

# **HHS Public Access**

Author manuscript Qual Life Res. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Qual Life Res. 2016 October ; 25(10): 2429–2439. doi:10.1007/s11136-016-1307-5.

# **The development of a new computer adaptive test to evaluate chorea in Huntington disease: HDQLIFE Chorea**

**N. E. Carlozzi**1, **N. R. Downing**2, **S. G. Schilling**1,3, **J.-S. Lai**4,5, **S. M. Goodnight**1, **J. A. Miner**1, and **S. A. Frank**<sup>6</sup>

<sup>1</sup> Department of Physical Medicine and Rehabilitation, University of Michigan, North Campus Research Complex, 2800 Plymouth Road, Building NCRC B14, Room D037, Ann Arbor, MI 48109-2800, USA

2 College of Nursing, The University of Iowa, Iowa City, IA, USA

<sup>3</sup> Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

4 Center on Outcomes, Research and Education, Evanston Northwestern Healthcare, Northwestern University, Evanston, IL, USA

<sup>5</sup> Institute for Health Services Research and Policy Studies, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

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

# **Abstract**

**Purpose—**Huntington's disease (HD) is an autosomal dominant neurodegenerative disease associated with motor, behavioral, and cognitive deficits. The hallmark symptom of HD, chorea, is often the focus of HD clinical trials. Unfortunately, there are no self-reported measures of chorea.

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

Compliance with ethical standards

**Conflict of interest** Carlozzi, N.E. currently has research grants from the NIH; she is also supported by grant funding from the NIH, NIDILRR, and CHDI; she declares no conflicts of interest. Downing, N.R. declares no conflicts of interest. Schilling, S.G. has a research grant from NSF. He is also supported by grant funding from NIH. He declares no conflicts of interest. Lai J.-S. currently has research grants from the NIH; she declares no conflicts of interest. Goodnight, S.M. is supported by grant funding from the NIH and the Craig H. Neilsen Foundation; she declares no conflicts of interest. Miner, J.A. is supported by research grants from the NIH; she declares no conflict of interest. Frank, S. receives salary support from the Huntington Study Group for a study sponsored by Auspex Pharmaceuticals. There is no conflict of interest.

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

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

To address this shortcoming, we developed a new measure of chorea for use in HD, HDQLIFE Chorea.

**Methods—**Qualitative data and literature reviews were conducted to develop an initial item pool of 141 chorea items. An iterative process, including cognitive interviews, expert review, translatability review, and literacy review, was used to refine this item pool to 64 items. These 64 items were field tested in 507 individuals with prodromal and/or manifest HD. Exploratory and confirmatory factor analyses (EFA and CFA, respectively) were conducted to identify a unidimensional set of items. Then, an item response theory graded response model (GRM) and differential item functioning analyses were conducted to select the final items for inclusion in this measure.

**Results—**EFA and CFA supported the retention of 34 chorea items. GRM and DIF supported the retention of all of these items in the final measure. GRM calibration data were used to inform the selection of a 6-item, static short form and to program the HDQLIFE Chorea computer adaptive test (CAT). CAT simulation analyses indicated a 0.99 correlation between the CAT scores and the full item bank.

**Conclusions—**The new HDQLIFE Chorea CAT and corresponding 6-item short form were developed using established rigorous measurement development standards; this is the first selfreported measure developed to evaluate the impact of chorea on HRQOL in HD. This development work indicates that these measures have strong psychometric properties; future work is needed to establish test–retest reliability and responsiveness to change.

#### **Keywords**

Health-related quality of life; Neuro-QoL; PROMIS; HDQLIFE; Huntington's disease; Chorea; Patient-reported outcome (PRO); Motor function

# **Introduction**

Huntington's disease (HD) is an autosomal dominantly inherited neurodegenerative disease resulting from a trinucleotide expansion of cytosine-adenosine-guanine (CAG) in the HTT gene at 4p16.3 [1]. While HD involves cognitive and behavioral decline, the hallmark symptom of HD is chorea. Chorea is an abnormal involuntary movement defined as irregular, flowing movements of the face, trunk, and limbs of varying amplitude [2]. Age at onset of motor symptoms varies, with average age around 40 years, and involves progressive decline and premature death approximately 15–20 years later [3]. When the motor signs of HD begin, chorea may be subtle. As the disease progresses, chorea may become larger in amplitude and frequency, causing injury and impairing physical functioning. In addition, chorea, in combination with the other motor impairments associated with HD, is associated with increased injuries and falls [2, 4] and increases the risk of nursing home placement [5]. Choreic movements are often stigmatizing, as they are commonly mistaken for drunkenness [6]. Importantly, motor impairment and chorea are associated with lower self-reported HRQOL [7–10]. Although there is no cure for HD, medications for chorea can temporarily reduce the frequency and amplitude of movements. However, side effects from commonly used medications such as somnolence, akathisia, dysphagia, gait issues, and apathy may

negatively impact HRQOL [2]. Unfortunately, there are no measures to evaluate the impact of chorea on HRQOL.

Assessment of motor symptoms in HD is typically conducted by a trained motor rater using the Unified Huntington Disease Rating Scale (UHDRS) Motor scale [11]. The UHDRS Motor scale captures several aspects of the neurological examination that are often compromised in HD (i.e., voluntary and involuntary movements, eye movements, gait, and balance). Scores reflect a Total Motor Score (TMS) as well as a diagnostic confidence level (DCL) based on motor rater judgment of whether motor signs demonstrate unequivocal signs of HD. While the DCL is a reliable measure of motor diagnosis and the TMS is a good prognostic indicator of onset within 5 years of diagnosis [12], the TMS is a clinician-rated scale that does not include subjective input from the individual with HD. Thus, the TMS does not evaluate the impact of chorea on HRQOL. In fact, there are no existing patientreported outcome (PRO) measures that examine chorea. This gap in measurement is especially problematic given the Food and Drug Administration (FDA) recommendation to include patient-reported outcome (PRO) measures in clinical trials in order to support claims of a treatment's efficacy in improving HRQOL [13].

To this end, recent efforts have been devoted to developing state-of-the-art PRO assessments through the Neuro-QoL [14, 15] and patient-reported outcome measurement information system (PROMIS) [16, 17] HRQOL measurement systems. These measurement systems were developed to create standardized outcome measures for use in clinical trials that are based on patient report for persons with chronic diseases and neurological disorders [16, 18]. Because HD involves the unique motor impairment, chorea, it is important that PROs used in HD clinical trials include a HRQOL PRO related to chorea. Therefore, the purpose of this study was to develop a PRO measure of self-reported HRQOL related to chorea using established PROMIS methodology [19].

# **Methods**

#### **Participants**

Five hundred and seven individuals with prodromal HD (CAG 36, but did not yet have an HD clinical diagnosis) and/or manifest HD (had clinical diagnosis of HD) were included in our sample. Participants had to be at least 18 years of age and able to read and comprehend English, and have the ability to provide informed consent. Participants were recruited from specialized HD treatment centers at the University of Michigan, the University of Iowa, the University of California-Los Angeles, Indiana University, Johns Hopkins University, Rutgers University, Struthers Parkinson's Center, and Washington University. Participants were also recruited through electronic medical records [20], the National Research Roster for Huntington's Disease, articles/advertisements in HD-specific newsletters and Web sites, and the Predict-HD study [21].

#### **HDQLIFE Chorea item pool**

We began with an initial pool of 141 questions that were designed to evaluate how chorea impacts function and HRQOL using the same methodology used by the Neuro-QoL and

PROMIS [19]. Specific item content was based on focus group discussion among individuals at risk of HD, those with prodromal HD, individuals with manifest HD, nonprofessional caregivers of individuals with HD (e.g., family members), and professionals working with individuals with HD ( $n = 6$  groups with symptomatic HD;  $n = 5$  with individuals at risk of or prodromal for HD;  $n = 3$  non-clinical caregivers;  $n = 2$  groups with HD clinicians) [10]. Items were further refined via results from expert review, evaluation of item literacy level, and patient cognitive review to ensure adequate content coverage, appropriate reading levels, and comprehension level. The final item pool included 64 questions designed to evaluate how chorea affects HRQOL [22].

### **Clinician-rated measures**

Several measures from the UHDRS [11] were administered to all participants. This included the Total Functional Capacity (TFC) scale [23] which is a 5-item clinician-rated measure that provides an index of total functional capacity. Specifically, these five items evaluate day-to-day functioning across the domains of occupation, finances, domestic chores, activities of daily living, and care level. Scores range from 0 (lowest level of functioning) to 13 (highest level of functioning). In this study, TFC scores were used to classify participants with an HD diagnosis as either early stage (sum scores of 7–13) or later stage (sum scores of 0–6). In addition, participants completed the UHDRS motor examination, which includes 15 clinician-rated items designed to evaluate oculomotor function, dysarthria, motor task sequencing, rhythmic tapping, chorea, dystonia, gait, and postural stability. Scores range from 0 (no motor difficulties) to 120 (greater motor difficulties).

#### **Statistical analyses**

All analyses were conducted according to the established PROMIS measurement development guidelines [19]. We used factor analysis to examine the unidimensionality of items. We first randomly divided the sample into two datasets: one for exploratory factor analysis (EFA;  $n = 254$ ) and the other for confirmatory factor analysis (CFA;  $n = 253$ ). We used EFA with a PROMAX rotation to determine the number of factors within the item pool using the following criteria: (1) eigenvalues  $>1$  and (2) the number of factors before the break in the scree plot. Item loadings were used to determine items and their associated factor (criterion > 0.3). CFA was conducted to confirm the factor structure from EFA using the following criteria: (1) comparative fit index  $(CFI) > 0.90$ , (2) root mean square error of approximation (RMSEA) <0.1 [24–27], and (3) residual correlations <0.15 [28–30]. EFA and CFA analyses were conducted using MPLUS 6.11 [31].

Parameters of items that met unidimensional criteria were estimated by using an IRT model —Samejima's graded response model (GRM) [32]. Item parameters were used to estimate information functions at the level of individual items and at the level of the entire item bank and to characterize the precision of items and the overall scale on the measurement continuum. Differential item functioning (DIF) was used to evaluate the stability of an item's measurement properties across subgroups within certain variables using IRT-scaled scorebased ordinal logistic regression [33]; these analyses were implemented using LORDIF freeware [34]. Variables used for this study were gender, age ( $40 \text{ vs. } > 40$ ;  $50 \text{ vs. } > 50$ years), and education (high school graduate or less vs. >high school). Items that showed

significant DIF (criterion:  $p < 0.01$ ) of non-negligible magnitude (R2  $> 0.02$ ) in more than one comparison were candidates for removal from the chorea measure due to potential measurement bias. IRT-scaled scores were generated using the GRM and then converted into a standardized score utilizing a *t* metric (mean = 50,  $SD = 10$ ); these standardized scores were used for the rest of the analyses. IRTPRO 2.1 software was used to conduct these analyses [35]. CAT simulations were conducted using Firestar CAT simulation software [36].

#### **Preliminary validation data**

Pearson's correlations between the new HDQLIFE Chorea measure and the UHDRS Total Motor Score were calculated to examine convergent validity. To demonstrate adequate convergent validity, correlations between these measures should be moderate to large  $(r =$ 0.5–0.8) [37]. A univariate analysis was conducted to determine whether there were significant differences among the HD groups (prodromal vs. early-, vs. late-stage HD) on HDQLIFE Chorea. Tukey's honestly significant difference (HSD) post hoc analyses were used to identify significant between groups effects. We expect significant differences among the three groups with prodromal HD reporting less chorea-related functional difficulties than either early- or late-stage HD participants, and early-stage HD participants reporting less chorea-related functional difficulties than late-stage HD participants.

# **Results**

#### **Participants**

A total of 507 individuals with prodromal and/or manifest HD participated: 196 individuals had prodromal HD (CAG 36, but did not yet have an HD clinical diagnosis), 193 had earlystage HD (sum scores of 7–13 on the TFC), 117 had later-stage HD (sum scores of 0–6 on the TFC), and one individual was not classifiable (due to missing information); see Table 1 for a summary of demographic information. The mean age was  $49.01$  years (SD = 13.21; range 18–81), and the majority of participants were Caucasian (96.4 %) and female (59.2 %). Average education was 15.06 years (SD = 2.88; range 4–26). As expected, there were significant group differences for age (as symptoms are progressive with age),  $F(2, \theta)$  $503$ ) = 47.360,  $p < 0.0001$ . Prodromal participants ( $M = 42.60$ , SD = 12.04) were significantly younger than early-stage  $(M = 51.91, SD = 12.41)$  and late-stage participants  $(M = 55.07, SD = 11.89)$ , and the early-stage participants were younger than the late-stage individuals. There were no group differences for gender,  $X^2(2, N = 506) = 3.193$ ,  $p = 0.20$ , or ethnicity,  $X^2(2, N = 486) = 4.300$ ,  $p = 0.12$ . There were very small group differences for education,  $F(2, 501) = 14.781$ ,  $p < 0.0001$ ; early-stage HD ( $M = 14.74$ , SD = 2.78) and latestage HD ( $M = 14.22$ , SD = 2.62) had 1–1.5 years less education relative to prodromal HD participants ( $M = 15.88$  years, SD = 2.94).

#### **EFA and CFA findings**

EFA findings suggested that the data could be explained by five factors (Table 2). Factor 1 included 31 items that involve specific impact of chorea on various aspects of physical and social functioning. Factor 2 included 30 items that examined the impact of chorea on physical, social, and emotional functioning (11 items had substantial cross-loadings on

Factor 1). Factor 3 consisted of 13 items concerning tremors and shaking (2 items crossloaded on Factor 1; 2 items cross-loaded on Factor 2; 1 item cross-loaded on Factors 2 and 4). Factor 3 was not included for further consideration, as these items are more indicative of parkinsonism movements (i.e., tremors and shakiness) than chorea (i.e., fluid and dance-like movements). Factor 4 consisted of 16 items concerning chorea frequency and severity and the impact of chorea on physical and emotional functioning (4 items cross-loaded on Factor 1; 8 items cross-loaded on Factor 2; 1 item cross-loaded on Factor 3). Factor 5 consisted of 2

items involving the impact of chorea on driving (1 item cross-loaded on Factor 1). Factor 5 was not included for further consideration as two items are not appropriate for consideration as a scale.

Conceptually, Factors 1, 2, and 4 had substantial overlap (all three factors included items that reflected the impact/effect that chorea had on overall functioning). Thus, we focused on Factor 1 (which included the most items and accounted for the largest amount of variance), and retained items from Factors 2 and 4 that reflected the impact that chorea had on either physical or social functioning. Thus, 40 items were retained for further consideration in the CFA (we deleted 7 items from Factors 2 and/or 4 that reflected emotional functioning, 4 items from Factors 2 and 4 that had higher cross-loadings on Factor 3 or 5, and 5 items from Factor 4 that reflected chorea severity).

The initial CFA with the remaining 40 items revealed that 6 items had large residual correlations. These 6 items were removed in a subsequent confirmatory factor analysis with the 34 remaining items. Results from that analysis indicated that all 34 items examined fit the data well; CFI = 0.98, TLI = 0.98, RMSEA = 0.07, all  $r^2 > 0.03$ . In addition, all residual correlations were <0.15. Cronbach's alpha for this scale was 0.98, and all item-total correlations were >0.7.

#### **IRT analyses**

IRT parameter estimates for the 34 items indicated slopes ranging from 2.64 to 6.21 and thresholds ranging from –0.39 to 2.13 (Table 3). Information was good for scaled scores between  $-0.6$  and  $+2.7$  (see Fig. 1 for the scale information function), and the marginal reliability was 0.88. A 6-item calibrated short form was then selected using item calibration statistics (e.g., slope, item characteristic curves, item information, and average item difficulty), as well as input on clinical characteristics (e.g., items were selected that represent different clinical components of chorea difficulties). Specifically, we balanced the psychometric considerations with clinical content to ensure representativeness of the items that were selected for the short form (see Table 3).

We also examined differential item functioning (DIF) to ensure that selected items do not perform differently for different subgroups of participants when they should not (i.e., with relation to gender, age, and education). Specifically, items did not demonstrate DIF for age  $( $50 \text{ vs. } 50 \text{ and }  $40 \text{ vs. } 40$ ), gender (male vs. female), or education (some college and$$ lower vs. college degree and higher).

#### **Short-form scores**

The IRT-scaled scores (thetas) were converted into a standardized score utilizing a t metric (mean  $= 50$ ,  $SD = 10$ ); scores are based on HD sample means. Table 4 shows a summed score conversion table. The short-form scores had a marginal reliability of 0.78.

#### **CAT simulation**

The correlation between the CAT scores and the full item bank was 0.99, indicating that simulated CAT administration can produce results that are very similar to those obtained with administration of the entire 34-item set. Figure 2 shows the number of CAT items used for different scale scores in standard deviation units:  $At -1 SD$  units, the CAT always used all 34 items in the item bank; from +0.3 to +2.0 SD units, the CAT always used the minimum number of 4 items in the item bank; and at 3 SD units, the CAT used all 34 items in the item bank. Thus, the CAT simulation indicates that fewer items were needed to estimate scores for individuals with greater chorea than for individuals with less chorea.

#### **Preliminary validation data**

There was a significant positive correlation between HDQLIFE Chorea and the UHDRS Total Motor Score ( $r = 0.64$ ,  $p < 0.0001$ ) providing support for convergent validity. Univariate analysis also indicated significant group differences for HDQLIFE Chorea,  $F(2)$ , 489) = 159.2,  $p < 0.0001$ . Tukey's HSD analyses indicated that prodromal participants (M =  $43.45$ ,  $SD = 3.81$ ) reported significantly lower chorea-related functional problems than early-stage ( $M = 51.59$ , SD = 7.80) and late-stage participants ( $M = 57.05$ , SD = 8.20), and early-stage HD individuals indicated lower chorea-related functional problems than latestage HD individuals.

# **Discussion**

This study was designed to develop a new PRO that could sensitively evaluate the impact that chorea has on HRQOL in individuals with HD. The new HDQLIFE Chorea item bank includes a total of 34 items that evaluate the impact that chorea has on physical and social functioning. In addition, a corresponding 6-item short form was selected by a team of experts in chorea, HD, and measurement development. The chorea item bank and corresponding short form are available free of charge alongside PROMIS measures at [www.assessmentcenter.net.](http://www.assessmentcenter.net) Clinicians and researchers using this measure can generate scores on a t metric that indicates how his/her patient is functioning relative to other individuals with HD; higher scores indicate more self-reported chorea. In general, scores of 60 or above indicate that an individual is reporting significant concerns in physical and or social functioning due to his/her chorea (i.e., scores are higher than 68.27 % of individuals with HD). Scores of 70 or above indicate that the reported concerns are greater than 95.45 % of individuals with HD. Thus, any score  $\,60$  should warrant clinical follow-up. In addition, HD staging and other contextual variables (e.g., clinically rated motor functioning) should also be considered when interpreting these HDQLIFE scores. For example, for prodromal HD, scores >47 (which is 1 SD above the prodromal HD mean) may indicate a level of personal distress/reported functional impairment that warrants further consideration (especially in the absence of clinician-rated motor symptoms). For both early- and late-stage

HD, scores within normal limits (i.e., within  $\pm 1$  SD of mean) should be considered in conjunction with other contextual factors to determine what additional action, if any, is warranted. In this manner, the scoring metric for this measure provides clinical information that can be used to help guide clinical decision making and referrals.

The HDQLIFE Chorea item bank has several strengths. First, it was developed using wellestablished, state-of-the-art methodology [19] and meets established psychometric standards; it is a homogenous item set with excellent reliability. Items are also devoid of age, gender, and education bias. In addition, there is preliminary support for both convergent and known-groups validity. The HDQLIFE is also the first PRO system to include an assessment of chorea, and it is also the first time that CAT technology has been used to assess HRQOL in HD [22, 38, 39]. CAT offers several advantages to traditional test administration format in that only the most relevant items are administered to each participant, minimizing participant burden without sacrificing overall sensitivity. Furthermore, the corresponding static short form offers a more traditional test administration format, but since each item was developed and selected using IRT, each individual test item provides meaningful information. Thus, even if a participant only answers a single question, a meaningful score can still be derived (albeit the standard error of this score will be large). Finally, as mentioned above, the scoring is based on a t metric, which allows for straightforward interpretation of scores that are more than 1 SD above the mean (i.e., 60 or above).

While this study exhibits several strengths, we also acknowledge a number of weaknesses. First, although CAT administration is generally more efficient than traditional administration approaches, responders at either extreme end of the chorea spectrum (i.e., either very significant chorea or no chorea) may require more items to estimate a score. Furthermore, inconsistent responding will also require the administration of more items to estimate a score. Regardless, CAT simulation data suggest that the CAT performs well for individuals with chorea scale scores between  $-0.4$  and 2.3 (i.e., less than 10 items are administered). More work using prospective data is needed to confirm this. While rates for race/ethnicity of this HD sample were consistent with established prevalence rates [40–43] and other large HD research cohorts [44–46], this sample was primarily Caucasian, and therefore, generalizability to other race/ethnic groups is uncertain. Furthermore, additional work is needed to establish test–retest reliability and responsiveness to change data for this measure. Future work is also needed to examine the relationship of this new measure with more general self-reported measures of motor functioning.

Regardless, this is the first HD-specific PRO measure of chorea, and it is the first time that CAT has been used to evaluate HRQOL (or any other construct) in HD. This new PRO is a potential candidate for inclusion in HD clinical trials that target treatment of chorea. This is especially important given that the only medications that are currently labeled for treatment of HD target the treatment of chorea.

# **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 was supported by the NIH, National Institute of Neurological Disorders and Stroke (R01NS040068), the NIH, Center for Inherited Disease Research (provided supported for sample phenotyping), and the CHDI Foundation (award to the University of Iowa). We thank the University of Iowa, the Investigators and Coordinators of this study, the study participants, the National Research Roster for Huntington Disease Patients and Families, the Huntington Study Group, and the Huntington's Disease Society of America. We acknowledge the assistance of Jeffrey D. Long, Hans J. Johnson, Jeremy H. Bockholt, Roland Zschiegner, and Jane S. Paulsen. We also acknowledge Roger Albin, Kelvin Chou, and Henry Paulsen for the assistance with participant recruitment. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

**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 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. Ross CA, Tabrizi SJ. Huntington's disease: From molecular pathogenesis to clinical treatment. Lancet Neurology. 2011; 10(1):83–98. [PubMed: 21163446]
- 2. Jankovic J, Roos RA. Chorea associated with Huntington's disease: To treat or not to treat? Movement Disorders. 2014; 29(11):1414–1418. [PubMed: 25156927]
- 3. Walker FO. Huntington's disease. Lancet. 2007; 369(9557):218–228. [PubMed: 17240289]
- 4. Grimbergen YA, Knol MJ, Bloem BR, Kremer BP, Roos RA, Munneke M. Falls and gait disturbances in Huntington's disease. Movement Disorders. 2008; 23(7):970–976. [PubMed: 18381643]
- 5. Wheelock VL, Tempkin T, Marder K, Nance M, Myers RH, Zhao H, et al. Predictors of nursing home placement in Huntington disease. Neurology. 2003; 60(6):998–1001. [PubMed: 12654967]
- 6. Wexler A. Stigma, history, and Huntington's disease. Lancet. 2010; 376(9734):18–19. [PubMed: 20626094]
- 7. Ho AK, Gilbert AS, Mason SL, Goodman AO, Barker RA. Health-related quality of life in Huntington's disease: Which factors matter most? Movement Disorders. 2009; 24(4):574–578. [PubMed: 19097181]
- 8. Read J, Jones R, Owen G, Leavitt BR, Coleman A, Roos RA, et al. Quality of life in Huntington's disease: A comparative study investigating the impact for those with pre-manifest and early manifest disease, and their partners. Journal of Huntingtons Disease. 2013; 2(2):159–175.
- 9. Ready RE, Mathews M, Leserman A, Paulsen JS. Patient and caregiver quality of life in Huntington's disease. Movement Disorders. 2008; 23(5):721–726. [PubMed: 18175350]
- 10. Carlozzi NE, Tulsky 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]
- 11. Huntington Study Group. Unified Huntington's Disease Rating Scale: Reliability and consistency. Movement Disorders. 1996; 11(2):136–142. [PubMed: 8684382]
- 12. Paulsen JS, Long JD, Ross CA, Harrington DL, Erwin CJ, Williams JK, et al. Prediction of manifest Huntington's disease with clinical and imaging measures: A prospective observational study. Lancet Neurol. 2014; 13(12):1193–1201. [PubMed: 25453459]
- 13. US Department of Health and Human Services: Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims: Draft guidance. Health and Quality of Life Outcomes. 2006; 4
- 14. Cella D, Nowinski C, Peterman A, Victorson D, Miller D, Lai J-S, Moy C. The neurology quality of life measurement (Neuro-QOL) initiative. Archives of Physical Medicine and Rehabilitation, Supplement. 2011; 92(Suppl 1):S28–S36.
- 15. Gershon RC, Lai JS, 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. 2011

- 16. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested in its first wave of adult selfreported health outcome item banks: 2005–2008. Journal of Clinical Epidemiology. 2010; 63:1179–1194. [PubMed: 20685078]
- 17. Cella DF, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH Roadmap cooperative group during its first 2 years. Medical Care. 2007; 45(5 Suppl 1):S3–S11. [PubMed: 17443116]
- 18. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH Roadmap cooperative group during its first 2 years. Medical Care. 2007; 45(5 Suppl 1):S3–S11. [PubMed: 17443116]
- 19. PROMIS<sup>®</sup>. Instrument Development and Psychometric Evaluation Scientific Standards. [http://](http://www.nihpromis.org/Documents/PROMIS_Standards_050212.pdf) [www.nihpromis.org/Documents/PROMIS\\_Standards\\_050212.pdf](http://www.nihpromis.org/Documents/PROMIS_Standards_050212.pdf)
- 20. Hanauer DA, Mei Q, Law J, Khanna R, Zheng K. Supporting information retrieval from electronic health records: A report of University of Michigan's 9-year experience in developing and using the electronic medical record search engine (EMERSE). Journal of Biomedical Informatics. 2015; 55:290–300. [PubMed: 25979153]
- 21. Paulsen JS, Hayden M, Stout JC, Langbehn DR, Ayl-ward E, Ross CA, et al. Preparing for preventive clinical trials—the Predict-HD study. Archives of Neurology. 2006; 63(6):883–890. [PubMed: 16769871]
- 22. Carlozzi NE, Schilling SG, J.-S. L. Paulsen JS, Hahn EA, Perlmutter JS, Ross CA, Downing NR, Kratz AL, McCormack MK, Nance MA, Quaid KA, Stout J, Gershon RC, Ready R, Miner JA, Barton SK, Perlman SL, Rao SM, Frank S, Shoulson I, Marin H, Geschwind MD, Dayalu P, Foroud T, Goodnight SM, Cella D. HDQLIFE: Development and assessment of health-related quality of life in Huntington disease (HD). Quality of Life Research. Under Review.
- 23. Shoulson I, Fahn S. Huntington disease—clinical care and evaluation. Neurology. 1979; 29(1):1–3. [PubMed: 154626]
- 24. Kline, RB. Principles and practice of structural equation modeling. 2nd ed.. Guilford Press; New York: 2005.
- 25. Bentler PM. Comparative fit indexes in structural models. Psychological Bulletin. 1990; 107(2): 238–246. [PubMed: 2320703]
- 26. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling-a Multidisciplinary Journal. 1999;  $6(1):1-55.$
- 27. Hatcher, L. A step-by-step approach to using SAS for factor analysis and structural equation modeling. SAS Institute Inc.; Cary, NC: 1994.
- 28. McDonald, RP. Test theory: A unified treatment. Lawrence Erlbaum Associates Inc.; Mahwah, NJ: 1999.
- 29. Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Quality of Life Research. 2007; 16(Suppl 1):19–31. [PubMed: 17479357]
- 30. Cook KF, Kallen MA, Amtmann D. Having a fit: Impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. Quality of Life Research. 2009; 18(4):447–460. [PubMed: 19294529]
- 31. Muthén, LK.; Muthén, B, O. Mplus user's guide. Muthén & Muthén; Los Angeles, CA: 2011.
- 32. Samejima, F.; van der Liden, WJ.; Hambleton, R. The graded response model.. In: van der Liden, WJ., editor. Handbook of modern item response theory. Springer; NewYork, NY: 1996. p. 85-100.
- 33. Crane PK, Gibbons LE, Jolley L, van Belle G. Differential item functioning analysis with ordinal logistic regression techniques. DIFdetect and difwithpar. Medical Care. 2006; 44(11 Suppl 3):S115–S123. [PubMed: 17060818]
- 34. 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. [PubMed: 21572908]
- 35. Cai, L.; Thissen, D.; du Toit, SHC. IRTPRO for windows [computer software]. Scientific Software International; Lincolnwood, IL: 2011.

- 36. Choi SW. Firestar: Computerized adaptive testing simulation program for polytomous item response theory models. Applied Psychological Measurement. 2009; 33(8):644–645.
- 37. Cohen, J. Statistical power analysis for the behavioral sciences. 2nd ed.. Academic Press; New York: 1988.
- 38. Carlozzi NE, Downing NR, McCormack MK, Schilling SG, Perlmutter JS, Hahn EA, Lai J-S, Frank S, Quaid KA, Paulsen JS, Cella D, Goodnight SM, Miner JA, Nance MA. New measures to capture end of life concerns in Huntington disease: Meaning and purpose and concern with death and dying from HDQLIFE (a patient reported outcomes measurement system). Quality of Life Research. Under Review.
- 39. Carlozzi NE, Schilling SG, Lai J-S, Perlmutter JS, Nance MA, Waljee JF, Miner JA, Barton SK, Good-night SM, Dayalu P. HDQLIFE: The development of two new computer adaptive tests, speech difficulties and swallowing difficulties. Quality of Life Research. Under Review.
- 40. Pringsheim T, Wiltshire K, Day L, Dykeman J, Steeves T, Jette N. The incidence and prevalence of Huntington's disease: A systematic review and meta-analysis. Movement Disorders. 2012; 27(9): 1083–1091. [PubMed: 22692795]
- 41. Folstein, SE. Huntington's disease: A disorder of families. Johns Hopkins University Press; Baltimore: 1989.
- 42. 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]
- 43. Narabayashi H. Huntington's chorea in Japan: Review of the literature. Advances in Neurology. 1973; 1:253–259.
- 44. Duff K, Paulsen J, Mills J, Beglinger LJ, Moser DJ, Smith MM, et al. Mild cognitive impairment in pre-diagnosed Huntington disease. Neurology. 2010; 75(6):500–507. [PubMed: 20610833]
- 45. Tabrizi SJ, Langbehn DR, Leavitt BR, Roos RA, Durr A, Craufurd D, 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]
- 46. Paulsen JS, Long JD, Johnson HJ, Aylward EH, Ross CA, Williams JK, 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]



# **Fig. 1.**

HDQLIFE Chorea test information plot. In general, we want total information to be >9.0 and standard error to be <0.33 (this provides a reliability of 0.9). This figure shows excellent total information and standard error for HDQLIFE Chorea scale scores between –0.6 and +2.7



#### **Fig. 2.**

HDQLIFE Chorea number of CAT items by CAT theta. This figure shows the number of CAT items used for different scale scores in standard deviation units: at –1 SD units, the CAT always used all 34 items in the item bank; from +0.3 to +2.0 SD units, the CAT always used the minimum number of 4 items in the item bank; and at 3 units, the CAT used all 34 items in the item bank

#### **Table 1**

# Demographic data for the HDQLIFE participants



Entries in the table represent percentage of participants unless otherwise specified

 $a<sub>T</sub>$ There were significant group differences for this variable

Author Manuscript

Author Manuscript

#### **Table 2**

# Exploratory factor analysis results of the HDQLIFE Chorea item pool







Bold values refers to item loadings that were >0.3

 $a<sup>2</sup>$  In the past 7 days

b<br>During the past 7 days

#### **Table 3**

# HDQLIFE Chorea item parameters





Items that are indicated in bold were selected for inclusion on the 6-item chorea short form

 $\frac{a}{\ln}$  the past 7 days

 $b$ During the past 7 days

Author Manuscript

**Author Manuscript** 

#### **Table 4**

# HDQLIFE Chorea short-form summed score to t score conversion table

