

**PHS PUBLIC ACCESS**

Author manuscript

Qual Life Res. Author manuscript; available in PMC 2019 March 01.

Published in final edited form as:

Qual Life Res. 2018 March ; 27(3): 811–822. doi:10.1007/s11136-017-1755-6.

Evaluating Cognition in Individuals with Huntington disease: Neuro-QoL Cognitive Functioning Measures

Jin-Shei Lai, PhD¹, Siera Goodnight, MPH², Nancy R. Downing, PhD³, Rebecca E. Ready, PhD⁴, Jane S. Paulsen, PhD⁵, Anna L. Kratz, PhD², Julie C. Stout, PhD⁶, Michael K McCormack, PhD⁷, David Cella, PhD¹, Christopher Ross, MD⁸, Jenna Russell, MPH², and Noelle E. Carlozzi, Ph.D.²

¹Department of Medical Social Sciences, Northwestern University, Chicago, IL, USA

²Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI, USA

³College of Nursing, Texas A&M, Bryan, TX, USA

⁴Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, MA, USA

⁵Departments of Psychiatry, Neurology, and Psychological and Brain Sciences, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA

⁶Institute of Cognitive and Clinical Neurosciences, School of Psychological Sciences, Monash University, Clayton, Victoria, AUSTRALIA

⁷Department of Pathology, Rowan University, Piscataway, NJ, USA

⁸Departments of Psychiatry, Neurology, Pharmacology and Neuroscience, Johns Hopkins University, Baltimore, MD, USA

Abstract

Purpose—Cognitive functioning impacts health-related quality of life (HRQOL) for individuals with Huntington disease (HD). The Neuro-QoL includes two patient-reported outcome (PRO) measures of cognition - Executive Function (EF) and General Concerns (GC). These measures have not previously been validated for use in HD. The purpose of this analysis is to evaluate the reliability and validity of the Neuro-QoL Cognitive Function measures for use in HD.

Address reprint requests to: Jin-Shei Lai, Ph.D., OTR/L, Northwestern University Feinberg School of Medicine, Department of Medical Social Sciences, 633 N St Clair, 19th Floor, Chicago, IL 60611, Phone: (312) 503-3370, FAX: 312-503-9800, js-lai@northwestern.edu.

Compliance with Ethical Standards**Disclosure of potential conflicts of interest**

All authors report no disclosures and no conflicts of interest relevant to the manuscript.

Ethical approval

This project was approved by the Institutional Research Board of all participating institutions. All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

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

Methods—Five-hundred-ten individuals with prodromal or manifest HD completed the Neuro-QoL Cognition measures, two other PRO measures of HRQOL (WHODAS 2.0 and EQ5D), and a depression measure (PROMIS Depression). Measures of functioning (the Total Functional Capacity [TFC] and behavior (Problem Behaviors Assessment) were completed by clinician interview. Objective measures of cognition were obtained using clinician-administered Symbol Digit Modalities Test (SDMT) and the Stroop Test (Word, Color, and Interference). Self-rated, clinician-rated, and objective composite scores were developed. We examined the Neuro-QoL measures for reliability, convergent validity, discriminant validity, and known-groups validity.

Results—Excellent reliabilities (Chronbach’s alphas = 0.94) were found. Convergent validity was supported, with strong relationships between self-reported measures of cognition. Discriminant validity was supported by less robust correlations between self-reported cognition and other constructs. Prodromal participants reported fewer cognitive problems than manifest groups, and early-stage HD participants reported fewer problems than late-stage HD participants.

Conclusions—The Neuro-QoL Cognition measures provide reliable and valid assessments of self-reported cognitive functioning for individuals with HD. Findings support the utility of these measures for assessing self-reported cognition.

Keywords

Huntington disease; Cognition; Neuro-QoL; Patient-centered outcomes

Huntington disease (HD) is a hereditary neurodegenerative disorder caused by a CAG triplet repeat expansion in the gene huntingtin. [1–4] Average prevalence rates for HD in North America are estimated to be 7.33 per 100,000 individuals. [5] Individuals with the HD gene expansion typically exhibit cognitive difficulties, which are both insidious and progressive. [6] Cognitive function is associated with health-related quality of life (HRQOL) for persons with HD. [7] HRQOL is a multidimensional construct reflecting the impact that a disease or disability has on mental, physical and social well-being. [8] Although there is no cure for HD, early identification and characterization of cognitive problems may help clinicians provide strategies to patients and their families to adapt their daily activities to improve function. [9; 10]

Investigators and clinicians commonly use standardized clinician-administered cognitive tests to monitor the cognitive status of patients with HD. For example, the Unified Huntington Disease Rating Scale (UHDRS) [11] includes objective tests with demonstrated sensitivity to early cognitive changes, [12] including the Symbol Digit Modalities Test, [13] the Stroop Interference Test, [14; 15] and the Verbal Fluency Test. [16] While neuropsychological tests provide precise measures of how a patient functions cognitively within a controlled environment, they provide more limited information about how an individual functions in day-to-day life, given the demands of the natural environment. [17–20] An alternative approach is to use self-reported cognition. Yet research in other clinical populations (e.g., cancer) suggests that patient perception of their own cognitive function, rather than their cognitive performance per se, is more relevant to HRQOL. [21] Monitoring perceived cognition using psychometrically-sound and clinically valid PRO measures is

important because cognitive complaints have a direct impact upon HRQOL. No disease-specific PRO measure is currently available for individuals with HD.

The Neuro-QoL [22; 23] was initiated by the National Institute of Neurological Disorders and Stroke (NINDS) to develop a HRQOL measurement system for people with neurological conditions. It was developed by gathering input from individuals with neurological conditions and experts, and establishing its psychometric properties using both classical and modern item response theory (IRT) approaches. [22; 24–27] The Neuro-QoL has not previously been validated in individuals with HD. Such a measurement tool may provide a viable assessment of self-reported cognition for this population, but reliability and validity data are needed to support this premise. To meet this need, this study aimed to establish the reliability and validity of the Neuro-QoL Cognitive Function measures in individuals with HD by comparing scores obtained from the Neuro-QoL Cognitive Function measures to those obtained from objective measures via clinician-administered neuropsychological tests, clinician-rated cognition, and self-reported cognition via validated questionnaires. This study also examined convergent validity, discriminant validity, and known-groups validity of the Neuro-QoL Cognitive Function measures using clinical information. As cognitive impairment was suspected as being a result of depression, [28] we evaluated the association of the Neuro-QoL Cognitive Function measures with depressive symptoms. We hypothesized that self-reported cognition tapped domains of both cognition and depression and thus would be correlated with objective and clinician-rated cognition as well as with depression (though with smaller magnitudes of the latter).

METHODS

Sample

This analysis uses data from 510 individuals with prodromal or manifest HD who participated in the HDQLIFE study. For a full description of the study, see Carlozzi, Schilling, Lai, et al., 2016. [29] Participants need to be at least 18 years old, able to read and understand English, and must have a positive test for the HD gene mutation but no clinical diagnosis based on their neurological exam (prodromal) and/or a clinical diagnosis of HD (made by a neurologist; manifest HD). The Total Functional Capacity (TFC) scale from the UHDRS, [30] a 5-item clinician-rated measure (score range: 0–13) with established reliably measuring functional decline with HD disease progression, [31; 32] was used to classify participants with a manifest HD diagnosis as either early-stage (scores: 7–13) or later-stage (scores: 0–6).

Participants were recruited from eight specialized treatment centers across the nation, the National Research Roster for Huntington's disease, existing online medical record data capture systems, [33] and articles/advertisements in HD-specific newsletters and websites, as well as through the Predict-HD research study. [34] Of the 510 participants, 198 had prodromal HD (CAG > 35, but no HD clinical diagnosis based on the UHDRS motor score), 195 had early-stage HD (sum scores of 7–13 on the TFC), and 117 had later-stage HD (sum scores of 0–6 on the TFC). Table 1 shows demographic information. Participants ranged in age from 18–81 years ($M = 49.1$, $SD = 13.3$), 40.8% of participants were male, and most were Caucasian (96.1%). Significant differences were seen for age, which was expected, F

(2, 507) = 46.466, $p < .0001$, since people in the prodromal stage of HD are typically younger than people in the early-stage and the late-stage HD groups. Additionally, the early-stage HD group was younger than the late HD group. Participants' education ranged from 4 to 26 years ($M = 15.1$, $SD = 2.9$). While there were group differences in education, $F(2, 505) = 15.756$, $p < .0001$, these differences were small; early- ($M = 14.7$, $SD = 2.8$) and late-stage HD ($M = 14.2$, $SD = 2.6$) had 1 to 1.5 years less education relative to the prodromal HD group ($M = 15.9$ years, $SD = 2.9$).

This project was approved by the Institutional Research Board of all participating institutions. All participants provided informed consent prior to participation.

Measures

Emotional Functioning Measures—Participants completed PROMIS Depression, [35–38] a self-report measure assessing sadness and hopelessness, using computerized adaptive testing (CAT). This measure is scored on a T-metric (with a mean of 50 and standard deviation of 10); the referent population is the general U.S. population. Higher scores indicate increased depression severity.

A single item from the Problem Behaviors Assessment (PBA-s) [39] was used to represent clinician-rated depression. The PBA-s is a clinician-administered semi-structured interview assessment of behavior that includes 11 items that assess depression, suicidal ideation, anxiety, irritability, aggression, apathy, obsessive-compulsive behaviors, perseverative thinking, paranoid thinking/delusions, hallucinations and disorientation. Each item is rated for both severity and frequency on a 5-point scale. We used the clinician-rated assessment of depression in this study.

Cognitive Measures

Self-reported Cognition: Participants completed *Neuro-QoL Applied Cognition – General Concerns (GC)* and *Neuro-QoL Executive Function (EF)* item banks, which have also been validated in samples of persons with neurological conditions, although not previously in HD. [22] Items included in the GC and EF are listed in the Appendix. The 18-item GC measures perceived difficulties in everyday cognitive abilities such as memory, attention, and decision making, while the 13-item EF emphasizes difficulties in applications of mental function related to planning, organizing, calculating, and working with memory and learning. Both GC and EF were administered as computerized adaptive tests (CATs) and static short forms online. The Neuro-QoL scores were reported using a T-score scoring system, in which the general neurological population mean=50 and standard deviation=10. Higher scores represent better cognitive function. The unidimensionality of these measures have previously been established. [22]

The WHODAS 2.0 [40] consists of 12 items assessing generic function-related HRQOL including: understanding and communication, self-care, mobility, interpersonal relations, work and household roles, and community and civic roles. The WHODAS 2.0 has been validated in an HD sample. [41] Items are rated on a scale of 0 to 4; higher scores indicate poorer health.

Clinician-administered Neurocognitive Tests: A certified cognitive examiner at each site administered the Symbol Digit Modalities Test (SMDT), [13] and the Stroop Test, [42] both of which have previously been associated with functional decline in prodromal and diagnosed HD. [43] The Symbol Digit Modalities Test [SMDT] [13] is a psychomotor measure that examines processing speed. This written test requires the participant to associate numbers and symbols using a key. The score reflects the number of items completed correctly in 90 seconds. Age and education corrected standardized scores ($M = 100$; $SD = 15$) were used in analyses. Higher scores indicate better cognitive functioning. The Stroop Test [42] provides a measure of executive function including cognitive flexibility and resistance to interference (i.e., the ability to inhibit over-learned verbal responses), and consists of three components: Stroop Word, Stroop Color, and Stroop Interference. Scores reflect number correct in 45 seconds; higher scores indicate better performance.

Clinician-rated Cognition: As mentioned above, the (PBA-s) [39] is a clinician-administered semi-structured interview assessment of behavior. In this study, Perseverative Thinking and PBA-s Disoriented Behavior severity scores were used with both of these severity scores reversed so that higher scores indicating better cognitive function (i.e., less perseverative thinking and disoriented behavior, respectively).

Composition scores: Composite scores for each type of cognition measure (neuropsychological test, clinician-rated, and self-reported) described above were generated to examine convergent and discriminant validity between the Neuro-QoL and other cognition measures that were validated in HD. The self-rated composite score was created using the two self-reported cognition items from the WHODAS⁴⁴: “In the last 30 days, how much difficulty did you have in: 1) Learning a new task, for example, learning how to get to a new place?” and 2) Concentrating on doing something for ten minutes?”. The objective composite was created using scores from the SMDT and Stroop Test (Word, Color and Interference). The clinician-rated composite was created using Perseverative Thinking and PBA-s Disoriented Behavior severity scores. All composite scores were created by reverse scoring the items when needed (i.e., higher scores representing better cognition), transforming scores for each measure to z-scores, taking the average of the scores, and transforming the score to a T-score with mean of 50 and SD of 10.

Analyses

Cronbach’s alphas were calculated to evaluate the reliability (internal consistency) of the Neuro-QoL Cognitive function measures (criterion: 0.70 [44; 45]). Floor and ceiling effects were used to describe whether the Neuro-QoL GC and EF sufficiently covered individuals’ perception of their cognition (criterion: proportion of participants with the lowest or the highest possible scores 20% [46; 47]). Correlation coefficients were used to evaluate the relationships between the Neuro-QoL GC and EF versus the three composite cognition scores. We defined convergent validity as high correlations (> 0.6) between the Neuro-QoL GC and EF and the self-reported composite previously described. [48] Discriminant validity would be supported by correlations that were lower in magnitude than those among the convergent validity correlations (by greater than or equal to 0.1 points). [48] Correlations were also calculated to examine the relationship between Neuro-QoL GC

and EF and depressive symptoms reported by individuals with HD and clinicians, in which a small to moderate relationship was expected. We also calculated correlations controlling for depression due to the potential impact of depression on cognitive function.

Analysis of variance (ANOVA) was used to evaluate whether Neuro-QoL GC and EF could significantly differentiate individuals with different stages within the disease (i.e., prodromal, early- or late-stage HD) to evaluate the known-groups validity. Partial eta squared, η^2 , was estimated to determine the strength of EF and GC being a predictor of staging groups. Partial eta squared (η^2) is defined as variance explained by X/(explained variance by X + total unexplained variance of Y). Any variation explained by other independent variables is removed from the denominator. This allows a researcher to compare the effect of the same variable across different studies, which contain different covariates or other factors. η^2 is considered small when its value is between 0.01 (inclusive) and 0.06, medium when between 0.06 (inclusive) and 0.14, and large when $\eta^2 \geq 0.14$. [49; 50]

Criterion validity was examined using a receiver operating characteristic (ROC) analysis by comparing diagnostic performance of Neuro-QoL GC and EF between individuals with prodromal vs. manifest HD. The area under the curve (AUC) values, a measure of discriminatory ability of the test to correctly identify prodromal vs. manifest HD, are interpreted as 0.9 as excellent, 0.8 as good, 0.7 as fair, and < 0.7 as poor. [51] We used a logistic regression model to evaluate how well EF and GC discriminated between prodromal versus manifest HD as well as between participants with and without clinically impaired cognition (defined as 1 SD below the normative mean on the SDMT). This was conducted to determine whether individuals with HD were at greater risk for cognitive function difficulties than the general population. According to the normal curve, 16% of the scores are expected to fall 1 SD below the mean (i.e., impaired); therefore, impairment rates that exceed 16% indicate greater impairment than would be expected compared to demographically-comparable neurologically healthy peers. [52] Likelihood ratios (i.e., sensitivity/[1-sensitivity]) of 2 indicated a minimum standard for differentiating between HD groups. [53] We considered validity and reliability were supported if 75% of the results are in accordance with the hypotheses. [54]

RESULTS

Table 2 provides descriptive data and reliability data for the Neuro-QoL Cognitive Function measures. Average times for individuals to complete measures were between 40 and 69 seconds. Cronbach's Alpha exceeded *a priori* criterion for both Neuro-QoL Cognitive Function measures. Less than 5% (range: 0.8%–1.7%) of participants reported the lowest possible scores (ceiling effect; low functioning) while about 7.2% (GC via CAT) to 18.5% (EF via SF) reported highest possible scores (floor effect; high function). Though floor and ceiling rates met our *priori* criteria, CAT captured EF and GC better than SFs given its smaller ceiling and floor effects. The EF CAT scores showed that floor effects decreased along with disease severity, with 21.5%, 6.3% and 0% for prodromal, early- and late-stage HD, respectively. For GC CAT scores, floor effects were 9.8%, 5.2% and 5.1% for prodromal, early- and late-stage HD, respectively. Results from CAT administration were used for further analyses.

Table 3 shows the correlation matrix of EF, GC, the objective cognition composite, the self-rated cognition composite, the clinician-rated cognition composite, and PROMIS Depression. When data from all participants were analyzed together (Table 3a), correlations ranged from 0.342 (GC versus clinician-rated composite) to 0.739 (EF versus GC). The pattern of correlations supported convergent and discriminant validity. Consistent with proposed hypotheses, EF and GC had high correlations with one another and with the self-rated composite, and lower correlations with the clinician-rated measures (by .10). The highest correlations were among self-report measures, and lower correlations were found among self-rated and objective composites. Correlations for EF ranged from $r = 0.41$ (PROMIS Depression) to $r = 0.75$ (GC); correlations for GC ranged from $r = 0.34$ (clinician-rated composite) to $r = 0.74$ (EF). Findings were similar when depressive symptoms were controlled in the analysis, as shown in Table 3b. When data were analyzed by staging groups (Table 3c), EF was significantly correlated with all measures except objective and clinician-rated composites for late-stage HD. GC was significantly correlated with all measures except objective composite for prodromal and late-stage and clinician-rated composite for late-stage. The prodromal group had larger magnitudes of correlations between GC and EF with other scores than the late-stage group. EF and GC had the largest correlation coefficients with the objective composite in the early-stage group though the correlations were considered weak, 0.20 and 0.19 for EF and GC, respectively. It was noted that GC was moderately correlated with PROMIS Depression across groups. Depression scores were 48.9 (SD=9.3), 51.0 (SD=10.9) and 51.2 (SD=10.9) for prodromal, early-stage and late-stage groups, respectively.

Table 4 and Figure 1 provide support for known-groups validity. As hypothesized, individuals with late-stage HD consistently self-reported worse cognition than the other two groups, and their scores were about 1 SD below the normative population mean. These findings indicated the late-stage HD group had significantly ($p < 0.001$) poorer self-reported cognition than the other two groups. Large effect sizes ($\eta^2 = 0.16$) were found on all but GC, in which a moderate effect size was found ($\eta^2 = 0.10$). Similar conclusions were found when analyses controlled for depressive symptoms (not shown in Table 4).

ROC results showed that the Neuro-QoL EF demonstrated high sensitivity (85.2%) and moderate specificity (55.4%), with an AUC of 0.79 and likelihood ratio of 5.76 for distinguishing prodromal HD versus manifest HD (see Figure 2a). Logistic regression results showed an accurate classification rate of 73.2%. Neuro-QoL GC demonstrated high sensitivity (86.2%) and poor specificity (30.9%), with an AUC of 0.68 and a likelihood ratio of 6.25 for distinguishing prodromal HD versus manifest HD (Figure 2b). The accurate classification rate was 64%.

Neuro-QoL EF demonstrated moderate sensitivity (72.8%) and moderate specificity (62.3%), with an AUC of 0.75 and a likelihood ratio of 2.68 for distinguishing those with and without clinical impairment (239 participants classified as impaired) (Figure 2c); findings met our *a priori* criteria for clinical decision making. The accurate classification rate was 67.8%. Neuro-QoL GC demonstrated high sensitivity (75.7%) and poor specificity (56.4%), with an AUC of 0.69 and a likelihood ratio of 3.12 for distinguishing prodromal

HD versus manifest HD (Figure 2d); again, findings met our *a priori* criteria for clinical decision making. The accurate classification rate was 66.7%.

DISCUSSION

This study provides evidence to support the reliability and validity of the Neuro-QoL GC and EF in individuals with HD, as at least 75% of the results were in accordance with the hypotheses. [54] Floor and ceiling effects of these measures met *a priori* criterion, however, they were less discriminative at the higher functioning levels. In addition, both measures demonstrated significant correlations with neuropsychological tests and clinician-related cognition for individuals with prodromal and early-stage HD; convergent validity was supported by significant correlations among the self-report measures, and less robust correlations among self-report and clinician-rated measures supported discriminant validity. Criterion validity was generally supported by analyses that examined sensitivity and specificity. Specifically, Neuro-QoL EF and GC both met criterion for clinical decision making with regard to being able to differentiate between those with and without manifest HD as well as to differentiate between those with and without manifest HD clinical impairment. Known-groups validity was supported in that individuals with prodromal HD reported better overall cognition than either of the manifest HD groups, and those with early-HD reported better cognition than those with later-stage HD. In addition, these declines in self-report scores across groups tracked with the objective and clinician-rated declines across groups. The fact that EF and GC significantly discriminated among the different groups with moderate (GC) or large (EF) effect sizes suggests that monitoring Neuro-QoL GC and EF scores, along with other tests, in routine follow-up care has the potential to identify those with deteriorating cognition so they can be provided with timely remediation/intervention.

Unlike neuropsychological tests which assess cognition in a controlled environment using structured procedures, the Neuro-QoL GC and EF are designed to capture participant perception of their cognitive decrements and the impact of cognitive decrements on daily activities, respectively. Though GC and EF were highly correlated, we noted that strengths of correlations with other measures varied. As shown in Table 3a, compared to GC, EF showed a trend of having larger magnitudes of correlations with objective and clinician-rated cognitive measures, while GC showed a trend of having a larger magnitudes of correlations with depression than EF did and this relationship was consistently moderate regardless of HD stage. We speculated that this was because most EF items consisted of “concrete” tasks (e.g., “remembering where things were placed or put away (e.g., keys)?”), allowing individuals to respond to these items using their own experiences, while participants might need to provide their subjective impressions of their functioning to respond to “abstract” GC questions (e.g., “I had trouble thinking clearly”). The fact that significant correlations of GC with other cognitive measures remained when controlling for depression indicated GC might tap both depression and cognition. Thus, we speculated that the Neuro-QoL EF and GC measured cognition and depressive symptoms simultaneously to some extent, in which EF tapped more cognition than depression while GC tapped more depression than cognition. Future studies should be done to test this hypothesis.

It was also noted that the strength of their relationship with cognition decreased along with the worsening stage and the relationship between GC and depression remained similar. Additionally, as shown in Table 3c, the fact that the magnitude of correlation coefficients between EF and GC with clinician-rated and objective composites decreased when disease stage got worse suggests that individuals with more advanced stage HD might have more anosognosia due to their worsening cognitive functioning. We thus recommend clinician-rated cognition may provide a more reliable assessment of cognition than self-reported cognition for patients with late-stage HD. Our finding in this study and other work in HD [55] warrant further studies to evaluate whether the same findings can be replicated by using a different sample group.

There are advantages of implementing the validated Neuro-QoL GC and EF in clinical practice. The Neuro-QoL was developed via rigorous qualitative and quantitative approaches and can be administered via static short-forms or computerized adaptive testing (CAT), and scores obtained from both short-forms and CAT are comparable to the scores obtained from the full-length item banks. [56; 57] The promise that CAT testing holds for clinical monitoring has been documented in the literature, [56–58] and meets the clinical needs of individuals with HD. Using this approach, a precise estimate of GC and EF can be obtained with the presentation of only a few items with a short period of time, in this study population within one minute; such brevity is well-suited for individuals with HD. Because of the progressive nature of HD, longitudinal studies of cognition are common, and measures that can briefly and sensitively assess cognition can help patient and families understand their current status and make adjustments in their daily living.

We acknowledge several study limitations. First, this study utilized a convenience sample that targeted individuals who were recruited through other research studies and through large, established HD clinics and may not represent the HD population at large. This sample was primarily Caucasian, and therefore, generalizability to other racial/ethnic groups is uncertain. Furthermore, the majority of our sample had greater than a high school education. While findings may not be as generalizable for those with high school or less education, analyses that focused on clinical impairment utilized an objective cognitive test that corrected for age and education (somewhat mitigating these concerns). Also, more research is needed to establish test-retest reliability and responsiveness to change of the Neuro-QoL cognition measures in HD.

In conclusion, GC and EF as measured by using the Neuro-QoL measurement system were significantly correlated with objective (neuropsychological testing) and clinician-rated cognition in prodromal and early-stage HD. Both Neuro-QoL EF and GC demonstrated high sensitivity in distinguishing cognition reported by individuals with prodromal HD versus manifest HD. We suggest Neuro-QoL EF and GC can be used as complementary sources to objective- and clinician-rated cognition to monitor cognition of patients with prodromal and early-stage HD. The Neuro-QoL measures are ready to be used in research settings and are available in HealthMeasures, <http://www.healthmeasures.net>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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). Dr. Kratz was supported during manuscript preparation by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (1K01AR064275; PI: Kratz). 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, and Roland Zschiegner. 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, Anna L. Kratz, Stephen Schilling, Amy Austin, Matthew Canter, Siera Goodnight, Jennifer Miner, Nicholas Migliore (University of Michigan, Ann Arbor, MI); Jane S. Paulsen, Nancy Downing, Isabella DeSoriano, Courtney Hobart, 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).

Funding/Support

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). Dr. Kratz was supported during manuscript preparation by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (1K01AR064275; PI: Kratz).

References

1. MacDonald ME, Ambrose CM, Duyao MP, Myers RH, Lin C, Srinidhi L, Barnes G, Taylor SA, James M, Groot N. Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*. 1993; 72(6):971–983. [PubMed: 8458085]
2. Squitieri F, Griguoli A, Capelli G, Porcellini A, D'Alessio B. Epidemiology of Huntington disease: first post-HTT gene analysis of prevalence in Italy. *Clinical genetics*. 2016; 89(3):367–370. [PubMed: 25689972]
3. Ross CA, Aylward EH, Wild EJ, Langbehn DR, Long JD, Warner JH, Scahill RI, Leavitt BR, Stout JC, Paulsen JS. Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nature Reviews Neurology*. 2014; 10:204–216. [PubMed: 24614516]
4. Evans SJW, Douglas I, Rawlins MD, Wexler NS, Tabrizi SJ, Smeeth L. Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. *Journal of Neurology, Neurosurgery & Psychiatry*. 2013 jnnp-2012–304636.
5. Rawlins MD, Wexler NS, Wexler AR, Tabrizi SJ, Douglas I, Evans SJ, Smeeth L. The Prevalence of Huntington's Disease. *Neuroepidemiology*. 2016; 46(2):144–153. [PubMed: 26824438]
6. Paulsen JS. Early detection of Huntington's disease. *Future neurology*. 2010; 5(1):85–104.

7. Carlozzi NE, Tulskey DS. Identification of health-related quality of life (HRQOL) issues relevant to individuals with Huntington disease. *Journal of Health Psychology*. 2013; 18(2):212–225. [PubMed: 22427174]
8. U.S. Department of Health and Human Services. *Healthy People 2010: Understanding and Improving Health*. 2. Washington, D.C.: U.S. Government Printing Office; 2000.
9. Mickes L, Jacobson M, Peavy G, Wixted JT, Lessig S, Goldstein JL, Corey-Bloom J. A comparison of two brief screening measures of cognitive impairment in Huntington's disease. *Movement Disorders*. 2010; 25(13):2229–2233. [PubMed: 20721924]
10. Stout JC, Glikmann-Johnston Y, Andrews SC. Cognitive assessment strategies in Huntington's disease research. *Journal of neuroscience methods*. 2016; 265:19–24. [PubMed: 26719240]
11. Kremer HPH. Huntington Study Group. Unified Huntington's disease rating scale: reliability and consistency. *Movement Disorders*. 1996; 11:136–142. [PubMed: 8684382]
12. Stout JC, Paulsen JS, Queller S, Solomon AC, Whitlock KB, Campbell JC, Carlozzi N, Duff K, Beglinger LJ, Langbehn DR, Johnson SA, Biglan KM, Aylward EH. Neurocognitive signs in prodromal Huntington disease. *Neuropsychology*. 2011; 25(1):1–14. [PubMed: 20919768]
13. Smith, A. *Symbol Digit Modalities Test: Manual*. Los Angeles, Calif: Western Psychological Services; 1995.
14. Stroop JR. Studies of Interference in Serial Verbal Reactions (Reprinted from *Journal Experimental-Psychology*, Vol 18, Pg 643–662, 1935). *Journal of Experimental Psychology-General*. 1992; 121(1):15–23.
15. Stroop JR. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*. 1935; 18:643–662.
16. Benton, AL., Sivan, AB., Hamsher, KD., Varney, NR., Spreen, O. *Contributions to neuropsychological assessment*. 2. New York: Oxford University Press; 1994.
17. Conway MA. In defense of everyday memory. *American Psychologist*. 1991; 46(1):19–26.
18. Loftus EF. The glitter of everyday memory ... and the gold. *American Psychologist*. 1991; 46(1):16–18. [PubMed: 1996855]
19. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*. 2003; 13(4):181–197. [PubMed: 15000225]
20. Silver CH. Ecological validity of neuropsychological assessment in childhood traumatic brain injury. *The Journal of head trauma rehabilitation*. 2000; 15(4):973–988. [PubMed: 10876170]
21. Hutchinson AD, Hosking JR, Kichenadasse G, Mattiske JK, Wilson C. Objective and subjective cognitive impairment following chemotherapy for cancer: A systematic review. *Cancer Treatment Reviews*. 2012; 38(7):926–934. [PubMed: 22658913]
22. Cella D, Lai JS, Nowinski C, Victorson D, Peterman A, Miller D, Bethoux F, Heinemann A, Rubin S, Cavasos J, Reder A, Sufit R, Simuni T, Holmes G, Siderowf A, Wojna V, Bode R, McKinney N, Podrabsky T, Wortman K, Choi S, Gershon R, Rothrock N, Moy C. Neuro-QOL: Brief Measures of Health-related Quality of Life for Clinical Research in Neurology. *Neurology*. 2012; 78:1860–1867. [PubMed: 22573626]
23. Cella D, Nowinski C, Peterman A, Victorson D, Miller D, Lai JS, Moy C. The Neurology Quality of Life Measurement (Neuro-QOL) Initiative. *Archives of Physical Medicine and Rehabilitation*. 2011; 92(Suppl 1):S28–S36. [PubMed: 21958920]
24. Hambleton, RK., Swaminathan, H., Rogers, HJ. *Fundamentals of Item Response Theory*. Newbury Park, CA: SAGE Publications, Inc; 1991.
25. Wright, BD., Masters, GN. *Rating scale analysis: Rasch measurement*. Chicago: MESA Press; 1985.
26. Perez L, Huang J, Jansky L, Nowinski C, Victorson D, Peterman A, Cella D. Using focus groups to inform the Neuro-QOL measurement tool: exploring patient-centered, health-related quality of life concepts across neurological conditions. *Journal of Neuroscience Nursing*. 2007; 39(6):342–353. [PubMed: 18186419]
27. Gershon R, Lai J, Bode R, Choi S, Moy C, Bleck T, Miller D, Peterman A, Cella D. 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*. 2012; 21(3):475–486. [PubMed: 21874314]
28. Rock P, Roiser J, Riedel W, Blackwell A. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*. 2014; 44(10):2029. [PubMed: 24168753]
 29. Carlozzi NE, Schilling SG, Lai JS, Paulsen JS, Hahn EA, Perlmutter JS, Ross CA, Downing NR, Kratz AL, McCormack MK, Nance MA, Quaid KA, Stout JC, Gershon RC, Ready RE, Miner JA, Barton SK, Perlman SL, Rao SM, Frank S, Shoulson I, Marin H, Geschwind MD, Dayalu P, Goodnight SM, Cella D. HDQLIFE: development and assessment of health-related quality of life in Huntington disease (HD). *Quality of Life Research*. 2016; 25(10):2441–2455. [PubMed: 27522213]
 30. Shoulson I, Fahn S. Huntington disease clinical care and evaluation. *Neurology*. 1979; 29(1):1–1. [PubMed: 154626]
 31. Marder K, Zhao H, Myers RH, Cudkowicz M, Kayson E, Kiebertz K, Orme C, Paulsen J, Penney JB Jr, Siemers E, Shoulson I. Rate of functional decline in Huntington's disease. *Huntington Study Group. Neurology*. 2000; 54(2):452–458. [PubMed: 10668713]
 32. Paulsen JS, Wang C, Duff K, Barker R, Nance M, Beglinger L, Moser D, Williams JK, Simpson S, Langbehn D, van Kammen DP. Challenges assessing clinical endpoints in early Huntington disease. *Mov Disord*. 2010; 25(15):2595–2603. [PubMed: 20623772]
 33. Hanauer DA, Mei Q, Law J, Khanna R, Zheng K. 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*. 2015; 55:290–300. [PubMed: 25979153]
 34. Paulsen JS, Hayden M, Stout JC, Langbehn DR, Aylward E, Ross CA, Guttman M, Nance M, Kiebertz K, Oakes D, Shoulson I, Kayson E, Johnson S, Penziner E, Investigators PH. Preparing for preventive clinical trials - The Predict-HD study. *Archives of Neurology*. 2006; 63(6):883–890. [PubMed: 16769871]
 35. Amtmann D, Kim J, Chung H, Bamer AM, Askew RL, Wu S, Cook KF, Johnson KL. Comparing CESD-10, PHQ-9, and PROMIS Depression Instruments in Individuals With Multiple Sclerosis. *Rehabilitation Psychology*. 2014; 59(2):220–229. [PubMed: 24661030]
 36. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS): depression, anxiety, and anger. *Assessment*. 2011; 18(3):263–283. [PubMed: 21697139]
 37. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, Devellis R, Dewalt D, Fries JF, Gershon R, Hahn EA, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R, Lai JS. PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*. 2010; 63(11):1179–1194. [PubMed: 20685078]
 38. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M. The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap Cooperative Group During its First Two Years. *Medical Care*. 2007; 45(5 Suppl 1):S3–S11.
 39. Callaghan J, Stopford C, Arran N, Boisse MF, Coleman A, Santos RD, Dumas EM, Hart EP, Justo D, Owen G. Reliability and factor structure of the Short Problem Behaviors Assessment for Huntington's disease (PBA-s) in the TRACK-HD and REGISTRY studies. *Journal of neuropsychiatry and clinical neurosciences*. 2015; 27(1):59–64. [PubMed: 25716488]
 40. Üstün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, Saxena S, Korff Mv, Pull C. Developing the World Health Organization disability assessment schedule 2.0. *Bulletin of the World Health Organization*. 2010; 88(11):815–823. [PubMed: 21076562]
 41. Carlozzi NE, Kratz AL, Downing NR, Goodnight S, Miner JA, Migliore N, Paulsen JS. Validity of the 12-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) in individuals with Huntington disease (HD). *Qual Life Res*. 2015

42. Stroop JR. Studies of interference in serial verbal reactions (Reprinted from *Journal Experimental-Psychology*, Vol 18, Pg 643–662, 1935). *Journal of Experimental Psychology: General*. 1992; 121(1):15.
43. Beglinger LJ, O'Rourke JJ, Wang C, Langbehn DR, Duff K, Paulsen JS. Earliest functional declines in Huntington disease. *Psychiatry Res*. 2010; 178(2):414–418. [PubMed: 20471695]
44. Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. New York: Academic Press; 1988.
45. DeVellis, R. *Scale development: Theory and applications*. 4. Los Angeles, CA: Sage; 2017.
46. Cramer, D., Howitt, DL. *The Sage dictionary of statistics*. Thousand Oaks, CA: Sage; 2004.
47. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine & Rehabilitation*. 2000; 81(12 Suppl 2):S15–20. [PubMed: 11128900]
48. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*. 1959; 56(2):81–105. [PubMed: 13634291]
49. Miles, J., Shevlin, M. *Applying regression and correlation: A guide for students and researchers*. London: Sage Publications, Ltd; 2001.
50. Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, N.J: L. Erlbaum Associates; 1988.
51. Youngstrom EA. A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of pediatric psychology*. 2014; 39(2):204–221. [PubMed: 23965298]
52. Heaton, RK., Miller, SW., Taylor, JT., Grant, I. *Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults*. Lutz, FL: Psychological Assessment Resources, Inc; 2004.
53. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. *The Lancet*. 2005; 365(9469):1500–1505.
54. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007; 60(1):34–42. [PubMed: 17161752]
55. Duff K, Paulsen JS, Beglinger LJ, Langbehn DR, Wang C, Stout JC, Ross CA, Aylward E, Carlozzi NE, Queller S. “Frontal” behaviors before the diagnosis of Huntington’s disease and their relationship to markers of disease progression: evidence of early lack of awareness. *J Neuropsychiatry Clin Neurosci*. 2010; 22(2):196–207. [PubMed: 20463114]
56. Lai JS, Zelko F, Krull K, Cella D, Nowinski C, Manley P, Goldman S. Parent-reported cognition of children with cancer and its potential clinical usefulness. *Quality of Life Research*. 2014; 23(4): 1049–1058. [PubMed: 24197478]
57. Choi S, Reise S, Pilkonis P, Hays R, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. *Quality of Life Research*. 2010; 19(1):125–136. [PubMed: 19941077]
58. Lai JS, Zelko F, Butt Z, Cella D, Kieran MW, Krull KR, Magasi S, Goldman S. Parent-perceived child cognitive function: results from a sample drawn from the US general population. *Child’s Nervous System*. 2011; 27(2):285–293.

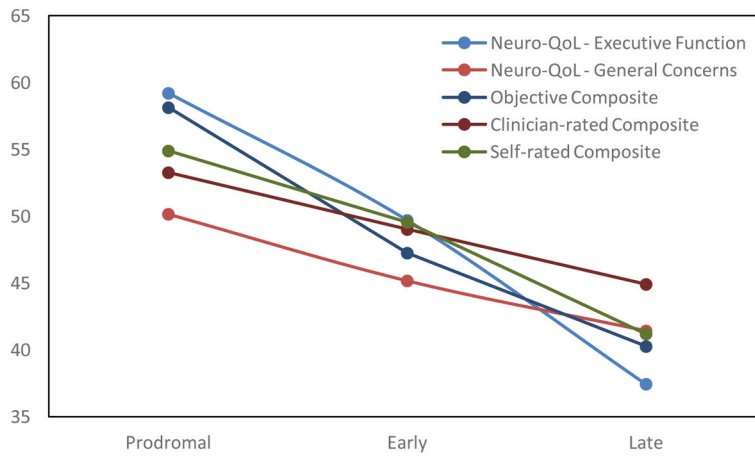
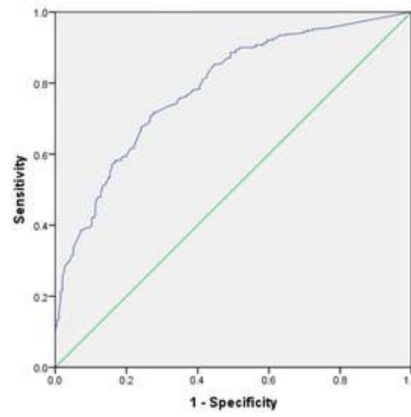


Figure 1. Comparisons of EF, GC, neuropsychological tests, and three composite scores across HD groups (prodromal, early- and late-stage HD)

2a. Executive Function: Prodromal vs. Manifest HD



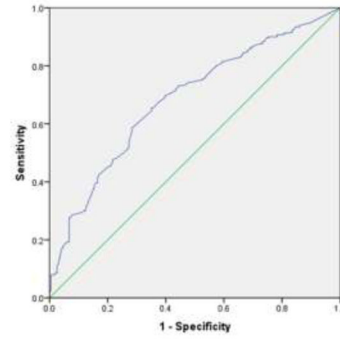
Author Manuscript

Author Manuscript

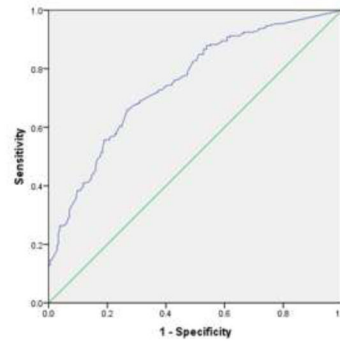
Author Manuscript

Author Manuscript

2b. General Concern: Prodromal vs. Manifest HD



2c. Executive Function: Clinical Impairment



2d. General Concern: Clinical Impairment

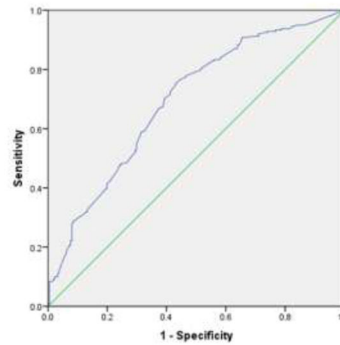


Figure 2. Receiver operating characteristic (ROC) of Neuro-QoL Executive Function (EF) and General Concern (GC)

Table 1

Sample Characteristics

Variable	Prodromal HD (n=198) M(SD)	Early-Stage HD (n=195) M(SD)	Late-Stage HD (n=117) M(SD)	All (N=510) M(SD)	p values
Age (years)	42.7 (12.2)	52.0 (12.4)	55.1 (11.9)	49.1 (13.2)	$p < .0001^a$
Years Since Diagnosis	-	3.1 (3.7)	6.0 (4.6)	4.1 (4.3)	$p < .0001^b$
Education (# of years)	15.9 (2.9)	14.7 (2.8)	14.2 (2.6)	15.1 (2.9)	$p < .0001^c$
Gender (%)					
Female	63.6	54.4	59.8	59.2	$p = .17$
Male	36.4	45.6	40.2	40.8	
Race (%)					
Caucasian	92.4	92.8	97.4	93.7	$p = .09$
Other	2	3.6	6.8	3.7	
Unknown	0.5	0	0	0.2	
Ethnicity (%)					
Not Hispanic or Latino	92.4	92.8	97.4	93.7	$p = .07$
Hispanic or Latino	1.5	4.1	0.9	2.4	
Not Provided	6.1	3.1	1.7	3.9	
Marital Status (%)					
Single, Never Married	16.2	14.9	12.0	14.7	$p = .005^d$
Married	67.2	52.8	61.5	60.4	
Separate/Divorced	13.6	24.6	23.1	20.0	
Widowed	0	3.1	3.4	2.0	
Living with Partner	3.0	4.1	0	2.7	
Unknown	0	0.5	0	0.2	

Note. HD = Huntington disease.

^aThe prodromal HD group was significantly younger than the early- and late-stage HD groups.

^bThe late-stage HD group were significantly more years since diagnosis than the prodromal or early-stage HD group.

^cThe prodromal HD group was significantly more educated than either the early- or late-stage HD group.

^dThe prodromal HD group was significantly more likely to be married than the early-HD group.

Table 2
Descriptive Information and Reliability Data for Neuro-QoL Cognitive Function Measures

Measures	Sample n	Cronbach's α	% of the sample with floor effects (high functioning)	% of the sample with ceiling effects (low functioning)	M (SD)	Administration Time in seconds Mdn (SD)
Neuro-QoL						
Applied Cognition – Executive Function CAT	469	-- ^a	11.2	0.8	50.9 (14.1)	56.0 (93.7)
Applied Cognition – Executive Function SF	469	0.94	18.5	0.8	51.3 (14.6)	69.0 (42.8)
Applied Cognition – General Concerns CAT	470	-- ^a	7.2	0.6	46.4 (10.5)	40.0 (66.4)
Applied Cognition – General Concerns SF	470	0.95	10.5	1.7	45.8 (10.5)	61.0 (44.7)

Note. CAT1 = Computer Adaptive Test; SF = Short Form; higher scores indicate better cognitive function;

^aFor CAT administration, individuals only endorsed items with the highest information on their estimated functional levels and did not complete all items included in the EF and GC. Therefore, alpha was not calculated.

$10.0 > d$
**

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Cognition T-scores by disease stage (known groups analysis)

	Prodromal			Early			Late			F	p	Partial η^2
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)				
Neuro-QoL Executive Function	195	59.21	12.38	187	49.70	11.42	103	37.45	9.67	122.87	<.0001	0.34
Neuro-QoL General Concerns	194	50.16	9.76	187	45.17	9.99	103	41.43	10.45	27.84	<.0001	0.10
Symbol Digit Modalities (standardized score)	195	98.86	16.82	188	71.77	21.49	85	50.86	20.11	204.40	<.0001	0.47
Color Naming (number correct)	194	76.71	13.07	190	52.02	15.44	105	33.97	13.56	338.57	<.0001	0.58
Word Reading (number correct)	196	97.47	16.61	191	65.37	19.28	102	43.14	17.49	343.14	<.0001	0.59
Interference (number correct)	194	46.60	15.38	189	28.91	10.03	93	17.78	8.77	198.59	<.0001	0.46
Objective Composite ^a	191	58.14	5.44	185	47.25	6.11	79	40.26	5.58	323.68	<.0001	0.59
Clinician-rated Composite ^b	185	53.27	5.03	173	49.02	8.13	100	44.91	9.52	42.99	<.0001	0.16
Self-rated Composite ^c	195	54.89	5.57	190	49.56	8.33	98	41.20	9.82	103.54	<.0001	0.30

Note: All scores calculated on a T-metric; general neurological population mean=50 and standard deviation=10. Higher scores represent better cognitive function. In all cases the late-stage HD group indicated more problems than either the early-stage or prodromal group and the early-stage HD group reported more problems than the prodromal group.

^aThe objective composite consisted of scores from Symbol Digit Modalities Test, Color Naming, Word Reading, and Interference

^bThe clinician-rated consisted of scores from PBAs Perseverative Thinking Severity and PBAs Disoriented Behavior Severity scores.

^cThe cognition self-rated composite consisted of scores from WHODAS Scores from Questions “In the last 30 days, how much difficulty did you have in: Learning a new task, for example, learning how to get to a new place?” and “In the last 30 days, how much difficulty did you have in: Concentrating on doing something for ten minutes?”