RESEARCH ARTICLE

HDQLIFE and Neuro-QoL Physical Function Measures: Responsiveness in Persons With Huntington's Disease

Noelle E. Carlozzi, PhD,[1*] Nicholas R. Boileau, MPH,[1] Kelvin L. Chou, MD,[2] Rebecca E. Ready, PhD,[3] David Cella, PhD,[4] Michael K. McCormack, PhD,[5,6] Jennifer A. Miner, MBA,[1] and Praveen Dayalu, MD[2]

[1]*Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, Michigan, USA*
[2]*Department of Neurology, University of Michigan, Ann Arbor, Michigan, USA*
[3]*Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, Massachusetts, USA*
[4]*Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA*
[5]*Department of Psychiatry, Rutgers University–Robert Wood Johnson Medical School, Piscataway, New Jersey, USA*
[6]*Department of Pathology, Rowan University-SOM (School of Medicine), Stratford, New Jersey, USA*

ABSTRACT: Background: Huntington's disease (HD) is a neurological disorder that causes severe motor symptoms that adversely impact health-related quality of life. Patient-reported physical function outcome measures in HD have shown cross-sectional evidence of validity, but responsiveness has not yet been assessed.

Objectives: This study evaluates the responsiveness of the Huntington Disease Health-Related Quality of Life (HDQLIFE) and the Quality of Life in Neurological Disorders (Neuro-QoL) physical function measures in persons with HD.

Methods: A total of 347 participants completed baseline and at least 1 follow-up (12-month and 24-month) measure (HDQLIFE Chorea, HDQLIFE Swallowing Difficulties, HDQLIFE Speech Difficulties, Neuro-QoL Upper Extremity Function, and/or Neuro-QoL Lower Extremity Function). Of the participants that completed the baseline assessment, 338 (90.9%) completed the 12-month assessment, and 293 (78.8%) completed the 24-month assessment. Standardized response means and general linear models evaluated whether the physical function measures were responsive to self-reported and clinician-rated change over time.

Results: Small to moderate effect sizes for the standardized response means supported 12-month and 24-month responsiveness of the HDQLIFE and Neuro-QoL measures for those with either self-reported or clinician-rated declines in function. General linear models supported 12-month and 24-month responsiveness for all HRQOL measures relative to self-reported declines in health, but generally only 24-month responsiveness was supported relative to clinician-rated declines in function.

Conclusions: Longitudinal analyses indicate that the HDQLIFE and the Neuro-QoL physical function measures are sensitive to change over time in individuals with HD. Thus, these scales exhibit evidence of responsiveness and may be useful outcome measures in future clinical trials. © 2019 International Parkinson and Movement Disorder Society

Key Words: health-related quality of life; Huntington's disease; patient-reported outcome (PRO); psychometric; validity

*__Correspondence to:__ Dr. Noelle E. Carlozzi, Department of Physical Medicine & Rehabilitation, University of Michigan, North Campus Research Complex, 2800 Plymouth Road, Building NCRC B14, Room G216Ann Arbor, MI 48109-2800; E-mail: carlozzi@med.umich.edu

Relevant conflicts of interests/financial disclosures: Carlozzi, N.E. is supported by research grants from the NIH, the Neilsen Foundation, and CHDI, as well as contracts from Goldfinch, LLC, and Health and Human Services – Centers for Medicare & Medicaid Services; she receives honoraria for her role on the CHDI scientific advisory board; she declares no conflicts of interest. Boileau, N.R. has been supported by research grants from the NIH, the Neilsen Foundation, and CHDI. Chou KL has received research support from the NIH (NS091856-01, NS10061102, NS107158), Parkinson Study Group (STEADY-PD III, SURE-PD3, NILO-PD), and Eli Lilly. He has been a consultant for Accordant, Boston Scientific, and Sunovion Pharmaceuticals. He also receives royalties from UpToDate and Springer Publishing. Ready, R.E. declares that she has no conflicts of interest. Cella, D. receives grant funding from the National Institutes of Health and reports that he has no conflicts of interest. McCormack, M.K. currently has grants from the NJ Department of Health; he declares no conflicts of interest. Miner, J.A. is supported by funding from the NIH, PCORI, the Craig H. Neilsen Foundation, and Goldfinch Bio; she declares no conflicts of interest. Dayalu, P. is supported by grants from Vaccinex and CHDI. He declares no conflicts of interest.

Huntington Disease Health-Related Quality of Life site investigators and coordinators are listed in the Appendix.

Received: 24 January 2019; **Revised:** 26 September 2019; **Accepted:** 4 October 2019

Published online 14 November 2019 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.27908

Huntington's disease (HD) is a progressive neurodegenerative disorder that causes profound cognitive, behavioral, and motor declines.[1-6] The motor disorder in HD is multifaceted and can include chorea, bradykinesia, rigidity, and dystonia; these symptoms adversely impact all body segments and limbs, with significant impacts on daily living and social participation.[7] In HD clinical trials, most motor outcomes are administered by clinicians (e.g., Unified Huntington Disease Rating Scale [UHDRS] Total Motor Score), but these measures correlate poorly with real-world function.[8] Patient-reported outcome (PRO) measures of motor symptoms and associated functional limitations are rarely used in HD, even though such outcomes are key measures of efficacy for new treatments.[9] In HD, where many therapies under development seek to slow the loss of function related to motor problems,[10-15] there is a need for meaningful and sensitive PRO measures that capture physical aspects of health-related quality of life (HRQOL).

The Huntington Disease Health-Related Quality of Life (HDQLIFE) measurement system[16-18] was designed to provide reliable and valid assessments of HRQOL in persons with HD. This system includes several HD-specific measures of HRQOL as well as generic HRQOL measures from the Quality of Life in Neurological Disorders (Neuro-QoL) system.[19,20] In a large cross-sectional study of persons affected with prodromal/premanifest, early-stage, and late-stage HD, the HDQLIFE physical function PROs demonstrated strong validity and reliability[21]; however, responsiveness over time has not yet been established. It is essential to determine the efficacy of experimental treatments, responsiveness, or the ability of a measure to detect meaningful change.[22] In the current study, we follow the same population for 2 years to determine whether these PROs can detect change over time. We hypothesized that the HDQLIFE and Neuro-QoL measures would be responsive to self-reported global changes and clinician-rated changes in health. Specifically, we hypothesized that: (1) 12-month and 24-month changes would be greater in magnitude for those with self-rated or clinician-rated declines in health relative to those reporting no change/improved health; (2) there would be small to moderate 12-month and 24-month effect sizes for participants with declines in health (based on either self-report or clinician ratings), negligible effect sizes for the group reporting no change, and negligible or small effect sizes for the group reporting improvements in physical health; and (3) there would be significant declines in PRO reports of physical HRQOL over time (for those individuals with self-reported declines or clinician-rated declines), whereas participants with no change or improvement in physical health would have no change or no change/small improvements, respectively.

Method

Study Participants

Data were collected through the HDQLIFE study, a longitudinal study examining HRQOL in persons affected with HD.[17] Participants were included in analyses if they completed at least 1 follow-up visit (n = 372 participants). A detailed description of the broader cohort study sample and recruitment methods are reported elsewhere.[17] Study eligibility included a positive gene test and/or a clinical diagnosis of HD as well as participant age of ≥18 years and ability to provide informed consent (cognitive status was confirmed using a standard assessment[23]).

Study Procedures

Participants completed assessments at baseline and at 12 and 24 months. Each visit involved an in-person assessment and several computer-based self-report surveys regarding HRQOL, which could be completed during the in-person visit or at home. A subsample of participants (n = 24) participated in a 1-day to 3-day retest of the self-report measures. All data were procured in accordance with the local institutional review boards, and the participants provided informed consent prior to participation in this study.

Measures

Clinician-Rated Assessments

All clinician-rated assessments were completed at each study visit. The UHDRS[24] was used to classify participants in this study. The final question on the UHDRS Total Motor Scale asks the rater to score a diagnostic confidence level (DCL) for the participant on a scale of 0 (normal) to 4 (symptoms unequivocal of HD with >99% certainty) to determine whether the participant has motor symptoms consistent with manifest HD. If the participant scored less than a 4, he or she was rated as having premanifest HD. The participants who opted to complete follow-up visits by telephone (~15% of the study sample) did not complete the Total Motor Scale at follow-up. The UHDRS Total Functional Capacity (TFC) scale was used to determine the stage for manifest HD participants. TFC scores range from 0 to 13, with higher scores indicating better function. Participants scoring between 7 and 13 on the TFC were classified as early HD, whereas those scoring between 0 and 6 rated were classified as late HD.[7,25] Baseline data were used to determine HD staging data for all analyses.

The TFC was also used to characterize clinician-rated longitudinal change. Baseline TFC scores were subtracted from 12-month and 24-month scores, respectively, to generate clinician-rated change scores in functioning for baseline to 12 months (M = −0.44; SD =1.49) and baseline to 24 months (M = −0.57; SD = 2.07). The distribution of

change scores were used to classify participants into 1 of the following 3 groups: those with declines in functioning (i.e., TFC scores that got worse at follow-up 1 SD greater than the sample; e.g., <−1.93 from baseline to 12 months and <−2.63 from baseline to 24 months), those with no change in functioning, or those with improvements in functioning (i.e., TFC scores that improved at follow-up 1 SD greater than the sample; e.g., >1.05 from baseline to 12 months and > 1.50 from baseline to 24 months). This resulted in n = 69 participants with declines, n = 238 with no change, and n = 23 with improvements in clinician ratings for baseline to 12 months as well as n = 45 participants with declines, n = 216 with no change, and n = 28 with improvements in clinician ratings for baseline to 24 months.

Self-Reported Assessments

The participants provided information about age, gender, marital status, race, and ethnicity. In addition, medical record data were used to confirm HD diagnosis and CAG repeat length for study participants.

Physical HRQOL was assessed using physical health measures from the HDQLIFE[16-18] and Neuro-QoL measurement systems. Specifically, we examined HDQLIFE Chorea[26] (which assesses the impact that chorea has on physical activity and participation), HDQLIFE Speech Difficulties[18] (which assesses how difficulty with speech, i.e., oral expression, language production, and articulation affects communication and well-being), HDQLIFE Swallowing Difficulties[18] (which assesses how problems with swallowing and choking impacts well-being and eating), Neuro-QoL Upper Extremity Function (which measures fine motor tasks and activities of daily living), and Neuro-QoL Lower Extremity Function (which measures mobility). The administration format for the HDQLIFE measures changed during the course of the study. Specifically, for the 259 participants (69.6% of the overall sample) completing all 3 study visits, 112 (43.2%) completed the full item pools for each HDQLIFE measure (64 items for Chorea, 27 items for Speech Difficulties, and 20 item for Swallowing Difficulties) and 147 (56.8%) completed the full item pools at baseline and 12 months, but switched to answering assessments as computer adaptive tests (CATs) plus short forms (SFs) at their 24-month visit. Of the 79 participants who only completed the baseline and 12-month assessments (but not the 24-month assessment), all completed the full item pools. Finally, of the 34 participants who only completed the baseline and 24-month assessments, 11 (32.4%) completed the full item pools at both assessments and 23 (67.6%) switched to the CATs plus SF administration at their 24-month visit. All Neuro-QoL measures were administered as a CAT plus SF at each time point. All HDQLIFE and Neuro-QoL scores are on a T-metric (M = 50, SD = 10); higher scores indicate more of the construct being assessed (i.e., high scores for HDQLIFE measures indicate worse physical HRQOL, whereas high scores on Neuro-QoL measures indicate better physical HRQOL). CAT scores were simulated for participants who completed the full item pools (rather than CAT administrations) using Firestar software.[27] The CAT and SF administration of these measures takes 1 minute or less to complete.[28]

At each follow-up visit (i.e., 12 and 24 months), participants rated 5 anchor items (1 for each of the 3 HDQLIFE and 2 Neuro-QoL measures) on a 5-point Likert scale, ranging from 1 (much worse) to 5 (much better), based on how they felt their condition was compared with the prior visit (Supplemental Table A). Each anchor item asked the participant about changes in either their chorea, speech, swallowing, ability to move their hands (i.e., upper extremities), or overall physical functioning (i.e., lower extremities). Based on their responses to each anchor item, participants were placed in 3 groups. Individuals who responded 1 (much worse) or 2 (worse) were included in the group with self-reported declines in health. Participants who responded 3 (same) were included in the group with no changes in health. Participants who responded 4 (better) or 5 (much better) were included in the group with improvements in health. To examine the change from baseline to 24 months (for which there was no specific anchor), we classified participants who reported poorer function during any of the visits in the poorer self-reported health group, those who did not report any changes between visits into the self-reported no change group, and participants who reported improved functioning during any of the visits were placed in the self-reported improvement group. In the case that a participant reported improvement at 1 follow-up visit and decline in another, the ratings were offset and the participant was placed in the no-change group.

Statistical Analysis

Descriptive Data

Statistical analyses were performed using SAS 9.4 software.[29] The data were normally distributed (according to Bulmer's criteria)[30] and therefore we used parametric tests to analyze the data. Group differences for demographic variables were examined using either chi-square (for categorical data; Fisher's exact tests when cells counts <5) or one-way analysis of variance (for continuous data). Descriptive data for each HRQOL measure were calculated for each HD group (premanifest, early HD, late HD). One-way analysis of variance with Bonferroni post-hoc analyses determined whether the 3 HD groups differed on the 5 HRQOL physical health measures at each of the 3 assessments (baseline and 12 and 24 months). We expected that the premanifest group would report better physical HRQOL than the early-HD group, who in turn

would report better physical HRQOL than the late-HD group.

Reliability and Measurement Error

Intraclass correlation coefficients (random 2-way consistency model) were calculated to examine 1- to 3-day test–retest reliability for the small subsample of participants; 95% confidence intervals (CI) were calculated for each reliability coefficient. Minimum acceptable criteria for test–retest reliability was set at ≥0.70 for intraclass correlations.[31]

In addition, minimal important difference (MID), detectable change (DC_{95}), and standard error of measurement (SEM) were calculated for the HDQLIFE and Neuro-QoL physical function measures. MID, or the smallest score changes that are perceived as being important,[32,33] were calculated using the means and standard deviations of participants who indicated that their physical function was either "a little worse" or "a little better" from baseline to 12 months (the absence of an anchor item to assess change from baseline to 24 months precluded the calculation of MIDs for this time frame). One-way analysis of variance with Bonferonni post-hoc comparisons were used to determine if group differences were significant. DC_{95} was calculated as a conservative estimate and identified reliable change scores (i.e., the amount of change due that can be detected with 95% confidence as not due to measurement error) from baseline to 12 months. DC_{95} was calculated according to the following formula:[34]

$$DC_{95} = SEM*1.96*\sqrt{2}.$$

Finally, SEMs were calculated to estimate the maximum difference between one's observed score and his or her true score for a given assessment.[35] SEM was calculated using the baseline data as follows:

$$SEM = SD*\sqrt{1-ICC}$$

where SD is the standard deviation of the sample and intraclass correlation coefficients is the test–retest reliability of the measure. SEM percentages (SEM divided by the mean of all observations across time points times 100) < 10% are indicative of good measurement error.[36]

Responsiveness

Guyatt's responsiveness statistic (RS) and standardized response mean (SRM) effect sizes were calculated to examine the responsiveness of the HRQOL measures. RS were calculated by dividing the mean change of each group by the standard deviation of change in the "no change" group.[37] RS were calculated relative to self-reported and clinician-rated changes in physical health. SRMs were calculated by dividing the average change from baseline to follow-up (12 and 24 months) by the standard deviation of the change.[38,39] SRMs were also calculated relative to self-reported and clinician-rated changes in physical health. For self-reported changes in health, we compared participants with self-reported declines in health with those with no change and those with improvements in health. For clinician-rated changes in health, we compared participants with clinician-rated declines in function relative to those with no change or those with improvements in clinician-rated function. Given findings in other clinical samples,[40] we hypothesized that RS and SRM effect sizes would be greater in magnitude for those who reported (or with clinician-rated) declines in health relative to those reporting no change/improved health. Effect sizes between 0.00 and |0.19| were considered "negligible," |0.20| to |0.49| were small, |0.50| to |0.79| were moderate, and ≥ |0.80| were large.[38] We hypothesized that participants who reported declines in health would have effect sizes ≤– 0.20 for positively worded concepts (i.e., higher scores indicate better HRQOL), or ≥ 0.20 for negatively worded concepts (i.e., when higher scores indicate worse HRQOL). For participants who reported no change, we predicted that they would have negligible SRMs (i.e., ≥|0.19|). For the group with improvements in either self-reported health or clinician rated function, we predicted that RS/SRMs would be either small (i.e., 0.20 to 0.49; we expected positive RS/SRMs for positively worded concepts and negative RS/SRMs for negatively worded concepts) or negligible (RS/SRMs ≥|0.19|); this hypothesis is based on previous literature that suggests globally reported improvements in HRQOL are smaller in magnitude that global ratings of declines.[40]

General linear models were used to examine change over time (from baseline to 12 months and baseline to 24 months) for each of the HDQLIFE and Neuro-QoL HRQOL measures relative to the respective self-reported anchor item or clinician-rated change. Each model included group status (i.e., declines in health, no change in health, or improvements) as a predictor of change in HRQOL. Least-square means and standard errors were calculated for each group to determine whether change over time significantly differed from zero. Responsiveness would be supported by significant declines in HRQOL relative to self-reported declines in health and clinician-rated declines in function.

We provide a summative table of the different analytical approaches examining responsiveness of the physical HRQOL measures. For each HRQOL measure, responsiveness will be supported if ≥75% of results are in accordance with the proposed hypotheses.[41]

Results

Study Attrition

At baseline, 152 participants had premanifest HD, 153 were early-stage, and 67 were late-stage HD. Of the

TABLE 1. Baseline Descriptive Data

Variable	Premanifest-HD, N = 152	Early-HD, N = 153	Late-HD, N = 67	Combined Sample, N = 372
Age, yr[a]				
M (SD)	43.0 (12.4)	53.0 (11.9)	55.0 (10.7)	49.3 (13.0)
Gender, %				
Female	64.5	53.6	56.7	58.6
Male	35.5	46.4	43.3	41.4
Race, %[a]				
White	98.0	95.4	92.5	95.8
African American	0.0	1.3	7.5	1.9
Other	1.3	3.3	0.0	1.9
Unknown	0.7	0.0	0.0	0.3
Ethnicity, %				
Not Hispanic or Latino	92.1	91.5	97.0	92.7
Hispanic or Latino	1.3	5.2	0.0	2.7
Not provided	6.6	3.3	3.0	4.6
Education, no. of yr[a]				
M (SD)	16.1 (2.8)	14.8 (2.8)	14.1 (2.5)	15.2 (2.8)
Marital status, %				
Single, never married	15.1	15.2	9.0	14.1
Married	69.1	57.6	65.7	63.8
Separated/divorced	13.2	19.9	22.4	17.6
Widowed	0.0	3.3	3.0	1.9
Living with partner	2.6	4.0	0.0	2.7
No. converted, baseline to 12 months	8	12	—	19
No. converted, baseline to 24 months	16	43	—	57
CAG repeats				
M (SD)	41.9 (2.6)	42.8 (3.9)	43.9 (6.3)	42.5 (3.7)

[a]Significant group differences: premanifest participants were on average 10 years younger than the early-HD group and 12 years younger than the late-HD group ($F_{2,344} = 37.63$; $P < 0.0001$); the premanifest group had approximately 1 more year of attainment in education ($F_{2,337} = 13.4$; $P < 0.0001$); the late-HD group had a significantly higher proportion of African Americans than the other 2 groups (Fisher's exact $P = 0.0048$). "No. converted" indicates the number of participants who moved to next stage from baseline to follow-up. Baseline to 24-month conversion also includes those who converted from 12 months to 24 months.
HD, Huntington's disease.

152 premanifest participants, 110 (72.4%) had a DCL of 0 to 1, whereas 42 (17.6%) had a DCL of 2 to 3 (prodromal). A total of 338 individuals (90.9%) completed the 12-month assessment, and 293 (78.8%) completed the 24-month assessment. Of those participants who were missing the 24-month visit (n = 78), 34 were lost to follow-up, 15 were unable to return or ineligible as a result of worsening symptoms, 12 withdrew consent or were unwilling to return, and 17 were lost because of "other" reasons (e.g., death, multiple reasons for termination, terminated by examiner). When compared with the participants who completed all 3 assessments, the participants who did not complete the 24-month assessment were more likely to have late HD ($\chi^2_2 = 23.4$; $P < 0.0001$). In addition, those who dropped out were more likely to be African American than other races (Fisher's exact $P = 0.0132$) or have a higher number of CAG repeats (odds ratio = 1.12; $P = 0.0005$).

Descriptive Data

Demographic characteristics for study participants are provided in Table 1. There were significant differences among the HD groups for each of the physical HRQOL measures at baseline and 12 and 24 months (Supplemental Table B). In all cases, group differences were in the hypothesized direction (i.e., the premanifest group reported better physical HRQOL than the early group, who in turn also reported better physical HRQOL than the late-HD group).

Reliability and Measurement Error

The 1-day to 3-day test–retest reliability for the HDQLIFE and Neuro-QoL measures was excellent. All intraclass correlation coefficients were > 0.85 (Supplemental Table B). MID and MCDs are presented in Table 2; MIDs generally ranged from 1 to 3 points depending on the measure and group that was being examined (the largest MIDs were seen for the groups with declines), whereas MDC_{95} values ranged from 4 to 11 points. All SEM values were < 10%, indicating that measurement error was within acceptable limits (Table 2).

Responsiveness

Table 3 displays 12-month and 24-month RS of physical HRQOL measures, relative to self-reported changes in physical health and clinician-rated changes in function. Effect sizes were consistent with the proposed hypotheses: they were greater in magnitude for those who self-reported declines in health relative to those with no changes or self-reported improvements in health and they were greater in magnitude for those with clinician-rated declines in

TABLE 2. The 12-Month Minimal Important Differences, Minimum Detectable Change, and Standard Error of Measurement

	MID				
	Decline in Self-Reported Health M (SD)	Improvement in Self-Reported Health M (SD)	$DC\%_{95}$ (LDC, UDC)	SEM	SEM %
Chorea CAT	1.03 (4.33)	−0.47 (5.63)	7.80 (−7.40, 8.20)	2.81	5.71
Chorea SF	0.89 (4.90)	0.18 (6.42)	7.54 (−7.04, 8.04)	2.72	5.49
Speech Difficulties CAT	1.36 (6.53)	−1.43 (7.99)	5.91 (−5.61, 6.21)	2.13	4.41
Speech Difficulties SF[a]	1.15 (6.19)	−3.36 (7.78)	7.28 (−7.18, 7.38)	2.62	5.42
Swallowing Difficulties CAT[b]	2.62 (4.46)	−1.00 (6.68)	4.13 (−3.63, 4.63)	1.49	2.99
Swallowing Difficulties SF[b]	2.76 (4.52)	1.08 (10.03)	5.77 (−4.77, 6.77)	2.08	4.18
Upper Extremities CAT[c]	−2.56 (5.96)	−0.41 (6.45)	5.65 (−6.75, 4.55)	2.04	4.59
Upper Extremities SF[b,c]	−2.87 (6.89)	−1.17 (4.85)	10.58 (−12.08, 9.08)	3.82	8.62
Lower Extremities CAT[c]	−2.12 (5.60)	−0.46 (5.92)	7.55 (−8.25, 6.85)	2.73	5.62
Lower Extremities SF[c]	−2.31 (5.72)	−0.25 (7.16)	6.31 (−7.61, 5.01)	2.28	4.65

LDC and UDC calculated in reference to 12-month change in Huntington Disease Health-Related Quality of Life.
[a]Health decline group differs from no change group.
[b]Health decline group differs from improvement group.
[c]Higher scores indicate better Huntington Disease Health-Related Quality of Life.
MID, minimal important difference; M, mean; SD, standard deviation; DC_{95}, detectable change (95% confidence); LDC, lower detectable change (95% confidence); UDC, upper detectable change (95% confidence); SF, Short Form; SEM, standard error of measurement; CAT, Computer Adaptive Test.

TABLE 3. Guyatt's Responsiveness Statistics for Changes in Huntington Disease Health-Related Quality of Life

	Self-Reported Changes in Physical Health (From Anchor Items)											
	Baseline to 12 Months						Baseline to 24 Months					
	No Change in Health		Declines in Health		Improvement in Health		No Change in Health		Declines in Health		Improvement in Health	
	N	SRM	N	SRM	N	SRM	N	SRM	N	SRM	N	SRM
Chorea CAT	223	**0.10**	74	**0.26**	24	**−0.30**	139	0.20	78	**0.30**	22	**0.15**
Chorea SF	220	**0.06**	73	**0.20**	24	**−0.19**	137	0.21	77	**0.44**	22	0.27
Speech Difficulties CAT	225	**0.05**	76	**0.21**	21	**−0.24**	142	**0.14**	83	**0.28**	18	**−0.02**
Speech Difficulties SF	224	**0.00**	75	0.18	21	**−0.35**	144	**0.12**	81	**0.26**	18	**−0.15**
Swallowing Difficulties CAT	239	**0.03**	64	**0.58**	21	**−0.22**	145	**0.09**	78	**0.61**	20	**0.10**
Swallowing Difficulties SF	241	**0.08**	61	**0.64**	21	**0.01**	145	**0.19**	76	**0.71**	20	0.40
Upper Extremities CAT[a]	210	**−0.11**	80	**−0.46**	33	**0.15**	127	−0.36	90	**−0.47**	28	**−0.12**
Upper Extremities SF[a]	211	**−0.12**	80	**−0.44**	33	**−0.02**	128	**−0.17**	90	**−0.43**	28	**−0.06**
Lower Extremities CAT[a]	197	**−0.14**	63	**−0.34**	63	**0.05**	125	−0.29	66	**−0.58**	56	**−0.06**
Lower Extremities SF[a]	197	**−0.13**	63	**−0.40**	63	**-0.06**	125	−0.26	66	**−0.46**	56	**0.08**
	Clinician-Rated Changes in Function (From TFC)											
	Baseline to 12 Months						Baseline to 24 Months					
	No Change in Health		Declines in Health		Improvement in Health		No Change in Health		Declines in Health		Improvement in Health	
	N	SRM	N	SRM	N	SRM	N	SRM	N	SRM	N	SRM
Chorea CAT	238	**0.06**	68	**0.22**	23	0.20	216	0.24	45	**0.31**	28	**−0.28**
Chorea SF	234	**0.04**	68	0.10	23	**0.07**	214	0.30	44	**0.46**	27	**0.00**
Speech Difficulties CAT	236	**0.01**	68	0.17	23	**−0.06**	216	**0.16**	44	**0.32**	28	**0.01**
Speech Difficulties SF	235	**−0.03**	68	0.14	23	**−0.09**	216	**0.12**	45	**0.37**	27	**−0.11**
Swallowing Difficulties CAT	236	**0.12**	68	0.11	23	**−0.25**	216	**0.17**	44	**0.53**	28	0.33
Swallowing Difficulties SF	236	**0.15**	69	**0.30**	22	**−0.15**	215	0.33	45	**0.54**	26	0.42
Upper Extremities CAT[a]	233	**−0.05**	66	**−0.51**	23	**−0.04**	215	-0.27	41	**−0.75**	26	−0.49
Upper Extremities SF[a]	235	**−0.15**	66	**−0.41**	23	**0.24**	217	-0.25	42	**−0.42**	26	**0.11**
Lower Extremities CAT[a]	233	**−0.07**	66	**−0.39**	23	**0.07**	215	-0.24	41	**−0.53**	26	−0.21
Lower Extremities SF[a]	234	**−0.10**	66	**−0.44**	23	**−0.01**	217	-0.28	42	**−0.48**	26	**−0.07**

Bold indicates effects size magnitudes that are consistent with the proposed hypotheses.
[a]Higher scores indicate better Huntington Disease Health-Related Quality of Life.
SRM, standardized response mean; TFC, Unified Huntington Disease Rating Scale Total Functional Capacity; CAT, Computer Adaptive Test; SF, Short Form.

TABLE 4. Responsiveness Relative to Self-Reported Changes in Health and Clinician-Rated Changes in Function

	12-Month Responsiveness, Least Squared Mean (SE)			24-Month Responsiveness, Least Squared Mean (SE)		
	No Change in Function	Declines in Function	Improvement in Function	No Change in Function	Declines in Function	Improvement in Function
Self-reported changes in physical health (from anchor items)						
Chorea CAT	0.50 (0.33)	1.12 (0.57)	−1.69 (1.01)	1.24 (0.51)[a]	1.66 (0.68)[a]	0.95 (1.28)
Chorea SF	0.29 (0.35)	0.95 (0.61)	−1.16 (1.06)	1.07 (0.46)[a]	2.53 (0.61)[a]	1.34 (1.14)
Speech Difficulties CAT	0.25 (0.40)	1.36 (0.68)[a]	−2.05 (1.30)	0.94 (0.56)	1.78 (0.73)[a]	−0.20 (1.57)
Speech Difficulties SF	0.02 (0.38)	1.12 (0.65)	−2.95 (1.23)*	0.79 (0.53)	1.61 (0.71)[a]	−1.02 (1.50)
Swallowing Difficulties CAT	0.17 (0.39)	2.81 (0.76)[a]	−1.81 (1.32)	0.59 (0.54)	3.55 (0.74)[a]	0.70 (1.46)
Swallowing Difficulties SF	0.49 (0.39)	3.19 (0.77)[a]	0.14 (1.31)	1.14 (0.50)[a]	4.24 (0.69)[a]	2.44 (1.34)
Upper Extremities CAT[b]	−0.65 (0.43)	−2.76 (0.69)[a]	1.20 (1.08)	−1.82 (0.51)[a]	−3.01 (0.61)[a]	−0.80 (1.09)
Upper Extremities SF[b]	−0.71 (0.43)	−3.03 (0.70)[a]	−0.12 (1.08)	−1.03 (0.57)	−3.05 (0.68)[a]	-0.39 (1.22)
Lower Extremities CAT[b]	−0.77 (0.42)	−1.94 (0.74)[a]	0.37 (0.74)	−1.85 (0.58)[a]	−3.72 (0.79)[a]	-0.40 (0.86)
Lower Extremities SF[b]	−0.69 (0.41)	−2.24 (0.72)[a]	-0.41 (0.72)	−1.43 (0.53)[a]	−3.25 (0.73)[a]	0.45 (0.79)
Clinician-rated changes in function (from TFC)						
Chorea CAT	0.32 (0.32)	1.13 (0.59)	0.66 (1.02)	1.34 (0.40)[a]	2.24 (0.87)[a]	−1.44 (1.11)
Chorea SF	0.21 (0.34)	0.63 (0.63)	0.33 (1.08)	1.50 (0.37)[a]	3.23 (0.81)[a]	−0.01 (1.03)
Speech Difficulties CAT	0.08 (0.39)	1.31 (0.72)	−0.38 (1.25)	1.04 (0.44)[a]	2.54 (0.98)[a]	0.05 (1.23)
Speech Difficulties SF	−0.14 (0.37)	1.03 (0.68)	−0.54 (1.18)	0.73 (0.42)	2.71 (0.92)[a]	−0.50 (1.19)
Swallowing Difficulties CAT	0.66 (0.40)	0.87 (0.75)	−1.36 (1.28)	0.97 (0.44)[a]	4.66 (0.97)[a]	2.33 (1.22)
Swallowing Difficulties SF	0.77 (0.39)	2.43 (0.73)[a]	-0.95 (1.28)	1.74 (0.41)[a]	4.58 (0.89)[a]	2.72 (1.17)[a]
Upper Extremities CAT[b]	−0.28 (0.40)	−3.71 (0.76)[a]	−0.19 (1.28)	−1.45 (0.38)[a]	−5.19 (0.88)[a]	−2.77 (1.10)[a]
Upper Extremities SF[b]	−0.86 (0.40)[a]	−3.19 (0.76)[a]	1.29 (1.29)	−1.56 (0.44)[a]	−3.43 (1.00)[a]	0.56 (1.27)
Lower Extremities CAT[b]	−0.41 (0.38)	−2.62 (0.72)[a]	0.35 (1.22)	−1.41 (0.43)[a]	−4.33 (0.98)[a]	−1.41 (1.24)
Lower Extremities SF[b]	−0.53 (0.37)	−2.87 (0.70)[a]	-0.03 (1.18)	−1.44 (0.40)[a]	−3.67 (0.92)[a]	−0.55 (1.16)

Higher scores indicate worse Huntington Disease Health-Related Quality of Life unless indicated.
[a]Denotes that change significantly differs from 0 ($P < 0.05$).
[b]Indicates that measure is reverse scored (i.e., higher scores indicate better Huntington Disease Health-Related Quality of Life).
SE, standard error;; CAT, Computer Adaptive Test; SF, Short Form; TFC, Unified Huntington Disease Rating Scale Total Functional Capacity.

function relative to those with no changes or improvements in clinician ratings of function. For self-reported change, effect sizes were generally negligible for the group reporting no change regardless of time frame (as hypothesized). For those with self-reported improvement, effect sizes were also generally negligible or small (in the hypothesized direction) for both the 12-month and 24-month time frames (as hypothesized). For those with self-reported declines in physical health, effect sizes were generally small regardless of time frame (again as hypothesized). A similar pattern of results was seen for self-reported change using SRM (see Supplemental Table C).

For clinician-rated change, there were generally negligible effect sizes for the group with no change and the group with improvement for the baseline to 12-month time frame (as hypothesized), and there were generally small effect sizes for the group with clinician-rated declines (as hypothesized). From baseline to 24 months, there were typically negligible to small effect sizes for the groups with no change or improvement, and small effect sizes for those with declines. Again, a similar pattern of results was seen for the clinician-rated change using SRM (see Supplemental Table C).

Findings for the general linear models are included in Table 4. Twelve- and 24-month responsiveness was supported for all HRQOL measures relative to self-reported declines in health. In addition, although 24-month responsiveness for all of the physical HRQOL measures was supported relative to clinician-rated declines in function, 12-month responsiveness was generally not supported relative to clinician-rated declines (notable exceptions included supported for 12-month responsiveness for Swallowing Difficulties [SF only], Upper Extremities [CAT and SF], and Lower Extremities [CAT and SF]). As hypothesized, participants with no change or improvements in physical health (based on both self-reported and clinician ratings) did not have significant 12-month changes on the HRQOL measures. Furthermore, the group with improvements (based on both self-reported and clinician ratings) also generally did not have significant 24-month change on the HRQOL measures (as hypothesized). Finally, with regard to 24-month responsiveness, and not as expected, the no-change group (based on both self-reported and clinician ratings) generally had changes in HRQOL score, albeit these changes were small.

Discussion

This investigation provides evidence that the HDQLIFE and Neuro-QoL Physical Function measures demonstrate responsiveness to health status changes for persons with HD. The majority of the HDQLIFE and Neuro-QoL physical HRQOL measures met our a priori criterion for

responsiveness (see the Supplemental Table D). First, 12-month and 24-month effect sizes for patient-reported physical health were larger for individuals with worsening health relative to participants with no change or improvements in physical health; this was true for self-reported decline and clinician-rated decline. Our results are consistent with the responsiveness of the Neuro-QoL measures in other neurological conditions, including Parkinson's disease and adult epilepsy.[42-44]

In addition, although the group with self-reported and clinician-rated declines exhibited small declines in physical HRQOL at 12 months and 24 months—which is typical for PRO measures—there were a few measures that were especially responsive to declines in physical HRQOL. Relative to self-reported declines in health, Swallowing Difficulties and Upper Extremities exhibited moderate effect sizes for both 12-month and 24-month change over time. Relative to clinician-rated functional declines, Swallowing Difficulties and Upper Extremity Function were again noteworthy with regard to 24-month change over time.

As expected, there were negligible or small changes over time on HRQOL scores for participants with self-reported and clinician-reported improvements in health. These findings are consistent with previous literature in other populations that find that global improvements in HRQOL—using PROs—tend to be significantly smaller in magnitude than those global ratings of decline.[40] In addition, and also as expected, 12-month effect sizes were negligible for individuals with no self-reported or clinical-reported changes in health. Yet, contrary to the expectation of negligible change, there were generally small 24-month effect size declines for the group that did not change based on both self-report and clinician ratings. Thus, participants may experience small declines in HRQOL that either the patients and/or clinicians either are unable to detect and/or that do not affect function.

Some measures did not perform as well as expected. Although 24-month responsiveness for all of the physical HRQOL measures was supported relative to clinician-rated declines in function, support for 12-month responsiveness relative to clinician-rated declines was less robust. We believe that the large variability in TFC-related change (i.e., large standard errors) was likely responsible for the absence of significant 12-month group differences. Specifically, individuals with larger TFC-related declines are also more likely to have advanced disease, and the reliability of PRO scores are adversely affected by cognitive impairment.[45]

Our responsiveness data were more robust when examining self-reported HRQOL relative to self-reported change and less robust when examining self-reported HRQOL relative to clinician-rated changes. Findings from other populations suggest that clinicians tend to systematically underestimate patient symptoms and functional decline.[46-54] 54 This is not an issue when both the patient and clinician are providing similar reports for functional abilities and symptoms, but in cases where clinicians and patients are discordant, the path forward may be less clear. It may be that this discrepancy is simply a disagreement between the patient and the clinician about the relative importance of an aspect of function, but HD patients are also less able to communicate their problems as the disease progresses. Importantly, PROs and clinician ratings provide different information; the overall concordance (or lack thereof) between these reports may provide clinically meaningful information. Given this, and the fact that PROs are not typically included in HD clinical care, clinicians may miss important patient-centered information by relying primarily on clinician ratings of disease progression. Including PROs can help clinicians better care for patients by providing a more complete picture. MIDs and MCDs provided in Table 2 can be used to help guide clinical interpretation of change scores on these PROs.

Our study has some limitations. With regard to the study procedures, a small portion of the sample completed follow-up assessments via telephone, and thus clinician ratings for several measures were missing for these participants. Although there are data to support the equivalence of the different administration formats,[55-57] it is also important to acknowledge that the administration format for the HDQLIFE measures was not consistent across study visits (as detailed previously). Also, the participants with more advanced disease were more likely to be lost to follow-up, and differential attrition may have made it more difficult to identify significant declines in physical HRQOL. This is not an uncommon problem in HD, and most studies (including clinical trials) target individuals that are either premanifest or early in the disease process.[58-60] Potential strategies to help mitigate loss to follow-up can include engaging advocacy groups, participation in HD community events, and engagement of social media.[61,62] Given that individuals with more severe disease were more likely to discontinue participation as a result of worsening symptoms (and in many cases the inability to provide informed consent at a later study visit) and that these individuals are precisely the individuals who are experiencing the largest declines in health, we would expect this to negatively impact our findings such that effect sizes are smaller than would be expected for the HD population. This would be consistent with findings in this sample that suggest that the reliability of these PROs can be compromised in those in the later stages of the disease process.[63] Our results are based on a well-educated, primarily non-Hispanic White population, limiting possible generalizability to other racial/ethnic groups. In addition, self-reported changes were also closely matched to each physical HRQOL domain, whereas clinician-rated changes reflected functional changes—which are less closely tied to HRQOL constructs—precluding our ability to directly compare the clinical utility of these measures

across these 2 different types of anchors. In addition, we did not collect data about therapeutic treatments for study participants, and thus the impact of current therapeutic treatments is currently unknown. Furthermore, given that the baseline data informed the development of the HDQLIFE CAT and SF administrations, the baseline and 12-month study visits included significantly more items than the 24-month study visit (when most participants were able to complete the shorter administration formats of these measures).

Finally, changes over time were modest, as demonstrated by the small MIDs (Table 2). As often happens, MIDs were generally smaller than the associated DC_{95} for each measure. The use of DC_{95} to classify people as changed is a useful conservative approach when one wishes to be 95% confident that change has occurred. However, if the goal is to optimize accurate classification, then a value closer to the MID is likely to be more accurate.

Taken together, the findings support the responsiveness of the HDQLIFE and Neuro-QoL physical HRQOL measures to change over time. We observed small to moderate effect sizes in the absence of an intervention designed to improve HRQOL, so these analyses likely provide a conservative estimate of psychometric performance. As such, data would support using these new measures of HRQOL in both observational and experiemental research study designs in persons with HD, in conjunction with the more commonly administered clinician-rated measures of physical functioning. Findings also highlight the importance of complementing clinician ratings with patient reports of HRQOL to better understand the impact that an intervention has on an individual affected with HD. In this manner, a more patient-centered approach to HD treatment should include an assessment of HRQOL. As such, HRQOL physical function measures have the potential to provide clinically relevant information that should be considered in the context of the standard clinical exam (which does not typically include PROs) in persons with HD. ■

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

Appendix

Huntington Disease Health-Related Quality of Life 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).

References

1. Duff K, Paulsen J, Mills J, et al. Mild cognitive impairment in prediagnosed Huntington disease. Neurology 2010;75(6):500–507.
2. Snowden J, Craufurd D, Griffiths H, Thompson J, Neary D. Longitudinal evaluation of cognitive disorder in Huntington's disease. J Int Neuropsychol Soc 2001;7(1):33–44.
3. Paulsen JS, Ready RE, Hamilton JM, Mega MS, Cummings JL. Neuropsychiatric aspects of Huntington's disease. J Neurol Neurosurg Psychiatry 2001;71(3):310–314.
4. Duff K, Paulsen JS, Beglinger LJ, Langbehn DR, Stout JC, Predict-HD Investigators of the Huntington Study Group. Psychiatric symptoms in Huntington's disease before diagnosis: the Predict-HD study. Biol Psychiatry 2007;62(12):1341–1346.
5. Long JD, Paulsen JS, Marder K, et al. Tracking motor impairments in the progression of Huntington's disease. Mov Disord 2014;29(3):311–319.
6. Maroof DA, Gross AL, Brandt J. Modeling longitudinal change in motor and cognitive processing speed in presymptomatic Huntington's disease. J Clin Exp Neuropsychology 2011;33(8):901–909.
7. Marder K, Zhao H, Myers RH, et al. Rate of functional decline in Huntington's disease. Neurology 2000;54(2):452.
8. Paulsen JS, Wang C, Duff K, et al. Challenges assessing clinical endpoints in early Huntington disease. Mov Disord 2010;25(15):2595–2603.
9. Basch E. The missing voice of patients in drug-safety reporting. N Engl J Med 2010;362(10):865–869.
10. Busse M, Quinn L, Dawes H, et al. Supporting physical activity engagement in people with Huntington's disease (ENGAGE-HD): study protocol for a randomized controlled feasibility trial. Trials 2014;15:487.
11. Busse M, Quinn L, Debono K, et al. A randomized feasibility study of a 12-week community-based exercise program for people with Huntington's disease. J Neurol Phys Ther 2013;37(4):149–158.
12. Dawes H, Collett J, Debono K, et al. Exercise testing and training in people with Huntington's disease. Clin Rehabil 2015;29(2):196–206.

13. Khalil H, Quinn L, van Deursen R, et al. What effect does a structured home-based exercise programme have on people with Huntington's disease? A randomized, controlled pilot study. Clinical Rehabil 2013;27(7):646–658.
14. Quinn L, Busse M, Khalil H, Richardson S, Rosser A, Morris H. Client and therapist views on exercise programmes for early-mid stage Parkinson's disease and Huntington's disease. Disabil Rehabil 2010;32(11):917–928.
15. Quinn L, Hamana K, Kelson M, et al. A randomized, controlled trial of a multi-modal exercise intervention in Huntington's disease. Parkinsonism Relat Disord 2016;31:46–52.
16. Carlozzi NE, Downing NR, Schilling SG, et al. The development of a new computer adaptive test to evaluate chorea in Huntington disease: HDQLIFE Chorea. Qual Life Res 2016;25(10):2429–2439.
17. Carlozzi NE, Schilling SG, Lai JS, et al. HDQLIFE: Development and assessment of health-related quality of life in Huntington disease (HD). Qual Life Res 2016;25(10):2441–2455.
18. Carlozzi NE, Schilling SG, Lai JS, 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; 25(10):2417–2427.
19. Cella D, Nowinski C, Peterman A, et al. The Neurology Quality of Life Measurement (Neuro-QOL) Initiative. Arch Phys Med Rehabil 2011;92(suppl 1):S28–S36.
20. Gershon RC, Lai JS, Bode R, et al. Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon clinical and general population testing. Qual Life Res 2011;21(3):475–86.
21. Carlozzi NE, Ready R, Frank S, et al. Patient-reported outcomes in Huntington disease: Neuro-QOL and HDQLIFE physical function measures. Mov Disord 2017;32(7):1096–1102.
22. Pickard AS, Johnson JA, Feeny DH. Responsiveness of generic health-related quality of life measures in stroke. Qual Life Res 2005; 14(1):207–219.
23. Jackson W, Novack TA, Dowler RN. Effective serial measurement of cognitive orientation in rehabilitation: the Orientation Log. Arch Phys Med Rehabil 1998;79(6):718–720.
24. Huntington Study Group. Unified Huntington's Disease Rating Scale: reliability and consistency. Mov Disord 1996;11(2):136–142.
25. Shoulson I. Huntington disease: functional capacities in patients treated with neuroleptic and antidepressant drugs. Neurology 1981; 31:1333–1335.
26. Carlozzi NE, Downing NR, Schilling SG, et al. The development of a new computer adaptive test to evaluate chorea in Huntington disease: HDQLIFE chorea. Qual Life Res 2016;25(10):2429–2439.
27. Choi SW, Podrabsky T, McKinney N, Firestar D. Computerized adaptive testing simulation program for dichotomous item response theory models. Appl Psych Meas 2012;36(1):67–68.
28. Chang KW, Austin A, Yeaman J, et al. Health-related quality of life components in children with neonatal brachial plexus palsy: a qualitative study. PM R 2017;9(4):383–391.
29. *SAS 9.4 Language Reference Concepts* [computer program]. Cary, NC: SAS Institute; 2013.
30. Bulmer MG. Principles of Statistics. New York: Dover Publications; 1979.
31. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press; 1969.
32. Jaescke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407–415.
33. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharmacoeconomics Outcomes Res 2011;11(2):171–184.
34. Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome measures in rheumatology. Minimal clinically important difference. J Rheumatol 2001;28(2):400–405.
35. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52(9):861–873.
36. Flansbjer U, Holmback AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med 2005;37:75–82.
37. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40(2):171–178.
38. Cohen J. A power primer. Psychol Bull 1992;112(1):155–159.
39. Kopjar B. The SF-36 health survey: a valid measure of changes in health status after injury. Inj Prev 1996;2:135–139.
40. Dineen K, Cella D, Hahn EA, Yount S. Meaningful change in cancer-specific quality-of-life scores: differences between improvement and worsening. Clin Ther 2002;24:41–42.
41. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60(1):34–42.
42. Nowinski CJ, Siderowf A, Simuni T, Wortman C, Moy C, Cella D. Neuro-QoL Health-Related Quality of Life Measurement System: validation in Parkinson's disease. Mov Disord 2016;31(5):725–733.
43. Salsman JM, Victorson D, Choi SW, et al. Development and validation of the Positive Affect and Well-Being Scale for the Neurology Quality of Life (Neuro-QOL) measurement system. Qual Life Res 2013;22(9):2569–2580.
44. Victorson D, Cavazos JE, Holmes GL, et al. Validity of the Neurology Quality-of-Life (Neuro-QoL) measurement system in adult epilepsy. Epilepsy Behav 2014;31:77–84.
45. Carlozzi NE, Paulsen JS, Stout J, et al. Understanding patient-reported outcome measures in Huntington disease: at what point is cognitive impairment related to poor measurement reliability? Neurotherapeutics 2017;14(1):239–240.
46. Rodriguez LV, Blander DS, Dorey F, Raz S, Zimmern P. Discrepancy in patient and physician perception of patient's quality of life related to urinary symptoms. Urology 2003;62(1):49–53.
47. Justice AC, Chang CH, Rabeneck L, Zackin R. Clinical importance of provider-reported HIV symptoms compared with patient-report. Med Care 2001;39(4):397–408.
48. Kalarchian MA, Wilson GT, Brolin RE, Bradley L. Assessment of eating disorders in bariatric surgery candidates: self-report questionnaire versus interview. Int J Eat Disorder 2000;28(4):465–469.
49. Lindstrom E, Lewander T, Malm U, Malt UF, Lublin H, Ahlfors UG. Patient-rated versus clinician-rated side effects of drug treatment in schizophrenia—clinical validation of a self-rating version of the UKU Side Effect Rating Scale (UKU-SERS-Pat). Nord J Psychiat 2001;55:5.
50. Luoma ML, Hakamies-Blomqvist L, Sjostrom J, et al. Physical performance, toxicity, and quality of life as assessed by the physician and the patient. Acta Oncol 2002;41(1):44–49.
51. Neville C, Clarke AE, Joseph L, Belisle P, Ferland D, Fortin PR. Learning from discordance in patient and physician global assessments of systemic lupus erythematosus disease activity. J Rheum 2000;27(3):675–679.
52. Neville C, Joseph L, Clarke AE, Belisle P, Ferland D, Fortin PR. Learning from discordance in patient and physician global assessments of lupus disease activity. Arthritis Rheum 1997;40(9):845–845.
53. Sneeuw KCA, Aaronson NK, Sprangers MAG, Detmar SB, Wever LDV, Schornagel JH. Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. J Clin Epidemiol 1998;51(7):617–631.
54. Sneeuw KCA, Aaronson NK, Sprangers MAG, Detmar SB, Wever LDV, Schornagel JH. Evaluating the quality of life of cancer patients: assessments by patients, significant others, physicians and nurses. Brit J Cancer 1999;81(1):87–94.
55. Bjorner JB, Chang CH, Thissen D, Reeve BB. Developing tailored instruments: item banking and computerized adaptive assessment. Qual Life Res 2007;16(suppl 1):95–108.
56. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). J Clin Epidemiol 2008;61(1):17–33.

57. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE Jr. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. J Clin Epidemiol 2014;67(5):516–526.

58. Frost C, Mulick A, Scahill RI, et al. Design optimization for clinical trials in early-stage manifest Huntington's disease. Mov Disord 2017;32(11):1610–1619.

59. Paulsen JS, Lourens S, Kieburtz K, Zhang Y. Sample enrichment for clinical trials to show delay of onset in Huntington disease. Mov Disord 2019;34(2):274–280.

60. Paulsen JS, Long JD, Johnson HJ, et al. Clinical and biomarker changes in premanifest Huntington disease show trial feasibility: a decade of the PREDICT-HD Study. Front Aging Neurosci 2014;6:78.

61. Goodman L, Sia C, Carnes R, et al. Advocacy recruiting for Huntington's disease clinical trials. PLoS Curr 2011;3:RRN1230.

62. Kinel S, McCarthy L, Kayson E, et al. Project AWARE: assessing awareness, willingness, and ability of the Huntington disease community to participate in Huntington disease clinical trials. Paper presented at: Huntington Study Group 2011; Indianapolis, IN, 5 November, 2011. https://link.springer.com/content/pdf/10.1007/s13311-011-0083-8.pdf.

63. Carlozzi NE, Schilling S, Kratz AL, Paulsen JS, Frank S, Stout JC. Understanding patient-reported outcome measures in Huntington disease: at what point is cognitive impairment related to poor measurement reliability? Qual Life Res 2018;27(10):2541–2555.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.