OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Archives of Clinical Neuropsychology 26 (2011) 59–66

Archives of CLINICAL NEUROPSYCHOLOGY

Estimating Premorbid Functioning in Huntington's Disease: The Relationship between Disease Progression and the Wide Range Achievement Test Reading Subtest

Justin J.F. O'Rourke[1,2], William H. Adams[1], Kevin Duff[3], Joanne Byars[1], Peg Nopoulos[1], Jane S. Paulsen[1], Leigh J. Beglinger[1,*]

[1]*Department of Psychiatry, University of Iowa, Iowa City, IA, USA*
[2]*Department of Counseling Psychology, University of Iowa, Iowa City, IA, USA*
[3]*Department of Neurology, University of Utah, Salt Lake City, UT, USA*

*Corresponding author at: Department of Psychiatry, The University of Iowa, Psychiatry Research, 1-321 Medical Education Building, Iowa City, IA 52242, USA. Tel.: +1-319-335-8765; fax: +1-319-353-4438.

E-mail address: leigh-beglinger@uiowa.edu (L.J. Beglinger).

Accepted 30 October 2010

Abstract

The estimation of premorbid abilities is an essential part of a neuropsychological evaluation, especially in neurodegenerative conditions. Although word pronunciation tests are one standard method for estimating the premorbid level, research suggests that these tests may not be valid in neurodegenerative diseases. Therefore, the current study sought to examine two estimates of premorbid intellect, the Wide Range Achievement Test (WRAT) Reading subtest and the Barona formula, in 93 patients with mild to moderate Huntington's disease (HD) to determine their utility and to investigate how these measures relate to signs and symptoms of disease progression. In 89% of participants, WRAT estimates were below the Barona estimates. WRAT estimates were related to worsening memory and motor functioning, whereas the Barona estimates had weaker relationships. Neither estimate was related to depression or functional capacity. Irregular word reading tests appear to decline with HD progression, whereas estimation methods based on demographic factors may be more robust but overestimate premorbid functioning.

Keywords: Huntington's disease; movement disorders; basal ganglia; assessment; dementia

Introduction

Huntington's disease (HD) is a genetic neurodegenerative condition characterized by a triad of motor dysfunction, psychiatric symptoms, and cognitive decline. HD is formally diagnosed when an individual known to have the genetic mutation for the disease (i.e., CAG repeat expansion) or a family history of HD unequivocally manifests an extrapyramidal movement disorder (Walker, 2007). Once a diagnosis is made, the disease slowly progresses over an average of 15–20 years, resulting in death. Although motor dysfunction is the hallmark sign of HD, psychiatric symptoms and cognition also insidiously decline, causing severe functional impairment (Beglinger et al., 2010; Paulsen & Conybeare, 2005).

The signs and symptoms of HD are primarily caused by apoptosis of the basal ganglia's medium spiny neurons, which subsequently results in a disruption of the frontal-subcortical circuitry (Vonsattel & DiFiglia, 1998; Vonsattel, Keller, & Pilar Amaya, 2008). Cognitive symptoms commonly associated with the neuropathology of HD include retrieval-based memory deficits, executive dysfunction, and slowed cognitive efficiency (Tekin & Cummings, 2002), all of which coalesce into a subcortical dementia as the disease progresses. Cognitive symptoms are often present more than a decade prior to motor diagnosis and gradually decline as diagnosis approaches (Duff et al., 2010; Rowe et al., 2010). In addition to cognitive changes, striatal

doi:10.1093/arclin/acq088 Advance Access publication on 8 December 2010

degeneration is associated with a range of oculomotor and gross motor signs (e.g., choreiform movements), as well as overall functional incapacity (Douaud et al., 2006). Correlations between the psychiatric symptoms of HD and degeneration in the basal ganglia are less clear, although lesions in the frontal-subcortical circuitry are often associated with two of the most common neuropsychiatric conditions in HD: depression and apathy (Drevets, 2000).

Given the progressive nature of HD, researchers and clinicians often monitor their patients' cognitive decline over the course of the disease. Baseline neuropsychological test scores documenting prodromal cognitive abilities are not ordinarily available to clinicians; therefore, estimates of premorbid intellectual functioning are used as a proxy for comparison against the current assessment data. One of the most common methods of estimating premorbid IQ is through the use of irregular word pronunciation tools (Franzen, Burgess, & Smith-Seemiller, 1997). Word pronunciation is an accepted estimate because reading ability is strongly related to general intelligence, relatively resistant to neurological insult and cognitive decline, and allows access to previous knowledge and education without placing demands on other cognitive abilities (Willshire, Kinsella, & Prior, 1991).

The word reading subtest from the Wide Range Achievement Test (WRAT; Wilkinson, 1993; Wilkinson & Robertson, 2006) is one of the most widely used the measures of premorbid intelligence, with approximately 64% of professionals using it in their clinical work or research (Smith-Seemiller, Franzen, Burgess, & Prieto, 1997). The popularity of measures like the WRAT Reading subtest and the National Adult Reading Test (NART; Nelson & Willison, 1991) can be attributed to their simplicity and minimal patient burden, as well as previous empirical evidence suggesting that these tests remain valid in a range of neurological conditions, including dementia (Maddrey, Cullum, Weiner, & Filley, 1996; O'Carroll & Gilleard, 1986). Despite the popularity of single-word reading tests, research suggests that word reading ability may decline relatively early in some neurodegenerative diseases, which may lead to underestimation of premorbid IQ. For example, Taylor (1999) examined the NART in people with Alzheimer's disease (AD) or multi-infarct dementia and found that irregular word reading was related to the measures of disease progression (e.g., tests of memory and executive functioning). The authors concluded that word reading scores consistently underestimated premorbid abilities. Additional cross-sectional (Ashendorf, Jefferson, Green, & Stern, 2009) and longitudinal studies in AD have shown similar findings, with one study demonstrating that performances on a word reading test steadily declined concurrently with the Mini-Mental State Exam over multiple annual visits (Cockburn, Keene, Hope, & Smith, 2000). Inaccurate assessment of premorbid IQ is problematic because it can mask early cognitive declines and may result in classification errors.

Given the progressive nature of HD—including cognitive changes long before motor diagnosis—and the potential susceptibility of word reading tests to certain types of cognitive decline, the current study examined performances on the WRAT Reading subtest in a large sample of participants with mild-to-moderate HD. The utility of a demographically based regression equation, the Barona formula (Barona, Reynolds, & Chastain, 1984), was also examined to determine its utility in an HD population. Estimates produced by the WRAT were also compared with those of the Barona formula to establish a consistent standard to compare WRAT Reading scores against, since the Barona is based on static demographic variables that are not vulnerable to changes in cognitive functioning. The relationship between premorbid estimates and disease characteristics known to be correlated with striatal degeneration (i.e., motor dysfunction, depression, retrieval deficits, and functional decline) was also examined to determine if WRAT Reading scores were correlated with disease progression and dementia, which may indicate that word pronunciation tests are inaccurate measures of premorbid abilities in persons with manifest HD. Given the existing literature in other neurodegenerative conditions, it is expected that the premorbid estimate based on the reading of irregular words will be associated with disease progression.

Methods

Participants

Participants were 93 volunteers (42 men, 51 women) with a confirmed diagnosis of HD as determined by an independent neurologist and either a positive test for the HD gene expansion or a positive family history of HD. Approximately 96% ($n = 88$) of the sample was Caucasian, with the remainder being African American ($n = 3$) or Native American ($n = 2$). Data for 43 participants were obtained from clinical trials of atomoxetine ($n = 17$) and citalopram ($n = 26$), whereas the remaining participants ($n = 50$) were enrolled through annual clinical visits to the University of Iowa Huntington's Disease Society of America Center of Excellence. Some participants were in more than one study; in those cases, only the first evaluation for each participant was used to avoid redundancy. All participants were in the mild to moderate stages of disease progression according to the Shoulson and Fahn (1979) Total Functional Capacity scale (Mild: TFC Stage 1 $n = 44$, TFC Stage 2 $n = 35$; Moderate: TFC Stage 3 $n = 14$) within the Unified HD Rating Scale (UHDRS; Huntington's Study Group, 1996).

Inclusion criteria for the two clinical trials required participants to be in the mild stages of HD with TFC scores in Stage 1 or 2. Participants were also required to present with subjective cognitive complaints or cognitive problems on previous neuropsychological testing. Exclusion criteria included hypertension, tachycardia, cardiovascular or cerebrovascular disease, age >75, clinically significant head injury, neurological disorder or insult other than HD, learning disability, or other medical condition that was likely to affect cognitive function, history of ADHD symptoms in childhood, severe or unstable psychiatric conditions, and current substance abuse. Participants in the citalopram clinical trial were also excluded from enrollment if they were taking other antidepressants. No criteria for participation were applied to patients recruited through HD clinical appointments since they were receiving comprehensive clinical care for HD.

Procedure

Study procedures were approved by the Institutional Review Board at the University of Iowa and all participants gave written informed consent. Clinical patients signed consent for their data to be entered into a research database, but their care was not altered as part of participation in this study. Research assistants, trained and supervised by licensed neuropsychologists, administered a battery of neuropsychological tests judged to be sensitive to the progression of cognitive declines in HD. Tests included measures of executive functioning, attention, memory, processing speed, visuoperception, and language. Measures differed depending on which study the participants were enrolled in, so there is not complete overlap of the data for the three subject pools. Test batteries were administered in one session and typically took participants 60–90 min to complete. Data on psychiatric symptoms, overall functional capacity, and demographic background were collected through the UHDRS, a widely used research rating scale in HD studies. Participants were also evaluated by certified motor raters for the motor signs of HD using the standardized motor component of the UHDRS.

Measures

WRAT-3rd edition/4th edition: Reading subtest. The WRAT-3 Reading subtest is a 57-item oral word reading test that assesses the ability to correctly read letters and irregularly pronounced words (Wilkinson, 1993). The first 15 items are letter identification items followed by 42-word pronunciation items. The WRAT-4 Reading subtest (Wilkinson & Robertson, 2006) is a recently updated version with 70 parallel items presented in a similar format. Raw scores for both versions of the WRAT are converted to age-corrected standard scores ($M = 100$, $SD = 15$).

All but three of the original items from the WRAT-3 were retained in the WRAT-4, and the validation of the WRAT-4 was an extension of the WRAT-3 validation process since the WRAT-4 items were retained based on standardization data for the WRAT-3 (Wilkinson & Robertson, 2006). Unlike the WRAT-3, the WRAT-4 Reading subtest has not been formally established as an estimate of premorbid functioning in the research literature; however, its substantial overlap with the WRAT-3 and its shared validation process suggests that the two versions of the test are similar premorbid estimators. To ensure their similarity in the current study, we conducted a *t*-test comparing the WRAT-3 and WRAT-4 Reading subtest standard scores to establish that the two versions of the test could be collapsed into a single WRAT Reading subtest variable.

Barona IQ estimation formula. The Barona estimation formula (Barona, Reynolds, & Chastain, 1984) is a regression equation that predicts IQ using a variety of demographic factors, including age, education, race, gender, region of residence, and occupation. Within the regression equation, education, race, and occupation are the strongest predictors of premorbid IQ. Barona, Reynolds, and Chastain established convergent validity for the Barona formula with the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), where it demonstrated correlations of .62, .49, and .60 with the Verbal IQ, Performance IQ, and Full-Scale IQ scales, respectively. Additionally, the Barona formula has been shown to be fairly robust in clinical samples, even when compared with other estimates that incorporate demographic information (Powell, Brossart, & Reynolds, 2003).

UHDRS motor assessment. The motor component of the UHDRS is a 15-item assessment of characteristic HD motor signs (e.g., chorea, bradykinesia, oculomotor dysfunction, etc.). Trained motor examiners rate each item on a 5-point scale (0–4), with higher scores indicating greater impairment. The scale's total motor score (TMS) ranges from 0 to 124 since many items are scored for both left- and right-sided functioning.

UHDRS total functional capacity. The TFC scale is a gross measure of functional capacity that assesses occupation, finances, domestic chores, activities of daily living, and care level. Composite TFC scores range from 0 to 13 with higher scores

indicating greater levels of functioning. The TFC total score is widely used to classify the progression of HD and has been grouped into five categories: Stage 1, 11–13; Stage 2, 7–10; Stage 3, 3–6; Stage 4, 1–2; and Stage 5, 0.

UHDRS depression assessment. A two-part item from the psychiatric assessment component of the UHDRS was used to assess depression. The frequency and the severity of depressive symptoms within the previous month were rated by clinicians with input from the patient and family. Higher scores indicate more frequent or severe depression. Frequency was rated from 0 (depressed mood never or almost never) to 4 (depressed mood frequently, most of the time). Similarly, severity was rated from 0 (no mood disturbance) to 4 (significant suffering and loss of functioning). A single depression score was derived from the product of the frequency and severity ratings of the UHDRS depression item.

Measures of delayed memory. Two measures of delayed memory were used to assess for HD-related dementia. Clinical trial participants completed the delayed recall trial of the Hopkins Verbal Learning Test-Revised (HVLT-R; Brandt & Benedict, 2001) and HD clinical patients completed the Delayed Memory Index of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998). The HVLT-R delayed recall subtest requires that participants recall up to 12 words after a 25-min delay following an initial learning presentation of the list three times. The Delayed Memory Index of the RBANS is a composite of word list, story, and figure delayed recall tasks, as well as a word-list recognition trial. HVLT-R raw scores are converted to *T*-scores ($M = 50$, $SD = 10$), and RBANS scores are age-corrected standard scores ($M = 100$, $SD = 15$). Delayed memory scores from the RBANS and HVLT-R were transformed into *z*-scores for comparability purposes.

Analysis

Descriptive statistics and frequency distributions were analyzed for all demographic and clinical variables (i.e., race, sex, age, and education). Standard scores ($M = 100$, $SD = 15$) for the WRAT Reading subtest and Barona formula were calculated, as well as the discrepancy between the two (WRAT–Barona). Correlations and *t*-tests were used to examine group differences on the clinical measures of disease progression (i.e., TFC, TMS, UHDRS depression item, and delayed memory scores) based on race, sex, age, and education. The comparisons of demographic groups were conducted to ensure that clinical variables were not unduly affected by race, sex, age, and education because of the role these demographic factors have in the Barona formula and WRAT Reading scaled scores. Bivariate correlations were then used to examine relationships between estimates of premorbid intelligence (i.e., WRAT Reading, Barona) and clinical characteristics of disease. The HVLT-R and RBANS delayed memory scores were correlated separately since the RBANS memory score consisted of additional tests of delayed memory. Mean comparisons and effect sizes were calculated to characterize differences in WRAT score between individuals who were impaired and non-impaired according to the HD clinical measures found to be correlated with in WRAT Reading performance.

Results

Descriptive statistics for demographic and clinical characteristics are provided in Table 1 for all 93 participants. As expected, motor and functional capacity scores indicated that participants were in the mild-to-moderate stages of disease.

Table 1. Demographic and clinical characteristics of HD participants ($N = 93$)

Characteristic	Mean	*SD*	Range
Age (years)	46.4	12.6	19–80
Education (years)	13.8	2.18	8–19
Total motor score	31.4	16.5	2–69
Total Functional Capacity	9.66	3.13	3–13
UHDRS Depression[a]	2.58	4.17	0–16
Delayed Memory *z*-scores[b]	−1.71	1.39	−4.00 to 1.10
WRAT Reading[c]	92.9	11.2	63–117
Barona Index[c]	107.5	7.54	78.78–118.54
WRAT-Barona Discrepancy[c]	−14.6	11.1	−42.31 to 13.72

Notes: HD = Huntington's disease; UHDRS = Unified Huntington Disease Rating Scale; WRAT = Wide Range Achievement Test.
[a]Depression frequency × severity.
[b]*z*-scores based on normative information for the HVLT-R and RBANS.
[c]Standard scores with $M = 100$, $SD = 15$.

Average motor scores suggested that this sample had notable motor impairments and the majority of the sample ($n = 68$) reported functional decline to some degree. Delayed memory scores suggested that almost half of the sample ($n = 44$) also had clinically significant memory deficits (i.e., 2 *SD* below the mean of normal individuals). Depression scores were positively skewed, as most participants ($n = 53$) did not report any depressed mood.

Age, education, and sex were not related to each other; however, there were differences in the education level based on race ($F_{2,90} = 5.32$, $p = .007$), with Native Americans having fewer years of education ($M = 9.0$, $SD = 1.41$) than African American ($M = 13.0$, $SD = 3.61$) or Caucasian ($M = 13.80$, $SD = 2.05$) participants. Age and race were the only demographic factors not associated with any of the clinical characteristics of HD (i.e., TFC, TMS, UHDRS depression item, and delayed memory scores). There were sex differences—$t(91) = -2.33$, $p = .022$—in delayed memory scores with men having lower *z*-scores ($M = -2.08$, $SD = 1.30$) than women ($M = -1.41$, $SD = 1.40$). Similarly, years of education was related to delayed memory scores—$t(91) = 2.12$, $p = .037$—with college graduates having higher *z*-scores ($M = -1.16$, $SD = 1.31$) than those with <16 years of education ($M = -1.88$, $SD = 1.38$).

Table 2 presents participants according to the standard scores derived from the WRAT Reading subtest and the Barona formula. A *t*-test comparing WRAT-3 and WRAT-4 Reading subtest standard scores indicated that there was no statistical difference between the two forms—$t(85) = 1.36$, $p = .18$; therefore, the two WRAT Reading forms were collapsed into a single WRAT Reading subtest variable. WRAT Reading scores were normally distributed with a mean of 92.9 ($SD = 11.2$). The mean Barona score was higher at 107.5 ($SD = 7.54$). The distribution of the Barona scores was negatively skewed with almost half the sample ($n = 45$) having an estimated IQ >110. Accordingly, the Barona formula provided estimates of premorbid intelligence that were higher than WRAT Reading estimates for 89.3% ($n = 83$) of the participants. The average discrepancy between estimates was nearly 1 *SD* with a mean difference of −14.6 ($SD = 11.0$) standard score points. For the 10.7% ($n = 10$) of participants with higher WRAT Reading scores than the Barona, WRAT Reading scores were an average of 4.60 ($SD = 3.91$) points above the Barona estimate, with discrepancies ranging from 0.02 to 13.72 points. On HD symptom measures, there were no statistically significant differences between the group who obtained higher WRAT Reading scores than the Barona compared with the group who obtained higher Barona estimates than WRAT Reading scores.

Bivariate correlations between demographic variables, clinical measures, and premorbid intelligence estimates are found in Table 3. Poorer performances on delayed memory and TMS were related to lower WRAT Reading scores. Delayed memory scores, regardless of the measure used (i.e., HVLT-R vs. RBANS), accounted for a significant portion of the variance in WRAT Reading scores ($r^2 = .24$ for the entire sample). TMS also contributed to the variance in WRAT Reading scores, although to a lesser degree ($r^2 = .06$). In contrast to WRAT Reading scores, Barona scores had a weak relationship with delayed memory and were not related to TMS. UHDRS depression scores and total functional capacity were not related to either estimate of premorbid intelligence. Interestingly, delayed memory was the only clinical variable that correlated with discrepancies between the two measures, as it accounted for 11% of the variability in the discrepancy score. The Barona formula and the WRAT–Barona discrepancy score also correlated with demographic factors, which is to be expected since demographic variables are a key component of the Barona formula.

Because both motor dysfunction and memory scores were associated with WRAT Reading, we conducted brief follow-up analyses to examine group differences in WRAT Reading performance based on motor and delayed memory scores. Since there are no established cutoff scores for differing levels of motor impairment, we divided the sample into two extreme groups (i.e., highest and lowest 25%). Differences between the upper and lower quartile of participants based on motor score were not significant for the Barona, but were significant for WRAT Reading—$t(46) = 2.50$, $p = .016$. There was a spread of nearly eight standard score points between the quartiles on WRAT Reading (lower quartile $M = 97.4$, $SD = 10.4$; upper quartile $M = 89.6$, $SD = 11.3$) with an effect size of $d = 0.72$. We also examined group differences based on delayed memory scores. The sample was separated into those with scores indicative of memory impairment ($z = -2.00$) versus the remainder of the sample. Again, there were no differences between the groups' Barona estimates; however, WRAT Reading scores were markedly different—$t(91) = 5.52$, $p \leq .0001$—and approximately 11 points lower for participants below the cut point (nonimpaired $M = 98.2$, $SD = 8.9$; impaired $M = 87.0$, $SD = 10.7$) with an effect size of $d = 1.14$.

Table 2. The number of participants with HD in each standard score group on the WRAT-Reading subtest and Barona IQ estimation formula ($N = 93$)

Premorbid estimate	Standard scores					
	<70	70–79	80–89	90–99	100–109	110–119
WRAT Reading subtest	3	9	17	38	19	7
Barona Index	0	1	0	15	32	45

Notes: HD = Huntington's disease; WRAT Reading = Wide Range Achievement Test-Reading subtest.

Table 3. Bivariate correlations between demographic and HD clinical characteristics with estimates of premorbid intelligence ($N = 93$)

Characteristic	WRAT Reading	Barona	WRAT Reading/ Barona Discrepancy
Gender	0.143	−0.204*	0.286**
Race	−0.038	−0.372**	0.215*
Age	−0.019	0.309**	−0.231*
Education	0.372**	0.838**	−0.194
Total motor score	−0.240*	−0.138	−0.150
Total Functional Capacity	0.146	0.144	0.51
Depression[a]	−0.003	−0.200	0.134
Delayed Memory			
Entire Sample	0.488**	0.254*	0.330**
Clinical Trial Participants ($n = 43$; HVLT-R Delayed Memory Trial)	0.558*	0.234	0.378*
HD Clinic Participant ($n = 50$; RBANS Delayed Memory Index)	0.442**	0.103	0.399**

Notes: HD = Huntington's disease; HVLT-R = Hopkins Verbal Learning Test-Revised; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; WRAT Reading = Wide Range Achievement Test-Reading subtest.

[a]UHDRS Depression item (Frequency × Severity).

*$p < .05$.

**$p < .01$.

Discussion

The present study examined two estimates of premorbid intellect in diagnosed HD and compared them with markers of disease progression. Our findings were consistent with the literature suggesting that word reading tests are confounded by progressing disease manifestation (i.e., motor dysfunction and lower memory scores in this study). WRAT Reading produced a lower estimated IQ than the Barona in 89% of participants, with the majority of those discrepant by more than 15 points (1 *SD*), and as much as 42 points in one case. WRAT Reading scores were also significantly lower than what was expected based on our sample's level of formal education; the sample as a whole had almost 2 years of college and almost one third had a college degree or higher. Although the mean estimated IQ from the WRAT was at the low end of the average range, the average Barona was consistently at the high end of average. Such discrepancies indicate that WRAT Reading likely underestimates premorbid intelligence in HD due to confounding by disease-related variables. In contrast, the high estimates produced by the Barona indicate that it may cause the opposite problem by potentially overestimating premorbid intellectual functioning, which could lead to over diagnosing cognitive decline.

WRAT Reading performance was more strongly related to the clinical features of HD than the Barona. Poorer WRAT Reading was associated with greater motor dysfunction and, by extension, progressive neural degeneration. Although the influence of specific motor signs on reading ability was not directly assessed, we suspect that motor-speech dysfunction likely affected WRAT Reading performances since previous research has shown that dysarthria directly affects oral reading ability in HD (Podoll, Caspary, Lange, & Noth, 1988). Worse memory functioning was also notably correlated with lower WRAT Reading subtest scores. A positive relationship between WRAT scores and delayed memory was present even when the HVLT-R and RBANS were analyzed separately; indicating that the added delayed memory components of the RBANS did not overly complicate our results. These findings suggest that the subcortical dementia that typifies cognitive dysfunction in diagnosed HD affects irregular word reading ability, which is consistent with what is observed in classic cortical dementias such as AD (Cockburn et al., 2000; Taylor, 1999). For example, McFarlane, Welch, and Rodgers (2006) found that individuals with minimal and mild AD differed significantly from each other and healthy comparisons on the NART, even after controlling for demographic factors (e.g., education). They also found that demographically based estimates were consistently higher than word reading tests and were not related to disease severity. Similarly, we found that the Barona lacked an association with most HD signs and symptoms. Barona scores were associated with memory functioning; however, this relationship was not significant when the two memory measures were analyzed separately. Of note, neither premorbid estimate was associated with functional capacity or depression.

Our follow-up analyses provide further evidence that the WRAT Reading subtest is an unreliable estimate of premorbid intelligence in HD. Despite the small amount of variance accounted for by motor and memory scores in WRAT Reading performances (6% and 24%, respectively), a remarkable decline of nearly one half to two thirds a standard deviation occurred on WRAT Reading when motor and memory impaired and nonimpaired participants were compared. Effect sizes also ranged from medium to large between the groups on these clinical measures. On the basis of these findings, we conclude that WRAT Reading may become increasingly invalid as people progress from the mild-to-moderate stages of HD.

On a broader level, our results suggest that premorbid estimates reliant on current cognitive abilities may be less robust than estimates based on static variables as neurodegenerative conditions progress; a finding which may have important clinical implications. However, a caveat is in necessary when interpreting our results. Although estimates based on static variables (i.e., demographic characteristics) may be more robust, we cannot conclude that such estimates are accurate and reliable. The validity of both word reading tests and demographically based estimates have been called into question, especially for those at the tail ends of the IQ distribution (Powell, Brossart, & Reynolds, 2003; Spinks et al., 2009). The Barona formula itself has a tendency to overestimate IQ by six standard score points for individuals in the average range and 20 points for those who are below average (Basso, Bornstein, Roper, & McCoy, 2000). This tendency to overestimate IQ likely accounts for the large number of participants in our study that have the Barona estimates >110 standard score points. Again, the Barona appears to overestimate premorbid IQ in our sample, which may increase the probability of diagnosing cognitive decline when none is present. A possible solution to the flaws of estimates based solely on the current cognitive skills or demographics may be the development of measures that combine both types of estimates (e.g., the OPIE-3; Schoenberg, Duff, Scott, & Adams, 2003); however, further investigation is needed. Even "combined" estimates may have reduced accuracy in HD because they often rely on skills like reading and attention.

The integrity of irregular word reading tests also needs further investigation in the prodromal phase of HD since neuropathological changes can manifest a decade before diagnosis. Aylward and colleagues (2004) reported volumetric changes in the head of the caudate 11 years prior to diagnosis and in the putamen 9 years before diagnosis. Impaired episodic memory (Montoya et al., 2006; Solomon et al., 2007), psychiatric disturbances (Duff, Paulsen, Beglinger, Langbehn, & Stout, 2007; Folstein, Abbott, Chase, Jensen, & Folstein, 1983), and mild motor dysfunction (Biglan et al., 2009) have all been observed in the years prior to the development of clinical motor signs warranting a diagnosis. Continued investigation may demonstrate that the less severe signs and symptoms of prodromal HD do not have the same deleterious effects on irregular word reading tests as they do in diagnosed HD.

Limitations of our study include the absence of a true premorbid assessment of intellect for our sample. Without complete neuropsychological data prior to the onset of HD neither the accuracy of WRAT Reading nor the Barona estimates can be fully determined. Nonetheless, we were able to make inferences about their utility in HD by examining clinical correlates for the two premorbid estimates. Another limitation is the lack of formal inclusion/exclusion criteria for participants recruited through the HD clinic. As a result, there may have been a sampling bias against severely impaired patients whose poor cognitive status precluded neuropsychological assessment. Caution should also be used in the interpretation of the TFC and UHDRS psychiatric assessments since they are broad measures that may be better utilized for screening purposes. Lastly, the limited level of depression in our sample provided a restricted range of scores (nearly half the sample who participated in the clinical trials were screened and excluded for significant depression) and may not accurately reflect what is present in the general HD population. More work is needed to clearly establish whether and how aspects of daily functioning or depression in HD are related to premorbid estimates of intelligence.

Performances on irregular word reading tests progressively decline with HD disease progression and may be susceptible to the effects of dementia in HD. No single estimate will provide a perfectly accurate assessment of premorbid intelligence in HD, especially in light of the disease's insidious nature and its pervasive affect on cognitive, psychiatric, and motor functioning. Therefore, there is a need to develop measures that mitigate the effects of current cognitive abilities on premorbid estimates. Although irregular word reading tests are commonly used among neuropsychologists, they may lead to the improper classification, diagnosis, and treatment planning for patients with HD.

Funding

This work is supported by the National Institutes of Health (NS055733-01); Eli Lilly and Company; Cure Huntington's Disease Initiative, Inc.; and the Huntington Disease Society of America.

Conflict of interest

None declared.

Acknowledgements

We are grateful for the assistance of Anne Leserman and her coordination of the clinic at the University of Iowa Huntington's Disease Society of America Center of Excellence. We would also like to thank Alysse Flynn, Jana Hanson, and Nicole Ramza for their assistance with clinical trial participants. Finally, we also want to thank the HD families who volunteer their time to assist in clinical research.

References

Ashendorf, L., Jefferson, A. L., Green, R. C., & Stern, R. A. (2009). Test-retest stability on the WRAT-3 reading subtest in geriatric cognitive evaluations. *Journal of Clinical and Experimental Neuropsychology*, *31*, 605–610.

Aylward, E. H., Sparks, B. F., Field, K. M., Yallapragada, V., Shpritz, B. D., Rosenblatt, A., et al. (2004). Onset and rate of striatal atrophy in preclinical Huntington disease. *Neurology*, *63*, 66–72.

Barona, A., Reynolds, C. R., & Chastain, R. (1984). A demographically based index of premorbid intelligence test for the WAIS-R. *Journal of Consulting and Clinical Psychology*, *52*, 885–887.

Basso, M. R., Bornstein, R. A., Roper, B. L., & McCoy, V. L. (2000). Limited accuracy of premorbid intelligence estimators: A demonstration of regression to the mean. *The Clinical Neuropsychologist*, *14*, 325–340.

Beglinger, L. J., O'Rourke, J. J. F., Wang, C., Langbehn, D. R., Duff, K., Paulsen, J. S., et al. (2010). Earliest functional declines in Huntington's disease. *Psychiatric Research*, *17*, 414–418.

Biglan, K. M., Ross, C. A., Langbehn, D. R., Aylward, E. H., Stout, J. C., Queller, S., et al. (2009). Motor abnormalities in premanifest persons with Huntington's disease: The PREDICT-HD study. *Movement Disorders*, *24*, 1763–1772.

Brandt, J., & Benedict, R. H. B. (2001). *Hopkins Verbal Learning Test-Revised*. Lutz, FL: Psychological Assessment Resources.

Cockburn, J., Keene, J., Hope, T., & Smith, P. (2000). Progressive decline in NART score with increasing dementia severity. *Journal of Clinical and Experimental Neuropsychology*, *22*, 508–517.

Douaud, G., Gaura, V., Ribeiro, M. J., Lethimonnier, F., Maroy, R., Verny, C., et al. (2006). Distribution of grey matter atrophy in Huntington's disease patients: A combined ROI-based and voxel-based morphometric study. *Neuroimage*, *32*, 1562–1575.

Drevets, W. C. (2000). Functional anatomical abnormalities in limbic and prefrontal cortical structures in major depression. *Progress in Brain Research*, *126*, 413–431.

Duff, K., Paulsen, J. S., Beglinger, L. J., Langbehn, D. R., & Stout, J. C. (2007). Psychiatric symptoms in Huntington's disease before diagnosis: The Predict-HD study. *Biological Psychiatry*, *62*, 1341–1346.

Duff, K., Paulsen, J., Mills, J., Beglinger, L. J., Moser, D. J., Smith, M. M., et al. (2010). Mild cognitive impairment in prediagnosed Huntington disease. *Neurology*, *75*, 500–507.

Folstein, S., Abbott, M. H., Chase, G. A., Jensen, B. A., & Folstein, M. F. (1983). The association of affective disorder with Huntington's disease in a case series and in families. *Psychological Medicine*, *13*, 537–542.

Franzen, M. D., Burgess, E. J., & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology*, *12*, 711–738.

Huntington's Study Group (1996). Unified Huntington's Disease Rating Scale: Reliability and consistency. *Movement Disorders*, *11*, 136–142.

Maddrey, A. M., Cullum, C. M., Weiner, M. F., & Filley, C. M. (1996). Premorbid intelligence estimation and level of dementia in Alzheimer's disease. *Journal of the International Neuropsychological Society*, *2*, 551–555.

McFarlane, J., Welch, J., & Rodgers, J. (2006). Severity of Alzheimer's disease and effect on premorbid measures of intelligence. *British Journal of Clinical Psychology*, *45*, 453–463.

Montoya, A., Pelletier, M., Menear, M., Duplessis, E., Richer, F., & Lepage, M. (2006). Episodic memory impairment in Huntington's disease: A meta-analysis. *Neuropsychologia*, *44*, 1984–1994.

Nelson, H. E., & Willison, J. R. (1991). *National adult reading test (NART): Test manual* (2nd ed.). Windsor, UK: NFER-Nelson.

O'Carroll, R. E., & Gilleard, C. J. (1986). Estimation of premorbid intelligence in dementia. *British Journal of Clinical Psychology*, *25*, 157–158.

Paulsen, J. S., & Conybeare, R. A. (2005). Cognitive changes in Huntington's disease. *Advances in Neurology*, *96*, 209–225.

Podoll, K., Caspary, P., Lange, H. W., & Noth, J. (1988). Language functions in Huntington's disease. *Brain*, *111*, 1475–1503.

Powell, B.D., Brossart, D.F., & Reynolds, C.R. (2003). Evaluation of the accuracy of two regression-based methods for estimating premorbid IQ. *Archives of Clinical Neuropsychology*, *18*, 277–292.

Randolph, C. (1998). *Repeatable battery for the assessment of neuropsychological status*. San Antonio, TX: The Psychological Corporation.

Rowe, K. C., Paulsen, J. S., Langbehn, D. R., Duff, K., Beglinger, L. J., Wang, C., et al. (2010). Self-paced timing task detects subtle changes in prodromal Huntington disease. *Neuropsychology*, *24*, 435–442.

Schoenberg, M. R., Duff, K., Scott, J. G., & Adams, R. L. (2003). An evaluation of the clinical utility of the OPIE-3 as an estimate of premorbid WAIS-III FSIQ. *The Clinical Neuropsychologist*, *17*, 308–321.

Shoulson, I., & Fahn, S. (1979). Huntington disease: Clinical care and evaluation. *Neurology*, *29*, 1–3.

Smith-Seemiller, L., Franzen, M. D., Burgess, E. J., & Prieto, L. R. (1997). Neuropsychologists' practice patterns in assessing premorbid intelligence. *Archives of Clinical Neuropsychology*, *12*, 739–744.

Solomon, A. C., Stout, J. C., Johnson, S. A., Langbehn, D. R., Aylward, E. H., Brandt, J., et al. (2007). Verbal episodic memory declines prior to diagnosis in Huntington's disease. *Neuropsychologia*, *45*, 1767–1776.

Spinks, R., McKirgan, L. W., Arndt, S., Caspers, K., Yucuis, R., & Pfalzgraf, C. J. (2009). IQ estimate smackdown: Comparing IQ proxy measures to the WAIS-III. *Journal of the International Neuropsychological Society*, *15*, 590–596.

Taylor, R. (1999). National Adult Reading Test performance in established dementia. *Archives of Gerontology and Geriatrics*, *29*, 291–296.

Tekin, S., & Cummings, J. L. (2002). Frontal-subcortical neuronal circuits and clinical neuropsychiatry: An update. *Journal of Psychosomatic Research*, *53*, 647–654.

Vonsattel, J. P., & DiFiglia, M. (1998). Huntington disease. *Journal of Neuropathology and Experimental Neurology*, *57*, 369–384.

Vonsattel, J. P., Keller, C., & Pilar Amaya, M. D. (2008). Neuropathology of Huntington's disease. *Handbook of Clinical Neurology*, *89*, 599–618.

Walker, F. O. (2007). Huntington's disease. *Lancet*, *369 (9557)*, 218–228.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised: Manual*. San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S. (1993). *The Wide Range Achievement Test: Manual* (3rd ed.). Wilmington, DE: Wide Range.

Wilkinson, G. S., & Robertson, G. (2006). *WRAT4: Wide Range Achievement Test professional manual* (4th ed.). Lutz, FL: Psychological Assessment Resources.

Willshire, D., Kinsella, G., & Prior, M. (1991). Estimating WAIS-R IQ from the National Adult Reading Test: A cross-validation. *Journal of Clinical and Experimental Neuropsychology*, *13*, 204–216.