



Published in final edited form as:

Qual Life Res. 2018 October ; 27(10): 2541–2555. doi:10.1007/s11136-018-1912-6.

Understanding patient-reported outcome measures in Huntington disease: At what point is cognitive impairment related to poor measurement reliability?

N.E. Carlozzi¹, S. Schilling², A.L. Kratz¹, J.S. Paulsen^{3,4,5}, S. Frank⁶, and J.C. Stout⁷

¹Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI, USA

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

³Department of Psychiatry, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA

⁴Department of Neurology, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA

⁵Department of Psychology, The University of Iowa, Iowa City, IA, USA

⁶Beth Israel Deaconess Medical Center, Boston, MA, USA

⁷School of Psychological Sciences, Monash University, Clayton, Victoria, Australia

Abstract

Purpose—Symptom progression in Huntington disease (HD) is associated with cognitive decline which may interfere with the self-report of symptoms. Unfortunately, data to support or refute the psychometric reliability of patient-reported outcomes (PROs) as HD progresses is limited. This is problematic given that PROs are increasingly recognized as important measures of efficacy for new treatments.

Methods—We examined PRO data from the HDQLIFE Measurement System (Speech Difficulties; Swallowing Difficulties; Chorea) in 509 individuals with premanifest, early-stage or

Address reprint requests to Noelle E. Carlozzi, Ph.D., University of Michigan, Department of Physical Medicine & Rehabilitation, North Campus Research Complex 2800 Plymouth Road, Building NCRC B14, Room G216, Ann Arbor, MI 48109-2800, Phone: (734) 763-8917, Fax: (734) 763-7186, carlozzi@med.umich.edu.

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 S. 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; Judy Rokeach (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: The authors have no conflicts of interest to report.

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 Helsinki declaration and its later amendments or comparable ethical standards.

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

late-stage HD. Clinician administered assessments of motor functioning (items from the UHDRS) and standardized objective assessments of cognition (Stroop, Symbol Digit Modalities) were also collected. We examined item bias using differential item functioning (DIF) across HD stage (premanifest, early-, late-) and relative to cognitive performance. We also examined correlations between self-report and clinician ratings. Regression models that considered total cognitive ability were utilized to determine psychometric reliability of the PROs.

Results—Most PRO items were free from DIF for both staging and cognition. There were modest correlations between PROs and clinician report (ranged from -0.40 to -0.60). Modeling analyses indicated that psychometric reliability breaks down with poorer cognition and more progressed disease stage; split-half reliability was compromised (i.e. split-half reliability <0.80) when scores were <136 for Chorea, <109 for Speech Difficulties, and <179 for Swallowing Difficulties.

Conclusions—Results indicate that the psychometric reliability of PROs can be compromised as HD symptoms progress and cognition declines. Clinicians should consider PROs in conjunction with other types of assessments when total cognition scores exceed critical thresholds.

Keywords

Huntington disease; measurement; patient reported outcomes; cognition; HDQLIFE

Introduction

Huntington disease (HD) is an autosomal dominant neurodegenerative disease that affects approximately 1 in 10,000 individuals [1–4]. HD is both insidious and progressive. The motor, cognitive, and psychiatric symptoms that are characteristic of HD emerge gradually and worsen progressively [5] with the majority of individuals meeting criteria for HD diagnosis around age 40 (diagnosis is based solely on the presence of clinically significant motor symptoms) [6]. The typical course until death is ~ 15 – 20 years after clinical diagnosis [7].

Recently, there has been an increased emphasis on utilizing patient-reported outcomes (PROs), especially those that examine health-related quality of life (HRQOL) (i.e., physical, social, and emotional well-being [8]) in clinical research and care [9]. To this end, the HDQLIFE measurement system [10] was developed as a PRO appropriate for individuals with HD. HDQLIFE includes 12 generic HRQOL measures of HRQOL (from the Neuro-QoL [11, 12] and PROMIS [13, 14] measurement systems: Anxiety, Depression, Anger, Positive Affect and Well-Being, Emotional and Behavioral Dyscontrol, Physical Functioning-Upper Extremity, Physical Functioning – Lower Extremity, Applied Cognition – Executive Functioning, Applied-Cognition-General Concerns, Stigma, Ability to Participate in Social Roles and Activities, and Satisfaction with Social Roles and Activities), as well as five new HD-specific measures (Chorea [15], Speech Difficulties [16], Swallowing Difficulties [16], Concern with Death and Dying [17], and Meaning and Purpose [17]). While HDQLIFE is the first comprehensive PRO measurement system that includes both generic and HD-specific aspects of HRQOL, there remain significant concerns

that the cognitive decline that is characteristic of HD [18–30] may preclude the ability to utilize PRO measures in this population.

Unfortunately, there is no data to support or refute the applicability of using PROs throughout all stages of HD. Therefore, the purpose of this study was to determine whether it is appropriate to use PROs when significant cognitive decline is present. Since measurement science dictates that PROs should provide both reliable (i.e., repeatable) and valid (i.e., measure what was intended) information, this study focused on establishing the reliability and validity of PRO measures in individuals across the HD disease spectrum to determine if and during what disease stages these measures meet established measurement science standards for PRO reliability and validity. Optimally, items within the PRO should not exhibit item bias across HD stage nor for cognitive performance. Furthermore, PROs should meet minimal criteria for acceptable psychometric reliability (i.e., $.70$ [31]) and be related to observer reports of similar constructs. Thus, the overall purpose of these analyses was to determine whether and at what stage cognitive impairment and HD disease progression may limit the utility of PRO measures, as evidenced by low reliability and relatively high error variance.

Methods

Participants

We recruited 506 individuals with HD to participate in this study. Participants were recruited through HD specialty clinics (Los Angeles, CA; Iowa City, IA; Indianapolis, IN; Baltimore, MD; Ann Arbor, MI; Golden Valley, MN; St. Louis, MO; Piscataway, NJ), support groups and HD specialized nursing home units, and in conjunction with the PREDICT-HD study. [32] Additional recruitment resources included the National Research Roster for Huntington's Disease, articles/advertisements in HD-specific newsletters and websites online medical record data capture systems [33]. Inclusion criteria were: a positive gene test and/or a clinical diagnosis of HD, ≥ 18 years of age, able to read and understand English, and cognitive capacity to provide informed consent (confirmed by a standardized assessment [34] when in question). All data were collected in accordance with local institutional review boards (University of Michigan Medical School Institutional Review Board, HUM00055669, approved 02/01/2012; Cleveland Clinic Institutional Review Board, IRB 13-460, approved 04/26/2017; Indiana University Institutional Review Board [IRB-01], Protocol 1208009383, approved 09/07/2012; Johns Hopkins Medicine Institutional Review Board, Study NA_00079341, approved 12/13/2012; University of Medicine and Dentistry of New Jersey, subsumed by Rutgers University, Institutional Review Board, Study ID Pro2012002196, approved 04/04/2013; Park Nicollet Institutional Review Board, Study 04334-13-A, approved 11/15/2013; University of California San Francisco Institutional Review Board, IRB 13-10880 Reference 065701, approved 09/04/2013; University of California Los Angeles Institutional Review Board, IRB 12-000743, approved 06/12/2012; University of Iowa Institutional Review Board, IRB ID 201301724, approved 01/17/2013; and Washington University St. Louis Institutional Review Board, IRB ID 201206052, approved 08/14/2012). In addition, participants were required to provide informed consent

prior to study participation. Study participants completed both an in-person assessment and online computer-based assessment comprised of several PROs.

Measures

For the purposes of this study, we examined data from the HDQLIFE measurement system [10], as well as several standardized assessments from the Unified Huntington's Disease Rating Scales (UHDRS) [35]. These assessments were part of a larger study protocol designed to evaluate HRQOL in HD; more details about the full study protocol are detailed elsewhere [10]. For the purposes of this paper, we examined the baseline data from this study.

HDQLIFE [10] PROs—Three physical functioning items banks PROs from HDQLIFE were administered to study participants: HDQLIFE Chorea [15], HDQLIFE Speech Difficulties [16], and HDQLIFE Swallowing Difficulties [16]. HDQLIFE Chorea includes 34 items that assess the impact that chorea (which comprises irregular, random involuntary movements of varying amplitude affecting the face, trunk, and limbs) has on physical activity, participation and HRQOL in individuals with HD. HDQLIFE Speech Difficulties includes 27 items that examine the impact that perceived difficulties in oral expression, language production, and articulation have on communication and general well-being. HDQLIFE Swallowing Difficulties includes 16 items that examine the effect that problems with swallowing (preparatory, oral, and pharyngeal) and choking have on eating and overall well-being. All HDQLIFE PROs are scored on a T metric ($M = 50$, $SD = 10$); higher scores indicate worse self-reported physical function.

UHDRS [35] Clinician-Rated Assessments—Four different assessments from the UHDRS were administered to study participants: Total Functional Capacity items [36], the Total Motor Score (TMS), and two cognitive assessments (The Stroop Color Word Test [37] and Symbol Digit Modalities Test [SDMT; 38]). Total Functional Capacity (TFC) [36] provides an index of day-to-day functioning; scores range from 0–13 with high scores indicating better functional capacity. The TFC was used to determine HD group for manifest participants [39]: sum scores of 7–13 = early-stage HD (stages I–II) and sum scores of 0–6 = late-stage HD (stages III–IV). The TMS provides an index of oculomotor function, dysarthria, chorea, dystonia, gait, and postural stability; higher scores indicate worse motor function. Stroop Color Word Interference [37] provides a measure of psychomotor speed and executive function. There are three different parts to this assessment: Color Naming (which requires participants to name blocks of color [either red, green, or blue] as quickly as they can in 45 seconds), Word Reading (which requires participants to read as many words as they can [either red, green, or blue] in 45 seconds), and Color/Word Interference (which requires participants to name the color of ink that a word [red, green or blue] is written in where the word is written in the wrong color of ink [the word red written in green ink] as quickly as they can in 45 seconds); higher scores on each of these separate components reflect better cognitive performance. The SDMT [38] provides a measure of speed of processing, psychomotor integration and working memory; it requires participants to match symbols to numbers according to a provided key. Higher scores reflect better

cognitive performance. We also created a Total Cognition Score by summing together scores from the Stroop and SDMT raw scores.

Data Analysis

Differential item functioning (DIF) using IRT scaled-score based ordinal logistic regression[40] was used to examine item bias both across HD stage (premanifest HD, early-HD, late-HD) and relative to cognitive performance (Stroop Color Naming, Stroop word Reading, Stroop Interference, and SDMT); analyses were conducted using LORDIF freeware [41]. In general, items should not exhibit DIF. Some degree of uniform DIF (i.e., when one group consistently has advantage across all levels of ability) is considered acceptable. Non-uniform DIF (i.e., when group advantage differs across different levels of ability) was used to flag potentially problematic items. For both HD stage, and cognitive performance, we considered DIF to be negligible if p was $> .01$ and non-negligible when p was $.01$.

Pearson correlations between self-reported motor functioning measures and associated clinician ratings of motor functioning were examined; correlations were examined separately for each HD staging group (premanifest, early-, and late-HD). We expected moderate agreement between self-report and clinician ratings (i.e., r 's between 0.40 – 0.60). A pattern of less robust correlations as HD stage progresses would provide an indication that measurement may be breaking down.

Partial correlations between self-reported motor functioning measures and associated clinician ratings of motor functioning that controlled for total cognition scores were also examined; again correlations were examined separately for each HD staging group (premanifest, early-, and late-HD). A pattern of less robust correlations (relative to the correlations that did not control for cognitive performance) between HD groups would provide additional evidence that measurement may be breaking down.

Three separate sets of regression models were used to examine both the error variance and psychometric reliability of the HDQLIFE PROs through the examination of split-half reliabilities (see Figure 1). The first set of simple linear regression models regressed the second split-half reliability score on the first split-half reliability score for each of the HDQLIFE PROs. For these analyses, we would hypothesize that the majority of the variance should be accounted for (i.e., 90%). In addition, in these simple regression models, the variance, with regard to both staging and cognition would be held constant, and thus we would expect overall model fit to be less robust than for models where variance is allowed to vary (by either staging: second set of models or cognition: third set of models). Thus, we would expect better model fit for the second set of models, heterogeneous variance regression models (i.e., where the variance is allowed to vary for HD stage for each of the separate HDQLIFE PROs (each model was fit with different variances for each HD stage). Similarly, we would expect a better model fit for the third set of heterogeneous variance models, where the variance is allowed to vary for total cognition scores for each of the HDQLIFE PROs. In addition, for each model, we specified that reliability for split-half reliabilities for each PRO should meet minimal acceptable standards (reliability scores can

range from 0.0 – 1.0; 0.70 – 0.79 = acceptable, 0.80 – 0.89 = good, and 0.90 = excellent reliability) [42, 43].

Results

Participants

A total of 509 participants with premanifest and/or manifest HD participated: 197 individuals had premanifest HD (i.e., a positive gene test [CAG = 36], but no HD clinical diagnosis), 196 had early stage HD (sum scores of 7–13 on the TFC), 116 had later-stage HD (sum scores of 0–6 on the TFC). On average, participants were 49.0 years of age (SD = 13.2), and most were Caucasian (95.5%) and female (58.9%). There were significant group differences for age (this was expected since symptoms are progressive with age), $F(2, 506) = 48.35, p < .0001$, with premanifest participants being significantly younger than early- and late-stage participants, and early-stage participants being significantly younger than the late-stage participants. The three groups did not differ by gender, $X^2(2, N = 509) = 3.58, p = .17$, or ethnicity, $X^2(2, N = 489) = 4.168, p = .12$.

Item Bias

Most items were free from item bias with regard to the different cognitive tests and HD staging (Table 1). For HDQLIFE Chorea, no items consistently exhibited DIF for staging or cognitive variables; when present, DIF was negligible. For HDQLIFE Speech Difficulties, 5 items exhibited DIF for both cognition and staging, but none consistently exhibited non-uniform DIF. For HDQLIFE Swallowing Difficulties, 4 items consistently exhibited DIF for both staging and cognitive variables; but none consistently exhibited non-uniform DIF. Taken together, although there were some items for Speech and Swallowing that exhibited uniform DIF, these items did not exhibit non-uniform DIF and the overall magnitude of DIF was minimal.

Pearson and Partial Correlations

As hypothesized, correlations by staging group were lower for those individuals with late-stage HD relative to the other two groups and differences between groups were less robust after controlling for cognitive performance (Table 2).

Regression Models

Estimated split-half reliabilities for the PROs by HD stage (premanifest, early and late) are provided in Table 3; reliability was excellent for premanifest participants, good for early-HD and acceptable for late-HD. As hypothesized, simple regression models indicated that the majority of the variance (R^2) was accounted for when the second split half reliability score was regressed on the first split half reliability score (Table 4). Furthermore, as anticipated, both the models that allowed for heterogeneous variance (for staging and cognition) showed significantly better fit than the simple regression models (Table 5). These results can also be seen in Figures 2, 3 and 4 (which provide graphic representations of the residual data generated for HDQLIFE Chorea, Speech Difficulties, and Swallowing Difficulties, respectively). Specifically, within each figure, overall variability increases both by HD stage (premanifest participant residual scores have less variability than early-, who have less

variability that late-HD participants across the three different PROs), as well as by overall cognition (as total cognition scores decrease, the overall variability increases for residual scores for all three PROs). Critical cutoff scores (for total cognition) for ensuring minimal acceptable split-half reliability are provided in Table 6.

Discussion

Although PROs are gaining influence in both the clinical and research areas, questions remain about the reliability of using self-report measures where cognitive decline is present. As such, we examined the psychometric reliability of several PROs in individuals with HD where cognitive decline is commonly a problem. We sought to determine when it is appropriate to use PROs independently, and when other complementary assessments of functioning and HRQOL are needed. Our findings indicated that with more progressed HD and poorer cognition function, high error variance and low reliability can negatively affect the psychometric properties of PRO measures. We recommend clinical cutoffs (Table 6) for cognition that can be used to aid the researcher and clinician in interpreting PRO data in individuals with HD.

Specifically, findings indicated that most items within the PROs were free from item bias with regard to cognitive and HD disease status (Table 1). For the different PROs, item bias was negligible for HDQLIFE Chorea, and while there was some evidence for bias for a handful of items on HDQLIFE Speech Difficulties and HDQLIFE Swallowing Difficulties, this bias was not systematic minimizing the overall impact that this bias might have on overall clinical interpretation. Furthermore, as expected, correlations among the PROs and clinician-rated symptoms were less robust with more progressed HD stages suggesting increased discordance among self-report and clinician-report with disease progression. This was further supported by correlations that considered cognitive performance; in these cases, group differences were less robust. Similarly, estimated split-half reliability was less robust with each progressive HD stage, again indicating that measurement reliability is lower within increasing levels of HD symptom burden. Furthermore, when this variability in measurement reliability took either cognitive status and/or HD staging into consideration, there was better model fit indicating that both cognitive status and disease stage impacts psychometric reliability of each of the different PROs. Regardless of the combined evidence of decreased psychometric reliability with both increased disease stage and increased cognitive decline, PROs still typically met acceptable standards for reliability (i.e., $> .70$). [31] While these findings would suggest that PROs may remain appropriate for use in later-stage HD, we offer clinical cutoffs for cognitive scores that can be used to maximize PRO reliability among those with cognitive decline (Table 6).

Specifically, the clinical cutoff scores provided in Table 6 can be used to highlight when caution should be utilized in administering PROs in individuals with HD. For example, when using HDQLIFE Chorea to assess the impact that these motor symptoms has on HRQOL in HD, participants with combined cognitive raw scores (Stroop Color Naming raw score + Stroop Word Reading raw score + Stroop Interference raw score + SDMT raw score) of ≥ 77 meet “acceptable” standards for measurement reliability (i.e., $\geq .70$) and those with combined cognitive raw scores of ≥ 136 meet the psychometric standards for “good” (i.e.,

.80) overall test reliability [31]. In cases where cognitive scores do not meet critical cutoffs, PRO scores should only be considered in conjunction with other assessments. It should be noted that these cutoff scores are somewhat imprecise, however, and fail to consider the known influences of age and education.

It is also important to note that the recommended cutoffs vary for each of the different PRO measures. For example, while the minimal acceptable total cognitive raw score = 77 for HDQLIFE Chorea, there is no critical cutoff score for HDQLIFE Speech Difficulties (i.e., all participants in our sample exceeded minimal reliability cutoffs regardless of their cognitive status), and the critical cognitive cutoff score for HDQLIFE Swallowing Difficulties = 134. Thus, while the psychometric reliability of the HDQLIFE Speech Difficulties PRO was never in question for those individuals with HD with poor cognitive performance, there was a critical cutoff score for both Chorea and Swallowing Difficulties. In fact, cognitive performance cutoffs were substantially higher for Swallowing Difficulties (i.e., 134), than it was for HDQLIFE Chorea (i.e., 77). There are a number of potential explanations for this difference. One possibility is that the cognitive complexity (sentence structure, recall burden, etc.) for each of the different PROs is different. This argument is elucidated using the following HDQLIFE Swallowing Difficulties item: “In the past 7 days, how often did you have trouble finishing your meal because of your difficulty swallowing?” This exemplar item requires the participant to consider each meal that he or she had over the course of the past week, then consider if he/she had difficulty finishing the meal, and if yes, was this because of swallowing or some other reason. One can see how the cognitive complexity of this type of question may be problematic for an individual with cognitive difficulties, especially those that are characterized by retrieval or working memory deficits such as is the case in individuals with HD. As such the cognitive status required to answer a complex item such as the provided example, is likely higher than that of a more simplistic item (i.e, the HDQLIFE Speech Difficulties item: “In the past 7 days, it was difficult to speak clearly”). An alternative explanation is that this difference may be due to the nature of the domain itself, as well as associated change over time (or lack thereof). For example, while the rate of progression for swallowing difficulties (as well as speech difficulties) appears to be consistent over the course of the disease,[44, 45] the rate of progression for chorea declines in the more advanced stages.[46–49] As such, it is also possible that the different cognitive cutoff scores may be explained by the fact that chorea might be less problematic for individuals with more advanced HD. In fact, it seems especially plausible that differences in the rates of progression for these symptom domains, in conjunction with associated anosognosia (especially in more advanced disease),[50–54] may contribute to the differential performance of these PRO measures.

While these results highlight several important findings, it is also important to acknowledge study weaknesses. Although this study engaged participants across the United States, this convenience sample may not accurately represent the broader HD community, especially with regard to gender (as this sample included slightly more females than males), education (our participants were generally more educated than the general populations), and race/ethnicity (rates for race/ethnicity were consistent with established prevalence rates [55–58], this sample was primarily Caucasian). In addition, participants were allowed to complete surveys over a two-week time frame of the in-person (i.e., clinician-rated assessments),

which may have contributed to less robust correlations between self-report and clinician-reported functioning.

Regardless of these study limitations, this is the first study that we are aware of that focuses on understanding PRO in a clinical population where both anosognosia[50–54] and cognitive decline[18–30] are present. Not surprisingly, results suggested that high error variance and low reliability can negatively affect the psychometric properties of PRO measures, especially in those participants with late-stage HD and cognitive impairments. Nonetheless, psychometric reliability, although less robust among more progressed participants, typically still met established clinical standards of measurement. As such these measures may still provide valuable information about HRQOL, especially from the participant's perspective. Thus, while we would recommend using PROs throughout the HD disease course, these measures should be used in conjunction with either clinician-rated reports or observer ratings with more advanced stage people; the potential discrepancies between patient-report and other observer-reports can in and of itself provide clinically meaningful information that can help guide treatment recommendations for these individuals.

Acknowledgments

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.

Funding: 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).

References

1. The 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; 72:971–983. [PubMed: 8458085]
2. Ross CA, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nat Rev Neurol*. 2014; 10(4):204–16. [PubMed: 24614516]
3. Squitieri F, et al. Epidemiology of Huntington disease: first post-HTT gene analysis of prevalence in Italy. *Clin Genet*. 2015
4. Evans SJW, et al. Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. *Journal of Neurology Neurosurgery and Psychiatry*. 2013; 84(10):1156–1160.
5. Paulsen JS. Early Detection of Huntington Disease. *Future Neurol*. 2010; 5(1)
6. Walker FO. Huntington's disease. *Lancet*. 2007; 369(9557):218–28. [PubMed: 17240289]
7. Ross CA, et al. Huntington disease and the related disorder, dentatorubral-pallidolusian atrophy (DRPLA). *Medicine (Baltimore)*. 1997; 76(5):305–38. [PubMed: 9352736]
8. Cella DF. Measuring quality of life in palliative care. *Seminars in oncology*. 1995; 22(2 Suppl 3): 73–81. [PubMed: 7537908]

9. Basch E. The Missing Voice of Patients in Drug-Safety Reporting. *New England Journal of Medicine*. 2010; 362(10):865–869. [PubMed: 20220181]
10. Carlozzi NE, et al. HDQLIFE: Development and assessment of health-related quality of life in Huntington disease (HD). *Qual Life Res*. 2016; 25(10):2441–55. [PubMed: 27522213]
11. Cella DF, et al. The Neurology Quality of Life Measurement (Neuro-QOL) Initiative. *Archives of Physical Medicine and Rehabilitation, Supplement*. 2011; 92(Suppl 1):S28–S36.
12. Cella D, et al. Neuro-QOL brief measures of health-related quality of life for clinical research in neurology. *Neurology*. 2012; 78(23):1860–1867. [PubMed: 22573626]
13. Cella DF, et al. 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*. 2010; 63:1179–1194. [PubMed: 20685078]
14. Cella DF, et al. 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.
15. Carlozzi NE, et al. The development of a new computer adaptive test to evaluate chorea in Huntington disease: HDQLIFE Chorea. *Qual Life Res*. 2016
16. Carlozzi NE, et al. HDQLIFE: the development of two new computer adaptive tests for use in Huntington disease, Speech Difficulties, and Swallowing Difficulties. *Qual Life Res*. 2016
17. Carlozzi NE, et al. 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). *Qual Life Res*. 2016
18. Tabrizi SJ, et al. Biological and clinical changes in premanifest and early stage Huntington’s disease in the TRACK-HD study: the 12-month longitudinal analysis. *Lancet Neurology*. 2011; 10(1):31–42. [PubMed: 21130037]
19. Tabrizi SJ, et al. Biological and clinical manifestations of Huntington’s disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. *Lancet Neurology*. 2009; 8(9):791–801. [PubMed: 19646924]
20. Paulsen JS, et al. Clinical and biomarker changes in premanifest Huntington disease show trial feasibility: A decade of the PREDICT-HD study. *Frontiers in Aging Neuroscience*. 2014; 6:78. [PubMed: 24795630]
21. Paulsen JS, et al. Clinical markers of early disease in persons near onset of Huntington’s disease. *Neurology*. 2001; 57(4):658–62. [PubMed: 11524475]
22. Peavy GM, et al. Cognitive and functional decline in Huntington’s disease: dementia criteria revisited. *Mov Disord*. 2010; 25(9):1163–9. [PubMed: 20629124]
23. Beglinger LJ, et al. Cognitive change in patients with Huntington disease on the Repeatable Battery for the Assessment of Neuropsychological Status. *J Clin Exp Neuropsychol*. 2010; 32(6):573–8. [PubMed: 19882420]
24. Stout JC, et al. Evaluation of longitudinal 12 and 24 month cognitive outcomes in premanifest and early Huntington’s disease. *J Neurol Neurosurg Psychiatry*. 2012; 83(7):687–94. [PubMed: 22566599]
25. Williams JK, et al. Everyday cognition in prodromal Huntington disease. *Neuropsychology*. 2015; 29(2):255–67. [PubMed: 25000321]
26. Ward J, et al. A four-year prospective study of cognitive functioning in Huntington’s disease. *J Int Neuropsychol Soc*. 2006; 12(4):445–54. [PubMed: 16981596]
27. Roos RA. Huntington’s disease: a clinical review. *Orphanet J Rare Dis*. 2010; 5(1):40. [PubMed: 21171977]
28. Tabrizi SJ, et al. 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*. 2012; 11(1):42–53. [PubMed: 22137354]
29. Ho AK, et al. Profile of cognitive progression in early Huntington’s disease. *Neurology*. 2003; 61(12):1702–6. [PubMed: 14694033]
30. Kirkwood SC, et al. Progression of symptoms in the early and middle stages of Huntington disease. *Arch Neurol*. 2001; 58(2):273–8. [PubMed: 11176966]

31. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960; 20(1):10.
32. Paulsen JS, et al. Detection of Huntington's disease decades before diagnosis: the Predict-HD study. *Journal of Neurology, Neurosurgery & Psychiatry*. 2008; 79(8):874–880.
33. Hanauer DA, et al. 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. Jackson W, TA N. Effective serial measurement of cognitive orientation in rehabilitation: the Orientation Log. *Arch Phys Med Rehabil*. 1998; 79(6):718–720. [PubMed: 9630156]
35. Huntington Study Group. Unified Huntington's Disease Rating Scale: Reliability and consistency. *Movement Disorders*. 1996; 11(2):136–42. [PubMed: 8684382]
36. Shoulson I, Fahn S. Huntington disease clinical care and evaluation. *Neurology*. 1979; 29(1):1–1. [PubMed: 154626]
37. 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.
38. Smith, A. *Symbol Digit Modalities Test: Manual*. Los Angeles, Calif.: Western Psychological Services; 1995.
39. Marder K, et al. Rate of functional decline in Huntington's disease. *Neurology*. 2000; 54(2):452–452. [PubMed: 10668713]
40. Crane PK, et al. Differential item functioning analysis with ordinal logistic regression techniques. DIFdetect and difwithpar. *Med Care*. 2006; 44(11 Suppl 3):S115–23. [PubMed: 17060818]
41. Choi SW, Gibbons LE, Crane PK. 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*. 2011; 39(8):1–30.
42. Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd. Academic Press; New York: 1988.
43. DeVellis, R. *Scale development: Theory and applications*. In: Bickman, L, Rog, DJ, editors *Applied social research methods series*. 4th. Los Angeles, CA: Sage; 2017.
44. Heemskerk AW, Roos RA. Dysphagia in Huntington's disease: a review. *Dysphagia*. 2011; 26(1): 62–6. [PubMed: 20838817]
45. Kagel MC, Leopold NA. Dysphagia in Huntington's disease: a 16-year retrospective. *Dysphagia*. 1992; 7(2):106–14. [PubMed: 1533361]
46. Reilmann R, Leavitt BR, Ross CA. Diagnostic criteria for Huntington's disease based on natural history. *Mov Disord*. 2014; 29(11):1335–41. [PubMed: 25164527]
47. Ross CA, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nature Reviews Neurology*. 2014; 10:204–216. [PubMed: 24614516]
48. Dorsey ER, et al. Natural history of Huntington disease. *JAMA Neurol*. 2013; 70(12):1520–30. [PubMed: 24126537]
49. Pagan F, Torres-Yaghi Y, Altshuler M. The diagnosis and natural history of Huntington disease. *Handb Clin Neurol*. 2017; 144:63–67. [PubMed: 28947126]
50. McCusker E, Loy CT. The many facets of unawareness in huntington disease. *Tremor Other Hyperkinet Mov (N Y)*. 2014; 4:257. [PubMed: 25411649]
51. Deckel AW, Morrison D. Evidence of a neurologically based “denial of illness” in patients with Huntington's disease. *Arch Clin Neuropsychol*. 1996; 11(4):295–302. [PubMed: 14588934]
52. Vitale C, et al. Unawareness of dyskinesias in Parkinson's and Huntington's diseases. *Neurol Sci*. 2001; 22(1):105–6. [PubMed: 11487181]
53. Chatterjee A, et al. A comparison of self-report and caregiver assessment of depression, apathy, and irritability in Huntington's disease. *J Neuropsychiatry Clin Neurosci*. 2005; 17(3):378–83. [PubMed: 16179661]

54. Duff K, et al. "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]
55. Pringsheim T, et al. The incidence and prevalence of Huntington's disease: a systematic review and meta-analysis. *Mov Disord*. 2012; 27(9):1083–91. [PubMed: 22692795]
56. Folstein, SE. Huntington's disease: A disorder of families. Johns Hopkins University Press; 1989.
57. Hayden MR, MacGregor JM, Beighton PH. The prevalence of Huntington's chorea in South Africa. *South African medical Journal*. 1980; 58:193–196. [PubMed: 6447364]
58. Narabayashi H. Huntington's chorea in Japan: Review of the literature. *Advances in neurology*. 1973; 1:253–259.

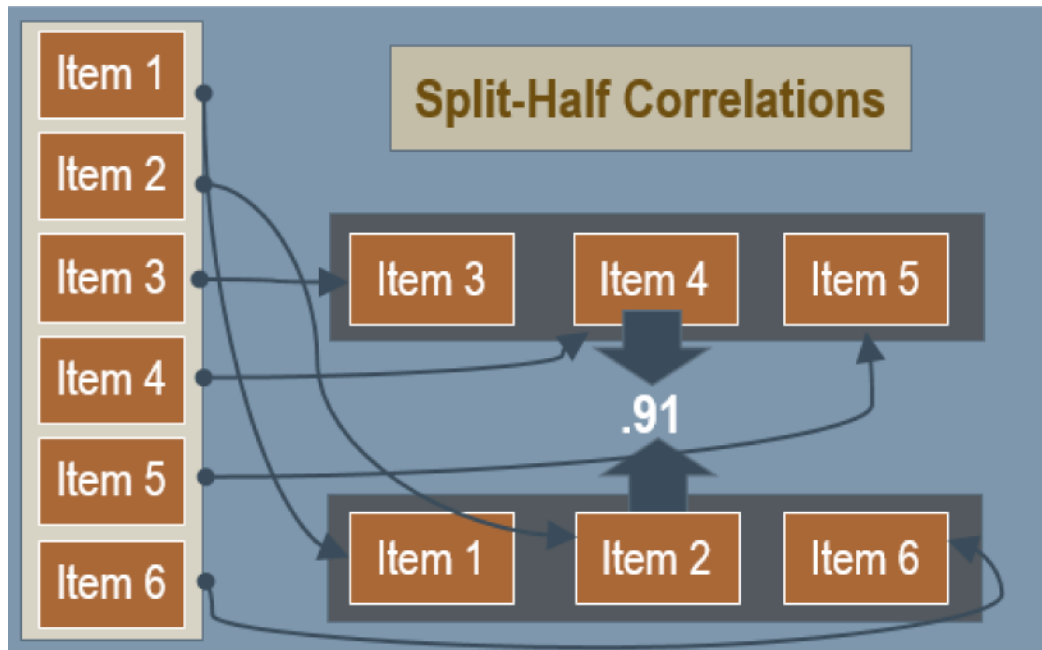


Figure 1. Example of split-half correlations

This figure provides a pictorial example of split-half reliability; for a 6-item measure, 3 items would be randomly selected and compared with the remaining 3 items to determine the consistency of results across items within the test.

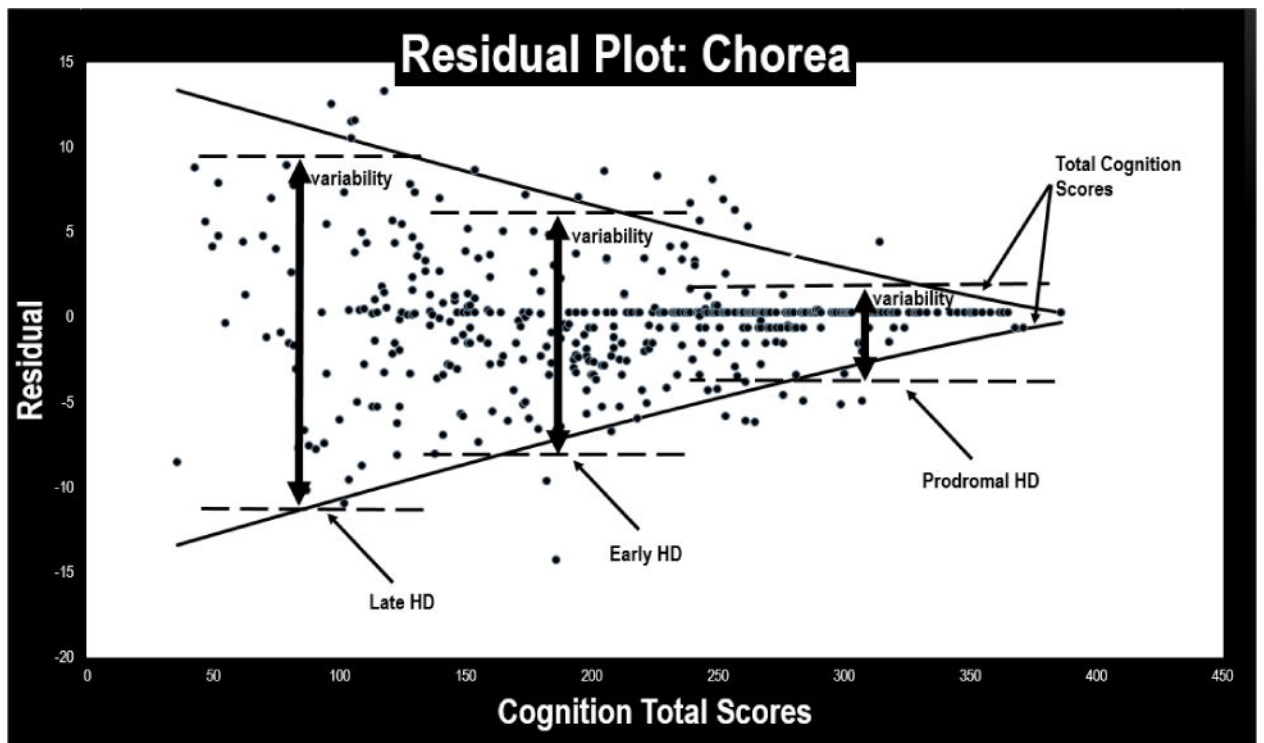


Figure 2. Residual Plot for HDQLIFE Chorea

This figure provides a scatterplot of the residual scores for the HDQLIFE Chorea measure: variability increases both by HD stage (premanifest participants have less variability in residual scores than early-, who have less variability than late-HD participants), and overall cognition (as total cognition scores decrease, the overall variability in residual scores increases).

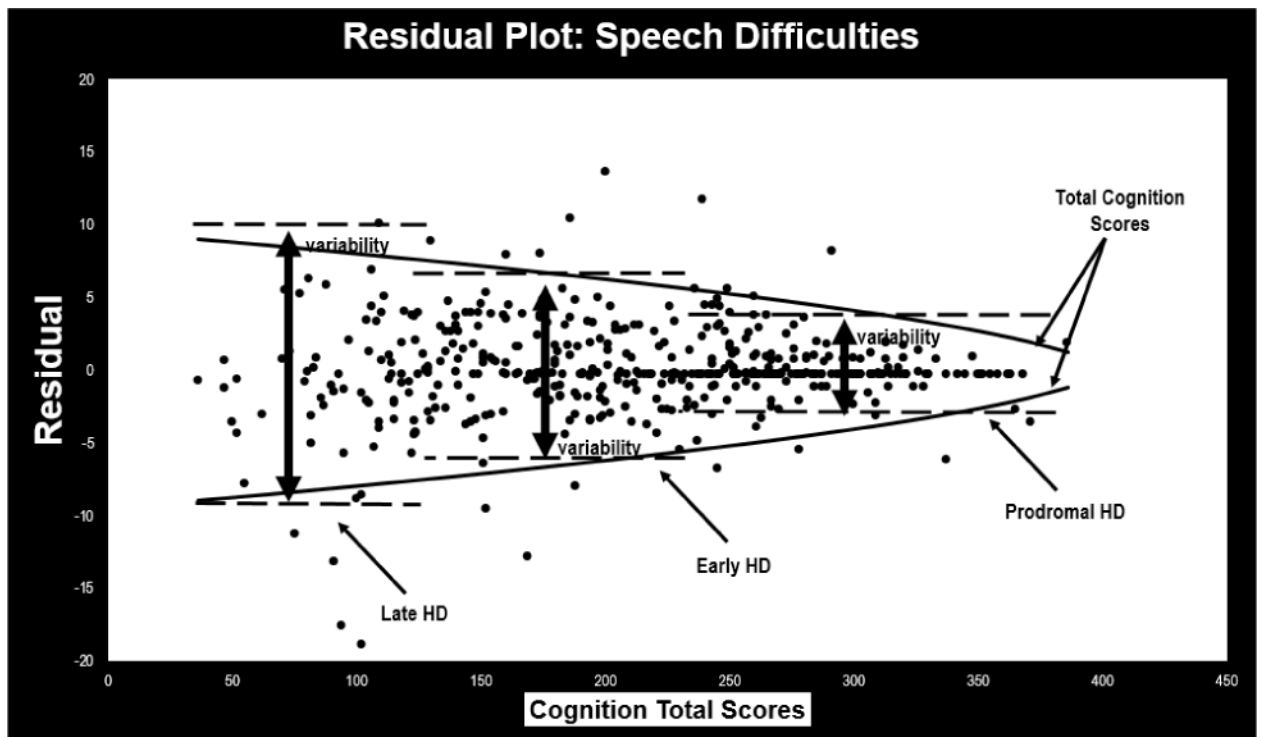


Figure 3. Residual Plot for HDQLIFE Speech Difficulties

This figure provides a scatterplot of the residual scores for the HDQLIFE Speech Difficulties measure: variability increases both by HD stage (premanifest participants have less variability in residual scores than early-, who have less variability than late-HD participants), and overall cognition (as total cognition scores decrease, the overall variability in residual scores increases).

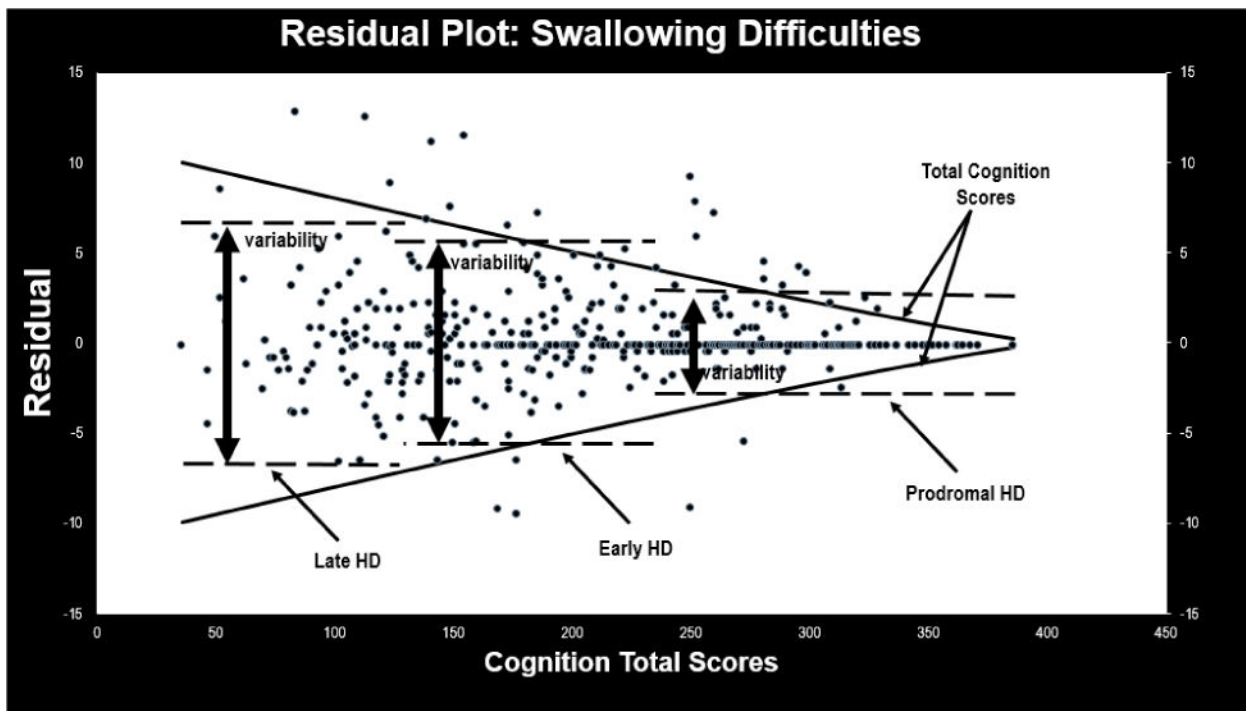


Figure 4. Residual Plot for HDQLIFE Swallowing Difficulties
 This figure provides a scatterplot of the residual scores for the HDQLIFE Swallowing Difficulties measure: variability increases both by HD stage (premanifest participants have less variability in residual scores than early-, who have less variability than late-HD participants), and overall cognition (as total cognition scores decrease, the overall variability in residual scores increases).

Table 1

Differential Item Functioning Analyses for HDQLIFE PROs

Item	SDMT	Differential Item Functioning			HD Disease Stage (premanifest, early-late-)
		Stroop Color Naming	Stroop Word Reading	Stroop Color/Word Interference	
HDQLIFE CHOREA ITEM BANK					
How often did your movements (e.g., chorea) impact your ability to enjoy the things you do for fun?					
How often did your movements (e.g., chorea) impact your ability to exercise?	Non-Uniform				
How often did your movements (e.g., chorea) interfere with your ability to do errands?	Uniform Positive			Uniform Positive	
How often did your movements (e.g., chorea) interfere with your ability to do your household chores?	Uniform Positive			Non-Uniform	
How often did your movements (e.g., chorea) interfere with your ability to eat?					
How often did your movements (e.g., chorea) interfere with your ability to participate in recreational activities?					
How often did your movements (e.g., chorea) interfere with your ability to socialize with your family					
How often did your movements (e.g., chorea) interfere with your ability to socialize with your friend	Uniform Negative	Uniform Negative	Uniform Negative		Uniform Negative **
How often did your movements (e.g., chorea) interfere with your ability to take a bath or shower?					
How often did your movements (e.g., chorea) interfere with your physical activities?					
How often did your movements (e.g., chorea) limit your physical activities?					
How often did your movements (e.g., chorea) make you fall?					Uniform Positive *
How often did your movements (e.g., chorea) prevent you from leaving the house?		Non-Uniform	Non-Uniform	Uniform Positive	Uniform Positive **
How often were you less effective at home due to your movements (e.g., chorea)?					

Item	SDMT	Differential Item Functioning			HD Disease Stage (premanifest, early-late-)
		Stroop Color Naming	Stroop Word Reading	Stroop Color/Word Interference	
How severe was your chorea on average?					
I had to limit my physical activity because of my movements (e.g., chorea).					
I had to limit my social activity because of my movements (e.g., chorea).					
I had trouble finishing things because of my movements (e.g., chorea).					
I had trouble starting things because of my movements (e.g., chorea).					
My movements (e.g., chorea) impacted my ability to bathe or shower.	Uniform Positive	Non-Uniform	Non-Uniform		
My movements (e.g., chorea) impacted my ability to get dressed.					
My movements (e.g., chorea) impacted my ability to eat.					
My movements (e.g., chorea) impacted my ability to feed myself.					
My movements (e.g., chorea) impacted my ability to walk.					
What was the severity of your movements (e.g., chorea) on most days?					
How often did your movements (e.g., chorea) impact your ability to hold things, like a glass or fork?					
How often did you feel unsteady when you were standing?		Uniform Positive			
How often did you limit your physical activities because of your movements (e.g., chorea)?					Uniform Negative **
How often did you limit your social activities because of your movements (e.g., chorea)?					Uniform Negative **
How often did your movements (e.g., chorea) interfere with your ability to get dressed?					Uniform Negative *
How often did your movements (e.g., chorea) interfere with your ability to walk?			Non-Uniform		
How often did your movements (e.g., chorea) interfere with your social activities?					

Item	SDMT	Differential Item Functioning			HD Disease Stage (premanifest, early-late-)
		Stroop Color Naming	Stroop Word Reading	Stroop Color/Word Interference	
How often did your movements (e.g., chorea) limit you at work (include work at home)?	Uniform Positive	Uniform Positive	Uniform Positive	Uniform Positive	Non-Uniform
I needed help doing my usual activities.	Uniform Positive	Uniform Positive	Uniform Positive	Uniform Positive	Non-Uniform
HDQLIFE SPEECH DIFFICULTIES ITEM BANK					
It was difficult for other people to understand me.	Uniform Positive**	Uniform Positive**	Uniform Positive**	Uniform Positive**	Uniform Positive**
How often did you slur your words when you spoke?	Uniform Positive	Non-Uniform	Uniform Positive*	Uniform Positive*	Non-Uniform
It was difficult to speak clearly.	Uniform Positive*	Uniform Positive*	Uniform Positive*	Uniform Positive*	Non-Uniform
How much difficulty do you currently have... saying what you want to say?	Uniform Positive*	Uniform Positive**	Uniform Positive**	Uniform Positive**	Uniform Negative**
How much difficulty do you currently have... speaking clearly?	Uniform Positive*	Uniform Positive**	Uniform Positive**	Uniform Positive**	Uniform Negative**
How much difficulty do you currently have... speaking?	Uniform Positive*	Uniform Positive**	Uniform Positive**	Uniform Positive**	Non-Uniform
How often did you have to speak slowly for other people to understand you?	Uniform Negative	Uniform Negative	Uniform Negative	Uniform Negative	Non-Uniform
How often did you limit your social activities because you had difficulty speaking?	Uniform Negative	Uniform Negative	Uniform Negative	Uniform Negative	Non-Uniform
How often did you mumble?	Uniform Negative*	Uniform Negative**	Uniform Negative**	Uniform Negative**	Uniform Positive**
How often did you worry about slurring your words when you spoke?	Uniform Negative*	Uniform Negative**	Uniform Negative**	Uniform Negative**	Uniform Positive**
How often did your speech difficulties interfere with your ability to socialize with your family?	Uniform Negative	Uniform Negative	Uniform Negative	Uniform Negative	Uniform Positive**
How often did you speech difficulties impact your ability to enjoy life?	Uniform Negative*	Uniform Negative**	Uniform Negative**	Uniform Negative**	Uniform Positive**
How often did your speech difficulties impact your ability to socialize with your family?	Uniform Negative*	Uniform Negative**	Uniform Negative**	Uniform Negative**	Uniform Positive**
How often did your speech difficulties impact your ability to socialize with your friends?	Uniform Negative*	Uniform Negative**	Uniform Negative**	Uniform Negative**	Uniform Positive**
How often did other people have difficulty understanding you?	Non-Uniform*	Uniform Positive**	Uniform Positive**	Uniform Positive**	Uniform Negative**
How often were you bothered by slurring your words when you spoke?	Uniform Negative*	Uniform Negative*	Uniform Negative*	Uniform Negative*	Uniform Positive*

Item	SDMT	Differential Item Functioning			HD Disease Stage (premanifest, early-late-)
		Stroop Color Naming	Stroop Word Reading	Stroop Color/Word Interference	
How often were you unable to maintain a conversation?	Non-Uniform	Non-Uniform**	Non-Uniform*	Non-Uniform*	
How often did you have to make an effort to carry on a conversation?					
How often were you bothered by your speech difficulties?			Uniform Negative		
My speech difficulties interfered with my ability to work (include work at home).					
During the past 7 days...I had to limit my social activity because of my speech difficulties					
During the past 7 days...I had to speak slowly for other people to understand me.					
During the past 7 days...I had trouble speaking.				Uniform Negative*	
During the past 7 days...I had trouble speaking clearly.					
During the past 7 days...I slurred my words when I spoke.	Uniform Positive**	Uniform Positive**	Uniform Positive*		Non-Uniform
During the past 7 days...I was frustrated by my speech difficulties.				Uniform Negative	
During the past 7 days...My speech difficulties made me feel self-conscious.				Uniform Negative*	Uniform Positive*
HDQLIFE SWALLOWING DIFFICULTIES ITEM BANK					
How often did you choke?					Uniform Positive*
How much difficulty do you currently have... swallowing?	Uniform Positive**	Uniform Positive*	Uniform Positive*	Uniform Positive**	Uniform Positive**
How much difficulty do you currently have... chewing?	Non-Uniform**	Non-Uniform	Non-Uniform	Uniform Positive*	Non-Uniform
How much difficulty do you currently have... eating?	Non-Uniform**	Non-Uniform	Non-Uniform*	Uniform Positive**	Non-Uniform*
I limited how much I ate because it was difficult to swallow.	Non-Uniform**				
How often did you have trouble finishing your meal ...?					
How often did you worry about choking?					

Item	Differential Item Functioning				HD Disease Stage (premanifest, early-late-)
	SDMT	Stroop Color Naming	Stroop Word Reading	Stroop Color/Word Interference	
How often did choking interfere with your ability to eat?					
How often did difficulty chewing interfere with your ability to eat?	Non-Uniform	Non-Uniform*	Non-Uniform**	Non-Uniform*	
How often were you bothered by your choking?					
How often were you unable to swallow?					
How often did you choke on average?					
I had trouble chewing.	Non-Uniform*	Uniform Positive*	Uniform Positive*	Uniform Positive*	Uniform Negative*
I had trouble eating because I choked.	Uniform Positive*				
I had to eat slowly to avoid choking.	Non-Uniform*	Non-Uniform	Non-Uniform	Uniform Positive**	Non-Uniform**

Note.

* = $p < .01$;

** = $p < .001$

Table 2

Correlations between PROs and clinician-rated motor functioning by HD group

Composite Scores	HDQLIFE PRO Measure											
	Chorea	Speech	Swallowing	Chorea	Speech	Swallowing	Chorea	Speech	Swallowing	Chorea	Speech	Swallowing
	Premanifest			Early-HD			Late-HD					
UHDRS Total Motor Score (simple correlation)	.40	.22	.31	.31	.21	.27	.22	.18	.27	.22	.18	.07
UHDRS Total Motor Score (partial correlation controlling for Total Cognition scores)	.35	.14	.25	.25	.15	.19	.27	.11	.19	.27	.11	.10

Note. PRO = patient-reported outcome; UHDRS = Unified Huntington's Disease Rating Scales

Table 3

Estimated PRO Split-Half Reliabilities by HD Stage

HDQLIFE PRO	Premanifest HD	Early-Stage HD	Late-Stage HD
Chorea	0.98	0.86	0.72
Speech Difficulties	0.98	0.85	0.69
Swallowing Difficulties	0.95	0.79	0.71

Note. PRO = patient-reported outcome; HD = Huntington disease

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Simple regression models that examine the ability for the first split half reliability score to predict the second split half reliability score

HDQLIFE PRO Measure	beta	R²	t
Chorea	0.97	0.94	83.22
Speech Difficulties	0.92	0.92	70.93
Swallowing Difficulties	1.24	0.84	49.18

Note. all $p < .0001$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Model fit results for HDQLIFE PRO measures for different regression models

Model	df	AIC	BIC	Chi-Square
HDQLIFE CHOREA				
Simple Regression	3	2291.75	2303.90	
Cognition	4	2112.89	2129.09	180.86*
HD Stage	5	2096.93	2117.18	198.81*
HDQLIFE SPEECH DIFFICULTIES				
Simple Regression	3	2330.83	2343.12	
Cognition	4	2239.79	2256.19	93.03*
HD Stage	5	2206.81	2227.30	128.02*
HDQLIFE SWALLOWING DIFFICULTIES				
Simple Regression	3	2136.32	2148.63	
Cognition	4	1993.47	2009.88	144.85*
HD Stage	5	2038.47	2058.94	101.85*

Note. AIC = Akaike information criterion; BIC = Bayesian information criterion; for AIC and BIC, smaller numbers indicate better model fit;

* = $p < .0001$

Table 6

Total Cognition Cutoffs for Reliability for Different HDQLIFE PROs

HDQLIFE PRO Measure	Reliability < 0.7	Reliability < 0.8
Chorea	< 77	< 136
Speech Difficulties	NA	< 109
Swallowing Difficulties	< 134	< 179

Note. M = 144.56 (SD = 77.31) for Total Cognition Scores for the combined sample

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript