HHS Public Access

Author manuscript

J Neurol. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

J Neurol. 2018 June ; 265(6): 1443–1453. doi:10.1007/s00415-018-8852-5.

Agreement between clinician-rated versus patient-reported outcomes in Huntington disease

Noelle E. Carlozzi, PhD[1], **Nicholas R. Boileau, MPH**[1], **Joel S. Perlmutter, MD**[2], **Kelvin L. Chou, MD**[3], **Julie C. Stout, PhD**[4], **Jane S. Paulsen, PhD**[5,6,7], **Michael K. McCormack, PhD**[8], **David Cella, PhD**[9], **Martha A. Nance, MD**[10,11], **Jin-Shei Lai, PhD, OTR/L**[9], and **Praveen Dayalu, MD**[3]

[1]Department of Physical Medicine and Rehabilitation, University of Michigan

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

Disclosure of Conflicts of Interest

Carlozzi, N.E. currently has research grants from the NIH; she is also supported by grant funding from the NIH and CHDI. She provides patient reported outcome measurement selection and application consultation for Teva Pharmaceuticals. She declares no conflicts of interest.

Boileau, N. R. is supported by grant funding from the NIH; he declares no conflicts of interest.

Perlmutter, J.S. currently has research funding from the NIH, HDSA, CHDI, Michael J Fox Foundation, Barnes Jewish Hospital Foundation, the Lauren & Lee Fixel Family Foundation, and the APDA. He has received honoraria from the University of Rochester, American Academy of Neurology, Movement Disorders Society, Toronto Western Hospital, Alberta Innovates, Parkinson Disease Foundation, Columbia University, St. Louis University, Harvard University, University of Michigan, Huntington Study Group, Stanford University, University of Florida at Gainesville, and World Parkinson Congress. He declares no conflicts of interest.

Chou, K. currently has funding from the NIH and Cavion, receives royalties from UpToDate, Springer Publishing and Demos Health and serves as a consultant for Accordant and Sunovion Pharmaceuticals. He declares no conflicts of interest.

Stout, J.C. has received research funding in the past three years from the Australian National Health and Medical Research Council, University College London, the CHDI Foundation, Prana Biotechnology, and the University of California, Davis. She is a Director of Stout Neuropsych Pty Ltd, which has received funding from Omeros, Teva Pharmaceuticals, Vaccinex, and Isis. She has been a consultant to Prana Biotechnology and Roche. She receives compensation as a member of the Board of the Huntington's Study Group. She declares no conflicts of interest.

Paulsen, J.S. currently has research grants from the NIH; she is also supported by grant funding from NIH, NINDS, and CHDI; she declares no conflicts of interest.

McCormack, M.K. currently has grants from the NJ Department of Health; he declares no conflicts of interest.

Cella, D. receives grant funding from the National Institutes of Health and reports that he has no conflicts of interest.

Nance, M.A. currently has funding from the NIH, HDSA, CHDI, and the Parkinson's Foundation. She has received research funding in the last three years from Teva Pharmaceuticals, Biotie, and Sunovion. She has received honoraria from WebMD, Worrell Inc, Optio Biopharma, and Augsburg College. She declares on conflicts of interest.

Lai J.-S. currently has research grants from the NIH and Neurofibromatosis Therapeutic Acceleration Program at John Hopkins University; she declares no conflicts of interest.

Dayalu, P. currently has research grants from the NIH, Astra-Zeneca, and Vaccinex. He declares no conflicts of interest.

Author Roles:

Carlozzi, N.E.: Study Principal Investigator; Data Collection Site; Oversight for Statistical Analysis; Initial draft of Method, Results and Discussion; Incorporation of revisions

Boileau, N.R. Primary Data Analyst Responsible for Majority of Statistical Analyses; Assistance with Methods and Results Sections; Review and Feedback on manuscript drafts

Perlmutter, J.S. Study Co-Investigator; Review and Feedback on manuscript drafts (extensive revisions to introduction and discussion)

Chou, K.L. Study Co-Investigator; Review and Feedback on manuscript drafts (extensive revisions to introduction)

Stout, J.C. Study Co-Investigator; Review and Feedback on manuscript drafts

Paulsen, J.S. Study Co-Investigator; Review and Feedback on manuscript drafts

McCormack, M.K. Study Co-Investigator; Review and Feedback on manuscript drafts

Cella, D. Study Co-Investigator; Review and Feedback on manuscript drafts (specific contributions to analysis interpretation)

Nance, M.A. Study Co-Investigator; Review and Feedback on manuscript drafts (specific contributions to interpretation for HD clinicians)

Lai, J-S. Study Co-Investigator; Review and Feedback on manuscript drafts

Dayalu, P. Study Co-Investigator; Initial Draft of the Introduction; Assistance Writing Discussion Including Integration of Findings with Existing HD Literature; Review and Feedback on manuscript drafts

[2]Neurology, Radiology, Neuroscience, Physical Therapy and Occupational Therapy, Washington University in St. Louis

[3]Department of Neurology, University of Michigan

[4]School of Psychological Sciences, Monash University, Clayton, Victoria, Australia

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

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

[7]Department of Psychological and Brain Sciences, The University of Iowa, Iowa City, IA, USA

[8]Department of Psychiatry, Rutgers – RWJMS, Piscataway, NJ and Department of Pathology, Rowan – SOM, Stratford, NJ

[9]Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University

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

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

Abstract

Background—Clinician-rated measures of functioning are often used as primary endpoints in clinical trials and other behavioral research in Huntington disease. As study costs for clinician-rated assessments are not always feasible, there is a question of whether patient self-report of commonly used clinician-rated measures may serve as acceptable alternatives in low risk behavioral trials.

Aim—The purpose of this paper was to determine the level of agreement between self-report and clinician ratings of commonly used functional assessment measures in Huntington disease.

Design—486 participants with premanifest or manifest Huntington disease were examined. Total Functional Capacity, Functional Assessment, and Independence Scale assessments from the Unified Huntington Disease Rating scale were completed by clinicians; a self-report version was also completed by individuals with Huntington disease. Cronbach's alpha was used to examine internal consistency, one-way analysis of variance was used to examine group differences, and paired t-tests, kappa agreement coefficients, and intra-class correlations were calculated to determine agreement between raters.

Results—Internal consistency for self-reported ratings of functional capacity and ability were good. There were significant differences between those with premanifest, early-, and late-stage disease; those with later-stage disease reported less ability and independence than the other clinical groups. Although self-report ratings were not a perfect match with associated clinician-rated measures, differences were small. Cutoffs for achieving specified levels of agreement are provided.

Conclusions—Depending on the acceptable margin of error in a study, self-reported administration of these functional assessments may be appropriate when clinician-related assessments are not feasible.

Keywords

Huntington disease; functioning; clinician-ratings; self-report ratings

Introduction

Huntington disease (HD) is an autosomal dominant inherited, progressive neurodegenerative disorder that causes profound cognitive, behavioral, and motor decline.[1, 2] The disease often manifests near age 40 and leads to death in 15–20 years.[1, 3, 4] Patients lose function in many areas, including occupational, social, financial, and activities of daily living.[5–7] Slowing functional loss is a major unmet need, and a prime goal of HD clinical trials.[8, 9] These trials require functional rating scales that, in some cases, may act as a primary endpoint.[5, 10]

The Unified Huntington's Disease Rating Scale (UHDRS) is the most widely used assessment in clinical HD research.[11] The UHDRS includes clinician-rated assessments of functioning, motor symptoms, and cognition. With regard to the UHDRS functional assessments it includes: the Total Functional Capacity (TFC), the Functional Assessment (FA) Scale, and the Independence Scale.[12, 13] The TFC in particular is a frequently used as a primary outcome measure. [e.g.,] [14–16] The entire UHDRS, including these scales, was designed and validated as a clinician-rated instrument.[11] The clinician completes the scales after focused interviews with the affect individual and/or a close informant.

Some behavioral and observation research trials in HD would substantially benefit if these scales could be accurately completed by patient self-report. Patients could then complete their own functional ratings electronically from home, for at least some time points, reducing the participant burden for studies that require frequent administrations of functional scales.[17] This could potentially decrease study costs and enhance recruitment and retention.[17, 18] Furthermore, patient-reported outcome measures, particularly as they pertain to health-related quality of life (HRQOL; the effect of a disease on one's physical, cognitive, emotional, or social well-being [19, 20]) are becoming increasingly important as measures of efficacy for new treatments.[21, 22]

On the other hand, relying on patient self-report has potential disadvantages. One key question is whether self-ratings would be similar to (and therefore a reasonable substitute for) a trained clinician's ratings using these functional rating scales. Cognitive dysfunction occurs even in premanifest HD (before motor symptoms manifest),[5, 23–26] and patient-reported and caregiver-reported functional ratings may diverge in the early stages of disease, possibly related to subtle loss of awareness on the patient's part.[27–32] These discrepancies might increase in late-stage HD, when dementia could interfere with a subject's insight and item comprehension, or lead to perseverative responses. Another potential source of error is misunderstanding of a question's intent. Most notably, many items in the UHDRS are intended to gauge the *capacity* to perform function X or Y, regardless of whether the subject actually performs X or Y with any regularity. This potential source of confusion could be mitigated with a trained clinician guiding the rating.

The broader literature comparing self-report to clinical interview data is extremely variable. For example, while there is often consensus between self-reported and clinician interview for clinical diagnoses,[33–42] consensus reporting between the patient and clinician for specific symptoms and overall functioning can be highly variable [43–47]. Furthermore, the accuracy of retrospective self-reported symptoms may suffer from recall bias or memory decay[48, 49], and may be subject to exaggeration, especially in cases where litigation or compensation is involved [39–42]. Several different factors may account for this variability including the objectivity/subjectivity of the symptom that is the subject of the report, social desirability, time length for recall, the age and education of the informant, and the level of self-awareness of the rater. More work is needed to understand how patient self-report in HD relates to clinician ratings.

The primary aim of this analysis was to compare clinician-reported and patient-reported ratings of UHDRS functional measures. We hypothesized that for premanifest and early HD subjects, self-report would correlate well with clinician-report, but that increased discrepancy may occur in later stage HD. In addition, we examined the reliability of individual items (and groups) across different raters and compared the internal consistency of clinician and self-reported measures to determine if one measure was more stable than the other. We hypothesized that clinician-ratings would have greater stability than self-report measures. Finally, we examined the ability for both the clinician-ratings and self-report ratings to differentiate between those with premanifest versus manifest disease (i.e., known groups validity). Since those with manifest disease have greater functional deficits than those with premanifest disease [50], this would provide evidence of validity.

Methods

Participants

Participants with either premanifest (gene-positive status for the HD CAG expansion and no clinical diagnosis), or manifest HD were invited to participate in a study about HRQOL (characterization details provided below). This report focuses on a subset of participants that were examined as part of a larger study; detailed data about the full sample are reported elsewhere.[51] Briefly, participants must have been ≥18 years of age, able to read and understand English, and able to provide informed consent. This convenience sample included recruitment efforts through local HD treatment centers (at the University of Michigan, University of Iowa, University of California-Los Angeles, Indiana University, Johns Hopkins University, Rutgers University, Struthers Parkinson's Center, and Washington University in St. Louis), the HD Roster, existing online medical record data capture systems, [52] and community outreach. In addition, a portion of this sample was recruited in conjunction with the Predict-HD research study.[53] All study activities were conducted with the approval of local Institutional Review Boards and have therefore been performed in accordance with the ethical standards in the 1964 Declaration of Helsinki. Consent from participants was also obtained before subject participation began.

Characterization of Huntington Disease (HD)

The Total Motor Scale and the Total Functional Capacity score from the Unified Huntington's Disease Rating Scale (UHDRS)[11] were used to classify HD participants in this study. The 15-item clinician administered TMS evaluates multiple aspects of motor functioning (scores range from 0–124 with higher scores indicating more motor difficulties). The final item of this scale requires the clinician to indicate a confidence rating for whether or not the participant has unequivocal signs of HD (the rating scale ranges from 0 = No abnormalities to 4 = ≥99% confidence). In order to be classified as having premanifest HD, participants were required to have a positive gene test for the HD CAG expansion (CAG > 35) and a clinician confidence rating of ≤ 4 on the last item of the TMS. In order to be classified as manifest HD, participants were required to have a clinician-based rating of 4 on the last item of the TMS. Staging for manifest participants was determined based on TFC scores (which includes five multiple-choice clinician-rated items addressing impairment or assistance required in occupation, finances, domestic chores, ADLs, and level of care). Scores range from 0–13 with higher scores indicating better functioning; early-HD (stages I-II) had TFC scores between 7 and 13 and late-HD (stages III-IV) had TFC scores between 0 and 6.

Measures

Demographics—Demographic information was collected using an online data capture system (Assessment Center^{SM}). Demographic data included age, gender, marital status, race, and ethnicity.

Medical Record Confirmation—Medical record data were retrieved to confirm HD diagnosis, date of HD diagnosis, and gene test results (including CAG repeat length). For manifest participants, the date of HD diagnosis was used to determine how many years had passed from diagnosis to the time of their study visit. For premanifest participants, the CAG repeat length and age at the time of their study visit were used to calculate their CAP score [54], which categorizes participants as either Low, Intermediate, and High risk of developing motor symptoms within the next five years.

The Unified Huntington's Disease Rating Scale (UHDRS). [11]—The UHDRS is a standardized rating scale that assesses mobility, cognition, functionality, and independence. [11] We examined three measures: Total Functional Capacity (TFC), the Functional Assessment (FA) Scale, and the Independence Scale. Two different administrations of these assessments were completed; the first was clinician-rated, and the second was patient self-reported (see appendix for self-report versions of these assessments). The TFC includes five multiple-choice questions addressing impairment or assistance required in occupation, finances, domestic chores, ADLs, and level of care. Both clinician-rated and self-report TFC scores range from 0–13 with higher scores indicating better functioning. [11] The Functional Assessment Scale includes 25 yes/no questions pertaining to a participant's ability to perform common tasks related to occupation, finances, activities of daily living (ADL), domestic chores, and level of care. Scores range from 0–25 for the clinician-rated and self-report version of this measure, with higher scores indicating higher functioning. [11] The clinician-rated Independence Scale is rated from 1 to 100 in intervals of five, with higher

scores indicating higher functioning and lower scores indicating worse functioning.[11] The self-report Independence scale differs slightly, in that participants rate themselves on a scale of 0 to 10, without an option to select increments of 0.5. For the purposes of this analysis, the clinician-rated Independence scale was transformed by dividing the clinician score (ranging from 0 – 100) by 10 to match the scale of the self-report measure (0–10).

Statistical Analysis

Statistical analyses were performed with SAS 9.4 and SPSS Version 22.0.[55, 56] Descriptive statistics were analyzed separately for both the clinician rating and self-reported scores.

Reliability—Cronbach's alpha coefficients were calculated to determine internal consistency of the UHDRS TFC and Functional Assessment for the clinician-rated and participant self-report. A critical cutoff of 0.70 was considered minimal acceptable reliability.[57] We used a dependent alpha (Feldt) t-test[58] to compare clinician and self-report measures of the same test. The Independence Scale is comprised of a single item; therefore, internal consistency cannot be calculated.

Known Groups Validity—Known groups validity was assessed using a one-way ANOVA to determine whether the clinician and self-reported FA and Independence Scales could differentiate among the three HD groups. We hypothesized that premanifest participants would report higher levels of independence and functioning than the early-stage HD group, and that the early-stage HD group should report more independence and better functioning than the late-stage HD group.[50] Bonferonni (Dunn) t-tests were used to examine mean differences among the three groups.[59] Known groups validity was not assessed for TFC scores, since TFC scores are required to determine the HD staging groups.

Measurement Agreement between Clinician-Rated and Self-Report Scores—The clinician-rated and self-report scores were compared using a paired t-test to determine the statistical significance of the mean difference in total TFC, FA, and Independence Scale scores. Next, individual TFC and FA items were compared between the clinician-rated and participant self-report rating using Cohen's kappa coefficient.[60] A kappa score between 0.81 and 1 was considered excellent agreement, 0.61–0.80 was considered substantial (very good) agreement, 0.41–0.60 was moderate agreement, 0.21–0.40 was deemed fair agreement, and lower than 0.20 indicated no agreement.[60, 61] A kappa coefficient for the Independence Scale could not be calculated because the clinician and the participant-rated scales differed (i.e., the clinician was allowed to choose intervals of 5). Intra-class correlation coefficients (ICCs) were calculated to analyze the group level reliability of the clinician rated items and the self-reported items. Overall agreement was measured by finding the percent of participants whose self-reported scores and clinician rated assessment scores were an exact match. Then, agreement between clinician-rated and self-reported Functional Assessment scores and TFC scores was examined at ascending discrepancy allowances to assess how much discrepancy would allow for different levels (75%, 90%, 95%, and 100%) of agreement.

Results

Descriptive Data

We examined four hundred and eighty-six individuals with either premanifest (n=195) or manifest HD (early stage n=190 or late stage n=101). Groups did not differ on gender (X^2_2 = 4.0, p=. 1341). As expected, age significantly differed across groups (F[2, 483] = 45.3, p<0001); the average age of the premanifest group (*M* = 42.9; *SD* = 12.2) was nine years younger than the early group (*M* = 52.0; *SD* = 12.3) and thirteen years younger than the late stage group (*M* = 55.5; *SD* = 11.6). This is understandable given the progressive nature of the disease and the usual age of onset. Race (Fisher's Exact *p*=.0053) and ethnicity (Fisher's Exact *p*=. 0002) also differed across groups; the late-stage group had a higher proportion of African Americans than the other two groups, the late-stage group included a higher proportion of non-Hispanic participants than the other two groups and the premanifest group had a larger proportion of individuals whose ethnicity was not provided relative to the manifest groups. A description of the sample can be found in Table 1. The average score for the clinician rating of TFC was 9.54 (SD=3.68). This significantly differed (t[469]=6.10; p<0001) from the self-report ratings of the same measure (M=9.22; SD=3.97). For Functional Assessment, the clinician rating (M=19.02; SD=6.30) did not significantly differ from the self-report ratings (M=20.90; SD=5.57; t[302]=1.01; p=3141). Finally, clinician ratings on the Independence Scale (M=8.55; SD=1.61) were significantly lower (t[481]=6.24; p<0001) than self-reported scores (M=8.84; SD=1.58).

Reliability

Internal consistency data is provided in Table 2. Internal consistency for the clinician-rated TFC was good (Cronbach's α = 0.89) and FA was excellent (Cronbach's α = 0.95). Similarly, internal consistency was good for the self-reported TFC (Cronbach's α = 0.86) and excellent for the FA (Cronbach's α = 0.95). The internal consistency for the self-reported TFC (Cronbach's α = 0.86) was significantly lower than the clinician-rated TFC (t[468]=4.75; p<0001). Similarly, the self-report administration of the FA (Cronbach's α = 0.94) was also significantly lower than the clinician assessment (t[301]=2.16; p=.03), though numerically and clinically this difference is small.

Known Groups Validity

For the FA and the Independence Scale, both the clinician-rated measure and self-report measure were differentiated among the HD groups (Table 3). All findings were in the hypothesized direction (i.e., premanifest participants reported higher levels of independence and functioning than the early-stage HD group, and the early-stage HD group reported more independence and functioning than the late-stage HD group).

Measurement Agreement between Clinician-Rated and Self-Report Scores

The clinician-rated TFC significantly differed from self-report TFC (t[469]=6.10; p<.0001; Table 2). These significant differences were seen for both the premanifest (t[189]=4.31; p<. 0001) and early (t[180]=5.87; p<.0001) groups; in both cases clinicians rated participants as higher functioning than the participants themselves did. However, TFC scores did not

significantly differ in the late-stage group (t[98]=0; p=1). At the individual level, there was 21.5% agreement between clinician-rated and self-reported scores. To achieve 75%, 90%, 95%, and 100% agreement, clinician-rated and participant self-report scores had to be within 2, 4, 5, and 7 points of each other, respectively (Table 4). Although the kappa was .14 (indicating no inter-rater reliability) and agreement for the individual items was minimal to weak (kappa scores ranged from 0.17 to 0.58), the intra-class correlations for the total score was good (ICC = 0.88; 95% CI: 0.85, 0.90; see Table 5).

For the FA scale, clinician-ratings did not differ from the self-report scores for the overall sample (t[303]=1.01; p=.3141; Table 2). FA ratings significantly differed among the early-stage participants (t[166]= 3.75; p=.0002), but not for the premanifest (t[44]= 1.59; p=.12) or late-stage participants (t[90]= −1.62; p=.11; Table 3). For the FA total score, there was 29.9% exact agreement between the clinician and subjects, and a 75% agreement within a 2-point window. To achieve 75%, 90%, 95%, and 100% agreement, FA scores needed to be within 2, 5, 7, and 13 points, respectively (Table 4). Although the kappa was 0.23 (indicating minimal inter-rater agreement), the intra-class correlation coefficient was excellent for the total score (ICC = 0.90; 95%CI: 0.92, 0.95; Table 5). With regard to individual items, kappa scores ranging from 0.30 to 0.85, indicating minimal to strong rater agreement (Table 5).

For the Independence Scale, clinician-rated significantly differed from self-report scores (t[481]=6.24; p<.0001; Table 2). With regard to the different HD stages, there were also group differences for the early-stage group (t[188]=5.96; p<.0001) and the late stage group (t[97]=4.35; p<.0001), but not the premanifest group (t[193]=0.74; p=.46; Table 3). There was 52% agreement between the clinician-rated and self-report Independence Scale scores. To achieve 75%, 90%, 95%, and 100% agreement, clinician-rated and self-report scores need to be within 1 point, 2, points, 2 points, and 6 points, respectively (Table 4). Intra-class correlations for the Independence Scale were acceptable (ICC = 0.88; 95% CI: 0.85, 0.90; Table 5).

Discussion

The purpose of this study was to examine the accuracy of self-report reported functioning in individuals with HD relative to associated clinician-ratings of the same construct. While self-report ratings and clinician-ratings were not a perfect match, self-ratings provided a reliable and valid alternative to clinician-ratings, especially when in person assessments are impractical or cost prohibitive for a particular study, or for observational or behavioral trials that are low risk. Our results indicated that self-reported ratings of functional capacity and functional ability internal consistency were good. Whereas clinician-reported ratings of the same constructs were excellent (and consistently higher than self-report ratings), the fact that the self-report ratings exceeded minimal acceptable criterion for internal consistency (i.e., > .70[62]) supports the reliability of the self-reported assessments. In addition, construct validity of self-reported functional ability and independence was supported by significant differences between those with premanifest, early-, and late-stage HD, such that those individuals with manifest HD (the early- and late-HD groups) reported less ability and independence than those with manifest HD, and those with late-stage HD reported less

ability and independence than those with early-HD. This pattern of findings was identical for the associated clinician rating scores for both functional ability and independence.

Although self-reported ratings significantly differed from clinician-ratings of the TFC (these differences were small, i.e., within .25 points of one another), and the agreement between individual items were not ideal, total scores agreed well on this measure. Self-report did not significantly differ from clinician-rated functional ability. When HD stage was considered, group differences were also not found between self- and clinician-reports for those with premanifest or late-stage HD; there were group differences on self- and clinician-report for those with early-HD, but this difference was small (less than 1 point). Furthermore, whereas agreement at the item-level was poor, FA total scores showed good agreement. Self-report ratings significantly differed from clinician ratings on the Independence Scale, but differences were small (within 0.5 of a point of one another). When HD stage was considered, self- and clinician-report differed for the two manifest groups but not the premanifest group; again differences were small (within 0.5 of a point for the early-group and 0.75 of a point for the late-HD group). Furthermore, item-level agreement was not optimal, but total score agreement was acceptable. Thus, while these findings indicate that the self-report TFC, FA and Independence do not perfectly match the associated clinician-rated measures, the agreement is close enough to use with confidence in research studies that are behavioral or observational in design and are low risk, or where clinician-ratings are resource prohibitive or impractical.

Finally, Table 4 provides clinicians and researchers with recommendations for achieving specified levels of agreement. For example, if a clinician or researcher conducts a study utilizing self-reported TFC, they could be 75% confident that the self-reported score was within 2 points of the clinician-rated score, 90% confident that the clinician-rated score was within 4 points, 95% confident that it was within 5 points and 100% confident that it was within 7 points of clinician-rated TFC. On a scale where scores range from 0–13, one could rely solely on self-reported scores if 75% confidence was acceptable. If one wanted to achieve 90, 95, or 100% confidence, they would likely not rely on self-reported scores alone (as would likely be the case with medication based clinical trials). For FA, where scores range from 0 to 25, a clinician or researcher might be most comfortable using self-report if they required 90% agreement, in which case could be confident that the self-reported score was within 5 points of the clinician rated score which is ±20% of the score. Similarly, on Independence, where scores range from 0 to 100, a clinician or researcher might be most comfortable using self-report if they required 90%-95% agreement (in which case we could be confident that the self-reported score was within 20 points of the clinician rated score which is ±20% of the score). Thus, while these discrepancies may preclude the use of these self-report measures in pharmacological trials (where measurement error can be especially problematic[63, 64]), depending on the research question and the acceptable margin of error, Table 4 can be utilized to decide whether or not the benefits of using a self-reported score (efficiency, cost, etc), exceeds the margin of error that would be acceptable for any given research questions.

This study provides support for self-reported functional capacity, ability and independence, and helps determine when one might consider only using self-report and not clinician ratings

of these scales. The study has some limitations. First, cognitive impairment is common in HD,[28, 30, 65, 66] yet we did not administer any assessments to assess it directly, or estimate patient overall awareness in this study. Furthermore, we did not consider how other specific psychological or motivational factors such as depression, anxiety or apathy, may affect self-report or the relationship between self-report and clinician rating.

Regardless of these limitations, these findings provide important information for the clinical utility of the self-reported TFC, FA and Independence Scale. While the self-reported and clinician-rated scores are not identical, the self-reported versions demonstrated acceptable reliability. Any differences between scores were small, and there was generally good agreement for total scores supporting their clinical utility. Specifically, HD clinicians and researchers can consider administering the self-reported versions of these assessments when the study design permits, and obtaining clinician ratings if not feasible (either because of cost or availability of clinician time).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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

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

References

1. Roos RA. Huntington's disease: a clinical review. Orphanet J Rare Dis. 2010; 5(1):40. [PubMed: 21171977]
2. Paulsen JS. Early Detection of Huntington Disease. Future Neurol. 2010; 5(1)
3. Ross CA, et al. Huntington disease and the related disorder, dentatorubral-pallidoluysian atrophy (DRPLA). Medicine (Baltimore). 1997; 76(5):305–38. [PubMed: 9352736]
4. Walker FO. Huntington's disease. Lancet. 2007; 369(9557):218–28. [PubMed: 17240289]

5. Tabrizi SJ, et al. Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. Lancet Neurology. 2012; 11(1): 42–53. [PubMed: 22137354]
6. Frank SA, et al. Functional decline due to chorea in Huntington's Disease. Neurology. 2004; 62(7, Suppl. 5):A204.
7. van Vugt JP, et al. Quantitative assessment of daytime motor activity provides a responsive measure of functional decline in patients with Huntington's disease. Mov Disord. 2001; 16(3):481–8. [PubMed: 11391742]
8. Mestre T, et al. Therapeutic interventions for symptomatic treatment in Huntington's disease. Cochrane Database Syst Rev. 2009; (3):CD006456. [PubMed: 19588393]
9. Van Walsem MR, et al. Health-related quality of life and unmet healthcare needs in Huntington's disease. Health Qual Life Outcomes. 2017; 15(6):1–10. [PubMed: 28069015]
10. Paulsen JS, et al. Challenges assessing clinical endpoints in early Huntington disease. Mov Disord. 2010; 25(15):2595–603. [PubMed: 20623772]
11. Unified Huntington's Disease Rating Scale: reliability and consistency. Mov Disord. 1996; 11(2): 136–42. [PubMed: 8684382]
12. Huntington Study Group. Unified Huntington's Disease Rating Scale: Reliability and consistency. Movement Disorders. 1996; 11(2):136–42. [PubMed: 8684382]
13. Klempir J, et al. Unified Huntington's disease rating scale: clinical practice and a critical approach. Funct Neurol. 2006; 21(4):217–21. [PubMed: 17367582]
14. Huntington Study Group. A randomized, placebo-controlled trial of coenzyme Q10 and remacemide in Huntington's disease. Neurology. 2001; 57(3):397–404. [PubMed: 11502903]
15. Huntington Study Group. Tetrabenazine as antichorea therapy in Huntington disease. Neurology. 2006; 66(3):366–372. [PubMed: 16476934]
16. Landwehrmeyer B, et al. Riluzole in Huntington's disease: a 3-year, randomized controlled study. Annals of Neurology. 2007; 62:262–272. [PubMed: 17702031]
17. Dishman R, Washburn R, Schoeller D. Measurement of physical activity. QUEST. 2001; 53:295–309.
18. Paulhus D, Vazire S. The self-report method. Handbook of research methods in personality psychology. 2007:224–239.
19. Cella DF. Measuring quality of life in palliative care. Seminars in oncology. 1995; 22(2 Suppl 3): 73–81. [PubMed: 7537908]
20. World Health Organization, W. International Health Conference. New York: 1946. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference.
21. 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. [PubMed: 16423298]
22. Carlozzi NE, Tulsky DS. Identification of Health-Related Quality of Life (HRQOL) Issues Relevant to Individuals with HD. J Health Psychol. 2013; 18(2):212–225. [PubMed: 22427174]
23. Stout JC, et al. Evaluation of longitudinal 12 and 24 month cognitive outcomes in premanifest and early Huntington's disease. J Neurol Neurosurg Psychiatry. 2012; 83(7):687–94. [PubMed: 22566599]
24. Stout JC, et al. Neurocognitive signs in prodromal Huntington disease. Neuropsychology. 2011; 25(1):1–14. [PubMed: 20919768]
25. Tabrizi SJ, et al. Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. Lancet Neurology. 2009; 8(9):791–801. [PubMed: 19646924]
26. Tabrizi SJ, et al. Biological and clinical changes in premanifest and early stage Huntington's disease in the TRACK-HD study: the 12-month longitudinal analysis. Lancet Neurol. 2011; 10(1): 31–42. [PubMed: 21130037]

27. Hoth KF, et al. Patients with Huntington's disease have impaired awareness of cognitive, emotional, and functional abilities. J Clin Exp Neuropsychol. 2007; 29(4):365–76. [PubMed: 17497560]
28. Sitek EJ, et al. Poor insight into memory impairment in patients with Huntington disease. Neurol Neurochir Pol. 2012; 46(4):318–25. [PubMed: 23023430]
29. Sitek EJ, et al. Self-awareness of executive dysfunction in Huntington's disease: comparison with Parkinson's disease and cervical dystonia. Psychiatry Clin Neurosci. 2013; 67(1):59–62. [PubMed: 23331288]
30. Sitek EJ, et al. Self-awareness of motor dysfunction in patients with Huntington's disease in comparison to Parkinson's disease and cervical dystonia. J Int Neuropsychol Soc. 2011; 17(5): 788–95. [PubMed: 21729402]
31. Snowden JS, et al. Awareness of involuntary movements in Huntington disease. Arch Neurol. 1998; 55(6):801–5. [PubMed: 9626771]
32. Vitale C, et al. Unawareness of dyskinesias in Parkinson's and Huntington's diseases. Neurol Sci. 2001; 22(1):105–6. [PubMed: 11487181]
33. Stuart AL, et al. Comparison of self-report and structured clinical interview in the identification of depression. Compr Psychiatry. 2014; 55(4):866–9. [PubMed: 24467941]
34. Fischer A, et al. Diagnostic accuracy for major depression in multiple sclerosis using self-report questionnaires. Brain and Behavior. 2015; 5(9)
35. Laferton JAC, et al. Screening for DSM-5 Somatic Symptom Disorder: Diagnostic Accuracy of Self-Report Measures Within a Population Sample. Psychosomatic Medicine. 2017; 79(9):974–981. [PubMed: 28922210]
36. Hinkle SN, et al. Validation of Self-reported Diagnosis of Gestational Diabetes at 6-weeks Postpartum. Epidemiology. 2017; 28(5):747–752. [PubMed: 28570385]
37. Silverstein MJ, et al. How Informative Are Self-Reported Adult Attention-Deficit/Hyperactivity Disorder Symptoms? An Examination of the Agreement Between the Adult Attention-Deficit/Hyperactivity Disorder Self-Report Scale V1.1 and Adult Attention-Deficit/Hyperactivity Disorder Investigator Symptom Rating Scale. J Child Adolesc Psychopharmacol. 2017
38. Videm V, et al. Self-reported Diagnosis of Rheumatoid Arthritis or Ankylosing Spondylitis Has Low Accuracy: Data from the Nord-Trondelag Health Study. Journal of Rheumatology. 2017; 44(8):1134–1141. [PubMed: 28412703]
39. Russo AC, Fingerhut EC. Consistency of Self-Reported Neurocognitive Symptoms, Post-Traumatic Stress Disorder Symptoms, and Concussive Events From End of First Deployment to Veteran Health Administration Comprehensive Traumatic Brain Injury Evaluation by Operations Enduring Freedom/Iraqi Freedom/New Dawn Veterans. Archives of Clinical Neuropsychology. 2017; 32(2):184–197. [PubMed: 28365745]
40. DeViva JC, Bloem WD. Symptom exaggeration and compensation seeking among combat veterans with posttraumatic stress disorder. Journal of Traumatic Stress. 2003; 16(5):503–507. [PubMed: 14584635]
41. Rohling ML, Allen LM, Green P. Who is exaggerating cognitive impairment and who is not? CNS Spectr. 2002; 7(5):387–95. [PubMed: 15122110]
42. Krebber AM, et al. Prevalence of depression in cancer patients: a meta-analysis of diagnostic interviews and self-report instruments. Psychooncology. 2014; 23(2):121–30. [PubMed: 24105788]
43. Hamood R, et al. A feasibility study to assess the validity of administrative data sources and self-reported information of breast cancer survivors. Isr J Health Policy Res. 2016; 5:50. [PubMed: 27980719]
44. Lo T, et al. Discordance between self-reported arthritis and musculoskeletal signs and symptoms in older women. BMC Musculoskelet Disord. 2016; 17(1):494. [PubMed: 27905906]
45. Nazem S, et al. Actigraphic and Sleep Diary Measures in Veterans With Traumatic Brain Injury: Discrepancy in Selected Sleep Parameters. J Head Trauma Rehabil. 2016; 31(2):136–46. [PubMed: 26959667]
46. McIntyre JPR, et al. A description of sleep behaviour in healthy late pregnancy, and the accuracy of self-reports. Bmc Pregnancy and Childbirth. 2016:16. [PubMed: 26810220]

47. Kim SY, et al. Discrepancy between self-assessed hearing status and measured audiometric evaluation. Plos One. 2017; 12(8)
48. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol. 2008; 4:1–32. [PubMed: 18509902]
49. Van den Bergh O, Walentynowicz M. Accuracy and bias in retrospective symptom reporting. Current Opinion in Psychiatry. 2016; 29(5):302–308. [PubMed: 27427854]
50. Marder K, et al. Rate of functional decline in Huntington's disease. Neurology. 2000; 54(2):452. [PubMed: 10668713]
51. Carlozzi NE, et al. HDQLIFE: Development and assessment of health-related quality of life in Huntington disease (HD). Qual Life Res. 2016; 25(10):2441–55. [PubMed: 27522213]
52. Hanauer DA, et al. Supporting information retrieval from electronic health records: A report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE). Journal of Biomedical Informatics. 2015; 55:290–300. [PubMed: 25979153]
53. Paulsen JS, et al. Preparing for preventive clinical trials - The Predict-HD study. Archives of Neurology. 2006; 63(6):883–890. [PubMed: 16769871]
54. Zhang Y, et al. Indexing Disease Progression at Study Entry with Individuals At-Risk for Huntington Disease. Am J Med Genet b Neuropsychiatr Genet. 2011; 156(7):751–763.
55. Institute, S. SAS 9.4 language reference concepts. SAS Institute; Cary, NC: 2013.
56. Inc., S. Released. IBM Corp; Armonk, NY: 2013. IBM SPSS Statistics for Windows.
57. Cronbach LG. Coefficient alpha and the internal structure of tests. Psychometrika. 1951; 16:297–334.
58. Feldt LS. A test of hypothesis that Cronbach's alpha coefficient is the same for two tests administered to the same sample. Psychometrika. 1980; 49:99–105.
59. Dunn OJ. Multiple comparisons among means. Journal of American Statistical Association. 1961; 56(293):52–64.
60. Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960; 20(1):10.
61. McHugh M. Interrater reliability: the kappa statistic. Biochemia Medica. 2012; 22(3):7. [PubMed: 22384515]
62. Nunnally, J., Bernstein, I. McGraw-Hill. New York, NY: McGraw-Hill; 1994. Psychometric theory.
63. Travessa AM, et al. Fifteen Years of Clinical Trials in Huntington's Disease: A Very Low Clinical Drug Development Success Rate. J Huntingtons Dis. 2017; 6(2):157–163. [PubMed: 28671135]
64. Mestre T, et al. Therapeutic interventions for disease progression in Huntington's disease. Cochrane Database of Systematic Reviews. 2009; (3)
65. Deckel AW, Morrison D. Evidence of a neurologically based "denial of illness" in patients with Huntington's disease. Arch Clin Neuropsychol. 1996; 11(4):295–302. [PubMed: 14588934]
66. McCusker EA, et al. Unawareness of motor phenoconversion in Huntington disease. Neurology. 2013; 81(13):1141–1147. [PubMed: 23966256]

Key Statements

What is already known about the topic

Although clinician-rated assessments of functioning are often used as primary endpoints in Huntington disease research, there is a question of whether patient self-report may serve as acceptable alternatives in low risk observational or behavioral studies.

What the paper adds

This paper reports on the level of agreement between self-reported and clinician ratings of function in Huntington disease.

Implications for practice, theory or policy

While not an exact match for clinician-ratings of functioning, self-reported ratings of functioning may be appropriate for low risk observational or behavioral trials in Huntington disease when the costs associated with clinician ratings are not feasible.

Table 1

Demographic characteristics of Individuals with HD, stratified by HD group

Variable	Premanifest (N=195)	Early (N=190)	Late (N=101)	All (N=486)
Age (years) *				
M (SD)	42.9 (12.2)	52.0 (12.3)	55.5 (11.6)	49.0 (13.2)
Education (# of years) *				
M (SD)	15.9 (2.9)	14.7 (2.8)	14.3 (2.7)	15.1 (2.9)
Gender (%)				
Female	64.1	54.7	55.5	58.6
Male	35.9	45.3	44.5	42.4
Race (%) *				
White	97.4	97.4	94.1	96.7
African American	0.0	1.1	5.9	1.7
More than One Race	1.5	0.0	0.0	0.6
American Indian/Alaskan Native	0.0	0.5	0.0	0.2
Other	0.0	1.0	0.0	0.6
Not Provided	0.5	0.0	0.0	0.2
Ethnicity (%) *				
Not Hispanic or Latino	92.3	93.2	97.0	93.6
Hispanic or Latino	1.5	3.7	1.0	2.3
Not Provided	6.2	3.2	2.0	4.1
Marital Status (%) *				
Single, Never Married	15.9	15.3	9.9	14.4
Married	67.7	54.0	66.3	62.0
Separated/Divorced	13.3	23.3	20.8	18.8
Widowed	0.0	3.2	3.0	1.9
Living with Partner	3.1	4.2	0.0	2.9
CAG Repeats *				
M (SD)	42.1 (2.9)	43.0 (3.6)	44.5 (7.1)	42.8 (4.0)
Disease Burden (premanifest only) **		--	--	--
Low Risk (%)	84.6	--	--	--
Intermediate Risk (%)	0.0	--	--	--
High Risk (%)	15.4	--	--	--
Time since diagnosis in years M(SD) (manifest only) *	--	3.5 (3.6)	6.3 (4.5)	--

Note. Entries in the table represent percentage of participants unless otherwise specified;

* indicates significant group differences: age, (F[2, 483] = 45.3, p<.0001), early-HD participants were older than premanifest participants and late-HD was older than early-HD; education (# of years), premanifest participants had more years of education than the other two groups (F[2, 481]= 13.1; p<.0001); race (White, African American, More than One Race, American Indian/Alaskan Native, Other, Not Provided), late-HD had more African Americans than the other two groups, the premanifest group had more mixed race participants than the other groups (Fisher's Exact p = . 0053) ; ethnicity (Not Hispanic/Latino, Hispanic/Latino, Not Provided) , more premanifest participant did not provide an ethnicity, late-HD had more Non-Hispanic/Latinos than the other two groups (Fisher's Exact p=.0002); marital status (Single, Married, Separated/Divorced, Widowed,

Living with Partner), Premanifest participants were less likely to be separated/divorced or widowed than the other two groups (χ^2_8 = 21.9, p=.0051); and time since diagnosis (in years, manifest only), the late-HD group had approximately 3 more years since diagnosis than the early-HD group (t[135.5]= −4.7; p<.0001).

** Disease burden was measured using CAP scores[54] to determine low, intermediate, or high risk of developing motor symptoms within the next 5 years.

Table 2

Clinician-rated vs Self-Report Descriptive Statistics and Internal Consistencies for the Total Functional Capacity, Functional Assessment, and Independence Scale

		Clinician Ratings		Self-Report Ratings		Clinician vs. Self-Report ratings	
Variable	**N**	**Mean (SD)**	**Internal Consistency (Cronbach's Alpha)**	**Mean (SD)**	**Internal Consistency (Cronbach's Alpha)**	**Mean Scores *T* (df)**	**Internal Consistency *T* (df)**
Total Functional Capacity	470	9.54 (3.68)	0.89	9.22 (3.97)	0.86	6.10 (469) **	4.75 (468) **
Functional Assessment	303	19.02 (6.30)	0.948	20.90 (5.57)	0.942	1.01 (302)	2.16 (301) *
Independence Scale	482	8.55 (1.61)	-	8.84 (1.58)	-	6.24 (481) **	-

Note: T(df) denotes the t-value (degrees of freedom) from the paired t-tests comparing the clinician and self-report measures; Independence Scale has no Cronbach's Alpha as it is a single-item measure.

** $p < .0001$

* $p < .05$

Table 3

Known Groups Validity for the Functional Assessment and Independence Scales

Variable	Premanifest[1]	Early[2]	Late[3]	F-value	p-value	Partial η^2
Functional Assessment						
Clinician Rated Mean (SD) [a,b,c]	24.42 (1.29)	22.04 (2.60)	11.74 (5.59)	330.70	<.0001	0.65
Self-Report Mean (SD) [a,b,c]	24.22 (1.29)	21.35 (3.41)*	12.46 (5.97)	332.23	<.0001	0.59
Independence						
Clinician Rated Mean (SD) [a,b,c]	9.78 (0.59)	8.52 (0.96)	6.16 (1.23)	537.28	<.0001	0.69
Self-Report Mean (SD) [a,b,c]	9.77 (0.69)	8.98 (1.10)	6.71 (1.69)	248.21	<.0001	0.51

[1] Sample sizes for premanifest participants are 45 for the Functional Assessment and 194 for the Independence Scale

[2] Sample sizes for early-stage participants are 167 for the Functional Assessment and 189 for the Independence Scale

[3] Sample sizes for late-stage participants are 91 for the Functional Assessment and 98 for the Independence Scale

[a] Bonferonni (Dunn) analysis indicates that Premanifest group is significantly different from Early Group

[b] Bonferonni (Dunn) analysis indicates that Early group is significantly different from Late Group

[c] Bonferonni (Dunn) analysis indicates that Premanifest group is significantly different from Late Group

Table 4

Total score differences required to achieve specified levels of agreement

	75%	90%	95%	100%
Total Functional Capacity	2 Points	4 Points	5 Points	7 Points
Functional Assessment	2 Points	5 Points	7 Points	13 Points
Independence Scale	1 Point	2 Points	2 Points	6 Points

Table 5

Agreement (Kappa and Intra-class correlation coefficients) for individual items and total scores for the UHDRS measures

Variable	*Kappa*	*ICC (95% CI)*
Total Functional Capacity (TFC)		
Occupation	0.58	*0.92 (.90, .93)*
Chores	0.17	*0.12 (−.05, .27)*
ADL	0.34	*0.47 (.37, .56)*
Finances	0.47	*0.86 (.84, .87)*
Care Level	0.45	*0.13 (−.08, .30)*
TOTAL FUNCTIONAL CAPACITY TOTAL SCORE	**0.14**	***0.89 (.87, .91)***
Functional Assessment		
Could the subject engage in gainful employment in his/her accustomed work?	0.74	*0.85 (.81, .88)*
Could the subject engage in any kind of gainful employment?	0.64	*0.78 (.73, .83)*
Could the subject engage in any kind of volunteer or non-gainful work?	0.53	*0.69 (.62, .75)*
Could the subject manage his/her finances (monthly) without any help?	0.63	*0.78 (.72, .82)*
Could the subject shop for groceries without help?	0.66	*0.80 (.75, .84)*
Could the subject handle money as a purchaser in a simple cash (store) transaction?	0.78	*0.64 (.55, .71)*
Could the subject supervise children without help?	0.63	*0.78 (.72, .82)*
Could the subject operate an automobile safely and independently?	0.85	*0.92 (.90, .94)*
Could the subject do his/her own housework without help?	0.62	*0.76 (.70, .81)*
Could the subject do his/her own laundry (wash/dry) without help?	0.79	*0.88 (.85, .91)*
Could the subject prepare his/her own meals without help?	0.61	*0.76 (.70, .81)*
Could the subject use the telephone without help?	0.30	*0.46 (.33, .57)*
Could the subject take his/her own medications without help?	0.62	*0.77 (.71, .81)*
Could the subject feed himself/herself without help?	0.52	*0.69 (.61, .75)*
Could the subject dress himself/herself without help?	0.66	*0.80 (.75, .84)*
Could the subject bathe himself/herself without help?	0.77	*0.87 (.84, .90)*
Could the subject use public transportation to get places without help?	0.68	*0.81 (.76, .85)*
Could the subject walk to places in his/her own neighborhood without help?	0.73	*0.85 (.81, .88)*
Could the subject walk without falling?	0.35	*0.52 (.41, .62)*
Could the subject walk without help?	0.59	*0.75 (.68, .80)*
Could the subject comb hair without help?	0.40	*0.57 (.47, .66)*
Could the subject transfer between chairs without help?	0.49	*0.66 (.58, .73)*
Could the subject get in and out of bed without help?	0.46	*0.63 (.54, .70)*
Could the subject use the toilet/commode without help?	0.85	*0.92 (.90, .94)*
Could the subject's care still be provided at home?	0.55	*0.71 (.64, .77)*
FUNCTIONAL ASSESSMENT TOTAL SCORE	**0.23**	***0.94 (.92, .95)***
Independence Scale		
INDEPENDENCE SCALE TOTAL SCORE	**-**	**0.88 (.85, .90)**