and the American Board of Foot and Ankle Surgery; *R.E. Ready* declares that she has no conflicts of interest; *J.A. Miner* is supported by research grants from the NIH; she declares no conflict of interest; *S.K. Barton* is supported by grant funding from the Huntington's Disease Society of America, CHDI Foundation and the NIH. She declares no conflicts of interest; *S.L. Perlman* is supported by grant funding from the NIH, CHDI, FARA, NAF, and several pharmaceutical companies (Edison, Horizon, Pfizer, Reata, Retrotope, Shire, Teva); she declares no conflicts of interest; *S.M. Rao* has received research grants from NIH, Department of Defense, National MS Society, CHDI Foundation and Biogen, and honoraria from the International Neuropsychological Society, Biogen and Genzyme; he declares no conflicts of interest; *S. Frank* receives salary support from the Huntington Study Group for a study sponsored by Auspex Pharmaceuticals. There is no conflict of interest; *I. Shoulson* has received research grants from the Food and

Drug Administration (FDA), National Institutes of Health (NINDS, NHGRI) and the Parkinson's Disease Foundation (NY, NY). He has also received speaker honoraria from the American Academy of Neurology and *JAMA Neurology* as an associate editor. Since May 2014, Dr. Shoulson has been a non-executive director of Prana Biotechnology Ltd (Melbourne, Australia), for which he is compensated for director and consulting services but has no equity positions or stock options in the company. He declares no conflicts of interest as a co-author of the submitted research report. *H. Marin* currently has grants from the NJ Department of Health; he declares no conflicts of interest. *M.D. Geschwind* currently has research grants from the NIH/NIA and Quest Diagnostics; he is also supported by grant funding from Cure PSP and Tau Consortium. He does consulting for MedaCorp, Inc, Gerson-Lehrman Group, Best Doctors, Advance Medical, Inc. and Optio, LLC. He receives compensation for multiple Grand Round lectures. He also gets funding for his research work from the Michael J Homer Family Fund. He discloses no conflicts of interest. *P. Dayalu* currently has research grants from the NIH, Astra-Zeneca, and Vaccinex. He declares no conflicts of interest. *S.M. Goodnight* is supported by grant funding from the NIH and the Craig H. Neilsen Foundation; she declares no conflicts of interest. *D. Cella* receives grant funding from the National Institutes of Health and reports that he has no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. The Huntington's Disease Collaborative Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell, 72*, 971–983.
2. Ross, C. A., et al. (2014). Huntington disease: Natural history, biomarkers and prospects for therapeutics. *Nat Rev Neurol, 10*(4), 204–216.
3. Squitieri, F., et al. (2015). Epidemiology of Huntington disease: First post-HTT gene analysis of prevalence in Italy. *Clinical Genetics, 89*, 367–370.
4. Evans, S. J. W., et al. (2013). Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. *Journal of Neurology, Neurosurgery and Psychiatry, 84*(10), 1156–1160.
5. Paulsen, J. S. (2010). Early detection of huntington disease. *Future Neurology*, *5*(1), 85–104.
6. Ross, C. A., et al. (1997). Huntington disease and the related disorder, dentatorubral–pallidoluysian atrophy (DRPLA). *Medicine (Baltimore), 76*(5), 305–338.
7. Cella, D. F. (1995). Measuring quality of life in palliative care. *Seminars in Oncology, 22*(2 Suppl 3), 73–81.
8. World Health Organization, W. (1946). *Preamble to the constitution of the World Health Organization as adopted by the International Health Conference*. In *International health conference,* New York.
9. Campbell, A. J., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations, and satisfactions*. New York: Russell Sage Foundation.
10. Patrick, D. L., & Erikson, P. (1988). What constitutes quality of life? Concepts and dimensions. *Clinical Nutrition, 7*(2), 53–63.
11. Jankovic, J., & Roos, R. A. (2014). Chorea associated with Huntington's disease: To treat or not to treat? *Movement Disorders, 29*(11), 1414–1418.
12. Peavy, G. M., et al. (2010). Cognitive and functional decline in Huntington's disease: Dementia criteria revisited. *Movement Disorders, 25*(9), 1163–1169.
13. Solomon, A. C., et al. (2007). Verbal episodic memory declines prior to diagnosis in Huntington's disease. *Neuropsychologia, 45*(8), 1767–1776.
14. Ready, R. E., et al. (2008). Patient and caregiver quality of life in Huntington's disease. *Movement Disorders, 23*(5), 721–726.
15. Tibben, A., et al. (1993). Presymptomatic DNA-testing for Huntington disease—Pretest attitudes and expectations of applicants and their partners in the Dutch program. *American Journal of Medical Genetics, 48*(1), 10–16.
16. Hocaoglu, M. B., Gaffan, E. A., & Ho, A. K. (2012). The Huntington's disease health-related Quality of Life questionnaire (HDQoL): A disease-specific measure of health-related quality of life. *Clinical Genetics, 81*(2), 117–122.
17. Paulsen, J. S., et al. (2013). A review of quality of life after predictive testing for and earlier identification of neurodegenerative diseases. *Progress in Neurobiology, 110*, 2–28.
18. Cella, D., et al. (2011). The neurology quality of life measurement (Neuro-QOL) initiative. *Archives of Physical Medicine and Rehabilitation, 92*(Suppl 1), S28–S36.
19. Gershon, R. C., et al. (2012). Neuro-QOL: Quality of life item banks for adults with neurological disorders: Item development and calibrations based upon clinical and general population testing. *Quality of Life Research*, *21*(3), 475–486.
20. Cella, D., et al. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested in its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology, 63*, 1179–1194.
21. van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York: Springer.
22. Choppin, B. (1968). Item bank using sample-free calibration. *Nature, 219*(5156), 870–872.
23. Choppin, B. (1981). Educational measurement and the item bank model. In C. Lacey & D. Lawton (Eds.), *Issues in evaluation and accountability* (pp. 204–221). London: Methuen.
24. Nance, M. A. (2007). Comprehensive care in Huntington's disease: A physician's perspective. *Brain Research Bulletin, 72*(2–3), 175–178.
25. Lai, J. S., et al. (2011). How item banks and its applications can influence measurement practice in rehabilitation medicine: A PROMIS fatigue item bank example. *Archives of Physical Medicine and Rehabilitation, 92*(Supp 1), S20–S27.
26. Carlozzi, N. E., et al. (2016). New measures to capture end of life concerns in Huntington disease: Meaning and purpose and concern with Death and Dying from HDQLIFE (a patient-reported outcomes measurement system). *Quality of Life Research*. doi:10.1007/s11136-016-1354-y.
27. Carlozzi, N. E., et al. (2016). The development of a new computer adaptive test to evaluate chorea in Huntington disease: HDQLIFE Chorea. *Quality of Life Research.*. doi:10.1007/s11136-016-1307-5.
28. Carlozzi, N. E., et al. (2016). HDQLIFE: the development of two new computer adaptive tests for use in Huntington disease, Speech Difficulties, and Swallowing Difficulties. *Quality of Life Research*. doi:10.1007/s11136-016-1273-y.
29. PROMIS® *Instrument Development and Psychometric Evaluation Scientific Standards*. http://www.nihpromis.org/Documents/PROMIS_Standards_050212.pdf.

30. Carlozzi, N. E., & Tulsky, D. S. (2013). Identification of health-related quality of life (HRQOL) issues relevant to individuals with Huntington disease. *Journal of Health Psychology, 18*(2), 212–225.
31. Kisala, P., & Tulsky, D. (2010). Opportunities for CAT applications in medical rehabilitation: Development of targeted item banks. *Journal of Applied Measurement, 11*(3), 315–330.
32. Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, et al. (Eds.), *Cognitive Aspects of survey methodology: Building a bridge between disciplines* (pp. 73–100). Washington, DC: National Academy Press.
33. MetaMetrics. (1995). *The LEXILE framework for reading*. Durham, NC: MetaMetrics Inc.
34. Shoulson, I., & Fahn, S. (1979). Huntington disease—Clinical care and evaluation. *Neurology, 29*(1), 1–3.
35. Hanauer, D. A., et al. (2015). Supporting information retrieval from electronic health records: A report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE). *Journal of Biomedical Informatics, 55*, 290–300.
36. Paulsen, J. S., et al. (2006). Preparing for preventive clinical trials: The predict-HD study. *Archives of Neurology, 63*(6), 883–890.
37. Paulsen, J. S., et al. (2008). Detection of Huntington's disease decades before diagnosis: The Predict-HD study. *Journal of Neurology, Neurosurgery and Psychiatry, 79*(8), 874–880.
38. Paulsen, J. S., et al. (2014). Clinical and biomarker changes in premanifest huntington disease show trial feasibility: A decade of the PREDICT-HD study. *Frontiers in Aging Neuroscience, 6*, 78.
39. Huntington Study Group. (1996). Unified Huntington's Disease Rating Scale: Reliability and consistency. *Movement Disorders, 11*(2), 136–142.
40. Hogarth, P., et al. (2005). Interrater agreement in the assessment of motor manifestations of Huntington's disease. *Movement Disorders, 20*(3), 293–297.
41. Huntington Study Group. (2006). Tetrabenazine as antichorea therapy in Huntington disease: A randomized controlled trial. *Neurology, 66*(3), 366–372.
42. Siesling, S., et al. (1998). Unified Huntington's disease rating scale: A follow up. *Movement Disorders, 13*(6), 915–919.
43. Tabrizi, S. J., et al. (2011). Biological and clinical changes in premanifest and early stage Huntington's disease in the TRACK-HD study: The 12-month longitudinal analysis. *Lancet Neurology, 10*(1), 31–42.
44. Busse, M., et al. (2011). Utilisation of Healthcare and Associated Services in Huntington's disease: A data mining study. *PLoS Currents, 3*, RRN1206.
45. Carlozzi, N. E., et al. (2014). Understanding the outcomes measures used in huntington disease pharmacological trials: A systematic review. *Journal of Huntington's Disease, 3*(3), 233–252.
46. National Institute on Neurological Disorders and Stroke: NINDS Common Data Elements. [cited May 23, 2011]. Available from: http://www.commondataelements.ninds.nih.gov/.
47. Smith, A. (1982). *Symbol digit modalities test: Manual*. Los Angeles: Western Psychological Services.
48. Stroop, J. R. (1992). Studies of interference in serial verbal reactions (Reprinted from *Journal Experimental-Psychology*, Vol. 18, pp. 643–662, 1935). *Journal of Experimental Psychology-General*, **121**(1): 15–23.
49. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.
50. Craufurd, D., Thompson, J. C., & Snowden, J. S. (2001). Behavioral changes in Huntington disease. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology, 14*(4), 219–226.
51. World Health Organization, W. (2012). *The World Health Organization Disability Assessment Scale, WHODAS II*. Available from: http://www.who.int/icidh/whodas/generalinfo.html.
52. Rabin, R., & de Charro, F. (2001). EQ-5D: A measure of health status from the EuroQol group. *Annals of Medicine, 33*(5), 337–343.
53. Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillside, NJ: Erlbaum.
54. De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.
55. PARSCALE. (2003). Scientific Software International Inc.: Lincolnwood, IL. http://www.ssicentral.com/irt/downloads.html.
56. Muthén, L. K., & Muthén, B. O. (2011). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
57. McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
58. Cook, K. F., Kallen, M. A., & Amtmann, D. (2009). Having a fit: Impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Quality of Life Research, 18*(4), 447–460.
59. Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research, 16*(Suppl 1), 19–31.
60. Samejima, F., van der Liden, W. J., & Hambleton, R. (1996). The graded response model. In W. J. van der Liden (Ed.), *Handbook of modern item response theory* (pp. 85–100). NY: Springer.
61. Cai, L., Thissen, D., & du Toit, S. H. C. (2011). *IRTPRO for windows* [Computer software]. 2011, Lincolnwood, IL: Scientific Software International.
62. R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
63. Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte carlo simulations. *Journal of Statistical Software, 39*(8), 1–30.
64. Ustun, T. B., et al. (2010). Developing the World Health Organization Disability Assessment Schedule 2.0. *Bulletin of the World Health Organization, 88*(11), 815–823.
65. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
66. Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores (Psychometric Monograph No. 17)*. Richmond, VA: Psychometric Society.
67. Bryant, F. B., & Yarnold, P. R. (1995). Principal components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & R. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: American Psychological Association.
68. Everitt, B. S. (1975). Multivariate analysis: The need for data, and other problems. *British Journal of Psychiatry, 126*, 237–240.
69. Gorsuch, R. L., & Analysis, Factor. (1983). *Hillsdale*. NJ: Lawrence Erlbaum Associates.
70. Clauser, B. E., & Hambleton, R. K. (1994). Review of differential item functioning, P. W. Holland, H. Wainer. *Journal of Educational Measurement, 31*(1), 88–92.
71. Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
72. Tabrizi, S. J., et al. (2012). Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: Analysis of 24 month observational data. *Lancet Neurology, 11*(1), 42–53.
73. Basch, E. (2010). The missing voice of patients in drug-safety reporting. *New England Journal of Medicine, 362*(10), 865–869.

74. Smith, K. M., & Dahodwala, N. (2014). Sex differences in Parkinson's disease and other movement disorders. *Experimental Neurology, 259*, 44–56.
75. Tabrizi, S. J., et al. (2009). Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: Cross-sectional analysis of baseline data. *Lancet Neurology, 8*(9), 791–801.
76. Paulsen, J. S., et al. (2014). Clinical and biomarker changes in premanifest Huntington disease show trial feasibility: A decade of the PREDICT-HD study. *Front Aging Neurosci, 6*, 78.
77. Dunn, K. M., et al. (2004). Patterns of consent in epidemiologic research: Evidence from over 25,000 responders. *American Journal of Epidemiology, 159*(11), 1087–1094.
78. Burg, J. A., Allred, S. L., & Sapp, J. H, 2nd. (1997). The potential for bias due to attrition in the National Exposure Registry: An examination of reasons for nonresponse, nonrespondent characteristics, and the response rate. *Toxicology and Industrial Health, 13*(1), 1–13.
79. Eagan, T. M., et al. (2002). Nonresponse in a community cohort study: Predictors and consequences for exposure-disease associations. *Journal of Clinical Epidemiology, 55*(8), 775–781.
80. Haga, S. B., et al. (2013). Public knowledge of and attitudes toward genetics and genetic testing. *Genetic Testing and Molecular Biomarkers, 17*(4), 327–335.
81. Roberts, J. S., et al. (2004). Who seeks genetic susceptibility testing for Alzheimer's disease? Findings from a multisite, randomized clinical trial. *Genetics in Medicine, 6*(4), 197–203.
82. Huntington Study Group, P.I., et al. (2016). Clinical-genetic associations in the prospective Huntington at Risk Observational Study (PHAROS): Implications for clinical trials. *JAMA Neurology, 73*(1), 102–110.
83. Pringsheim, T., et al. (2012). The incidence and prevalence of Huntington's disease: A systematic review and meta-analysis. *Movement Disorders, 27*(9), 1083–1091.
84. Folstein, S. E. (1989). *Huntington's disease: A disorder of families*. Baltimore: Johns Hopkins University Press.
85. Hayden, M. R., MacGregor, J. M., & Beighton, P. H. (1980). The prevalence of Huntington's chorea in South Africa. *South African Medical Journal, 58*, 193–196.
86. Narabayashi, H. (1973). Huntington's chorea in Japan: Review of the literature. *Advances in Neurology, 1*, 253–259.
87. Duff, K., et al. (2010). Mild cognitive impairment in prediagnosed Huntington disease. *Neurology, 75*(6), 500–507.
88. Rutherford, C., et al. (2016). Mode of administration does not cause bias in patient-reported outcome results: A meta-analysis. *Quality of Life Research, 25*(3), 559–574.
89. Duff, K., et al. (2010). "Frontal" behaviors before the diagnosis of Huntington's disease and their relationship to markers of disease progression: Evidence of early lack of awareness. *Journal of Neuropsychiatry and Clinical Neurosciences, 22*(2), 196–207.
90. Bull, M. T., et al. (2014). A pilot study of virtual visits in Huntington disease. *Journal of Huntington's Disease, 3*(2), 189–195.