Test–Retest Reliability of Saccadic Measures in Subjects at Risk for Huntington Disease

Tanya Blekher,[1] Marjorie R. Weaver,[2] Xueya Cai,[3] Siu Hui,[3] Jeanine Marshall,[2] Jacqueline Gray Jackson,[2] Joanne Wojcieszek,[4] Robert D. Yee,[1] and Tatiana M. Foroud[2]

PURPOSE. Abnormalities in saccades appear to be sensitive and specific biomarkers in the prediagnostic stages of Huntington disease (HD). The goal of this study was to evaluate test–retest reliability of saccadic measures in prediagnostic carriers of the HD gene expansion (PDHD) and normal controls (NC).

METHODS. The study sample included 9 PDHD and 12 NC who completed two study visits within an approximate 1-month interval. At the first visit, all participants completed a uniform clinical evaluation. A high-resolution, video-based system was used to record eye movements during completion of a battery of visually guided, antisaccade, and memory-guided tasks. Latency, velocity, gain, and percentage of errors were quantified. Test–retest reliability was estimated by calculating the intraclass correlation (ICC) of the saccade measures collected at the first and second visits. In addition, an equality test based on Fisher's z-transformation was used to evaluate the effects of group (PDHD and NC) and the subject's sex on ICC.

RESULTS. The percentage of errors showed moderate to high reliability in the antisaccade and memory-guided tasks (ICC = 0.64–0.93). The latency of the saccades also demonstrated moderate to high reliability (ICC = 0.55–0.87) across all tasks. The velocity and gain of the saccades showed moderate reliability. The ICC was similar in the PDHD and NC groups. There was no significant effect of sex on the ICC.

CONCLUSIONS. Good reliability of saccadic latency and percentage of errors in both antisaccade and memory-guided tasks suggests that these measures could serve as biomarkers to evaluate progression in HD. (*Invest Ophthalmol Vis Sci.* 2009; 50:5707–5711) DOI:10.1167/iovs.09-3538

Saccades are rapid eye movements that bring an image from the periphery into the fovea. There is growing evidence to suggest that in several psychiatric[1–7] and neurodegenerative disorders, such as Huntington disease (HD),[8–9] abnormal saccades may be a sensitive prediagnostic biomarker in those at risk for disease.

HD is an autosomal dominant neurodegenerative disorder characterized by progressive deterioration of motor, cognitive, and psychiatric function. The disease-causing mutation is a trinucleotide (CAG) repeat expansion in the 5′-translated region of the huntingtin gene.[10] The average age of disease onset is 40 years, although onset has occurred as early as age 2 and as late as age 80. Disease onset is insidious, often with a long prediagnostic period with subtle symptoms. Typically, diagnosis is made only on the basis of unequivocal motor signs consistent with HD. Neurodegeneration in the striatum has been the focus of previous neuropathologic[11] and neuroimaging[12,13] studies; however, recent reports suggest the presence of abnormalities throughout the cerebrum, including cortical thinning[14] and decreased white matter volumes[15–17] especially in the prefrontal cortex.[18]

Significant efforts have been concentrated on the identification of sensitive biomarkers that can be used in the prediagnostic stage of HD to monitor disease progression and evaluate novel therapeutic agents. Studies of saccades in prediagnostic carriers of a huntingtin expansion (PDHD) have reported a significant deficit in volitional saccades elicited in response to a command.[8,9,19,20] The reaction time (latency) and error rates were significantly increased in persons with PDHD compared with normal control subjects (NC). The most pronounced deficits were observed in memory-guided saccades, suggesting that this deficit may be a promising biomarker for evaluation of progression in prediagnostic stages of HD.

Previous studies[21–25] have demonstrated good test–retest reliability of saccadic measures in healthy populations, as well as patients with psychiatric disorders, and their first-degree relatives. These studies have evaluated the reliability of saccades across a wide test–retest interval (from 2 weeks up to 33 months). Contrary to the results in healthy individuals and patients with schizophrenia, those with bipolar disorder did not show temporal stability in saccade measures.[26] There is some evidence to suggest that saccadic measures decline with the progression of the neurodegenerative disease over a long time interval.[8] However, no study has addressed short-term temporal stability of saccadic measures in patients with a neurodegenerative disorder. Such a determination appears essential for proposing a saccadic measure as a biomarker for monitoring prediagnostic stages of HD.

The goal of the present study was to evaluate the short-term temporal stability of saccadic measures in a sample of PDHD and NC subjects. We tested participants at two visits, separated by an approximate 1-month interval. We used these data to compute the test–retest reliability of saccadic measures for a wide range of testing paradigms. In addition, we examined the influence of sex and study group (PDHD or NC) on the temporal stability of saccadic measures.

METHODS

Subject Recruitment and Evaluation

PDHD participants were recruited through an ongoing study at Indiana University of individuals at risk for HD. Inclusion criteria included: a parent affected with HD; no diagnosis of HD; normal or corrected visual acuity; no history of eye surgery; and no current, significant, eye-related symptoms. All PDHD individuals had an expanded allele

(>38 CAG repeats; $n = 9$). NC ($n = 12$) subjects were recruited by advertisement, did not have a family history of HD, and were age matched to the PDHD individuals. They also had normal or corrected-to-normal visual acuity; no history of eye surgery; and no current, significant, eye-related symptoms. All participants completed two study visits within an approximate 1-month time interval. The study was conducted in accordance with the Declaration of Helsinki. The study was approved by the local institutional review board (IUPUI IRB Study No. 0109-12), and all participants provided written informed consent.

An experienced movement disorder neurologist (JW) administered the motor portion of the Unified Huntington's Disease Rating Scale[27] to all study participants. The neurologist was aware that some participants were at risk for HD, but was blinded to the results of all other study assessments, including the results of huntingtin gene testing. On the basis of the motor examination only, the neurologist assigned an overall confidence rating based on the likelihood that any observed abnormalities represented HD. The ratings were defined as: 0, normal (no abnormalities); 1, nonspecific motor abnormalities (<50% confidence); 2, motor abnormalities that may be signs of HD (50%–89% confidence), 3, motor abnormalities that are likely signs of HD (90% to 98% confidence), and 4, motor abnormalities that are unequivocal signs of HD (≥99% confidence). Potential subjects were excluded if they received an overall confidence rate indicating that the motor abnormalities were unequivocal signs of HD (rating of 4).

During the first visit, the participants completed questionnaires designed to collect medical history, substance use history, and visual health information. Participants did not report a psychiatric or neurodegenerative disorder diagnosis and by self-report were free from current alcohol or drug abuse.

Testing Procedure

At each of the two study visits, eye testing was performed in a 45-minute session, typically completed at approximately the same time of the morning. The vertical and horizontal positions of the participant's pupils were recorded binocularly with two ultraminiature, high-speed video cameras attached to a headband and digitized at 250 Hz for later analysis. Four sensors monitored head movements; eye positions were adjusted for small head movements (spatial resolution <0.1°; EyelinkII; SR Research Ltd., Mississauga, Ontario, Canada). Participants were seated in front of a 22-in. monitor in a standard ophthalmic examination chair. Visual targets (3-mm red spot with 1-mm black center) were displayed in a darkened room on a monitor placed 23.5-in. from the eye sensors. As part of the pretesting procedure, calibration and validation tests were completed. Five saccadic tasks were performed in a fixed order. Before each task, the examiner instructed the participant verbally and performed a practice to ensure that the participant understood the oral instructions correctly. Each of the five saccadic tasks consisted of 24 trials. In each trial the participant was instructed to fix his or her gaze on the target located in the center of a monitor (0°) and to redirect the gaze in response to the illumination of a peripheral target. Timing (2–3 seconds) and position of the peripheral spot (±7.5°, ±15° horizontally; ±10° vertically) were randomized.

Task 1: Visually-Guided (VG) Task. The participant was instructed to fixate on the target located in the center of a monitor (0°). The central target was extinguished simultaneously with the illumination of a peripheral target. The participant was instructed to visually track the target light as rapidly as possible.

Task 2: Antisaccade (AS) Task. The conditions of this task were similar to the VG task; however, the participant was instructed to look in the *opposite* direction of the peripheral target, at an equal distance from the center.

Task 3: Memory-Guided task—Version 1 (MG1). The participant was instructed to fix his or her gaze on the central target while a horizontal peripheral target flashed for 50 ms. The participant was asked to continue to fixate on the central target until it was switched off, after an additional delay of 1 to 2 seconds. Then, the participant was asked to look at the remembered position of the peripheral flash.

Task 4: Memory-Guided Task—Version 2 (MG2). This trial was a combination of the VG and memory-guided saccades. Initially, the participant was instructed to perform two sequential VG trials. In each VG trial, the participant was instructed to fixate on the target located in the center of a monitor (0°). The central target was extinguished simultaneously with the illumination of a peripheral target. After 1 to 2 seconds the peripheral target was extinguished simultaneously with the illumination of the central target. The participant was instructed to follow the light. After the two VG trials were completed, the participant was instructed to fixate on the central target until it was switched off after an additional delay of 1 to 2 seconds. Then, the participant was asked to repeat the two previous VG trials by memory (i.e., to look subsequently at the two remembered positions of the peripheral targets).

Task 5: Memory-Guided Task—Version 3 (MG3). This trial was similar to MG1 except that the flashes occurred sequentially on three peripheral positions, located both vertically and horizontally. The participant was instructed to look sequentially at the three remembered positions of the flashes.

We included three memory-guided tasks in our battery to evaluate whether the complexity of the task influences the test–retest reliability of saccadic measures. The MG1 task is considered less complex than the MG2 and MG3 tasks, since the MG2 and MG3 tasks include some element of sequencing.

Quantitative Measures of Saccades

An interactive computerized analysis[28] was performed to quantify the saccade measures for each task. Initially, the latency, peak velocity, and gain (i.e., ratio of saccade amplitude to target amplitude) were calculated for the first saccade in each correctly performed trial. Then, for each participant, the average value for these three measures (saccade latency, peak velocity of the idealized 15° saccade[29] and gain) were computed for each task. For the VG task, the horizontal and vertical saccades were analyzed separately; for other tasks, the vertical and horizontal saccades were combined. All participants made timing and directional errors in some of the AS, MG1, MG2, and MG3 trials. Particularly in the AS task, in some trials, the participant made the first saccade in the incorrect direction (toward the stimulus) and then looked in the opposite (correct) direction. In the MG1 trial, a typical timing error occurred when the participant initiated the first saccade before the central light was extinguished. The other possible error in the MG tasks was a missed peripheral flash. In this case, the participant did not perform a saccade to the remembered position of the flash (or a sequence of saccades for the MG2 and MG3 tasks). The percentage of trials with an error was computed for each individual for each task as the ratio of the number of trials performed with a mistake to the total number of trials completed in the task. Average latency, velocity, and gain were not calculated for the subjects who made errors in more than 75% of the trials in a particular task (i.e., made fewer than six correct saccades in the task).

Statistical Analysis

Our initial goal was to evaluate the test–retest reliability of saccadic measures (percentage of trials with errors and average latency, gain, and velocity) computed for each of the five tasks in each group (PDHD and NC). We performed point and interval estimations of the intraclass correlation coefficient (ICC) between the two visits for each measure and tested whether the ICC is significantly different from zero. Fisher's z-transformation was used for the confidence interval estimation. The ICC is the appropriate statistical method in reliability analysis for assessing temporal stability of the measures.[30] To compare our results with those from previous studies, we also computed the Pearson correlation coefficient.

TABLE 1. Demographic and Clinical Characteristics of the Two Study Groups

Demographic	NC ($n = 12$)	PDHD ($n = 9$)	P
Age, y (mean $\pm$ SD, range)	41 $\pm$ 14 (26–63)	45 $\pm$ 14 (26–63)	0.55
Education, y (mean $\pm$ SD)	15 $\pm$ 2	15 $\pm$ 3	0.94
Male/female ratio*	4/8	4/5	0.60
UHDRS motor assessment, confidence level (mean $\pm$ SD)*	0.7 $\pm$ 1.0	1.8 $\pm$ 0.8	0.02
Expanded CAG repeat	N/A†	42 $\pm$ 2	

* Male/Female ratio was evaluated by Fisher's exact test and, UHDRS confidence level was evaluated using the Cochran-Mantel-Haenszel (row mean scores) statistic. All other comparisons performed using T-test.

† N/A, not applicable, normal control subjects do not have a family history of HD and were not tested molecularly to determine the number of CAG repeats.

As a secondary analysis, for measures with significant ICC (ICC significantly different from 0), we tested whether there were significant group (PDHD and NC) or sex effects on the ICC. We performed the equality test of the ICCs based on Fisher's z-transformation.[31,32] A paired *t*-test was used for the comparison of saccadic measures between the first and second visits.

RESULTS

Demographic and clinical characteristics for the two study groups, PDHD and NC, are presented in Table 1. The range of the participant's age in both groups was 26 to 63 years. The groups did not differ significantly in age, education, or sex. As might be expected, the PDHD group had a higher mean diagnostic confidence score than the NC group had.

The MG2 task appeared to be unexpectedly complex for both the PDHD and the NC participants. Approximately 30% of the participants made errors in more than 75% of the trials. For these participants, we could not estimate the latency, gain, and velocity of correct saccades.

Test–Retest Reliability in the NC and PDHD Groups

There were no significant departures from normal for any of the saccadic measures. Table 2 displays the mean $\pm$ SD, as well as the ICC and results of the equality test, for the saccadic measures. We found moderate to high test–retest reliability for the percentage of errors in the AS and MG tasks (ICC = 0.63–0.88) in each group. Because there were so few trials with errors in the VG task, the ICC was not calculated. Figure 1A illustrates the strong association between the percentage of errors in the MG3 task at the first and second visits.

We also found moderate to high test–retest reliability of saccadic latency in the VG, AS, MG1, and MG3 tasks (ICC =

TABLE 2. Summary Statistics for the Five Saccadic Paradigms

	Saccade Measure*				Intraclass Correlation					
	PDHD Group		NC Group		PDHD Group		NC Group		Equality Test	
	1st Visit	2nd Visit	1st Visit	2nd Visit	ICC	CI	ICC	CI	|T|	P
Percentage of errors										
AS	37 $\pm$ 28	32 $\pm$ 30	31 $\pm$ 25	28 $\pm$ 26	0.83	0.43, 0.96	0.93	0.77, 0.98	0.95	0.34
MG1	27 $\pm$ 16	23 $\pm$ 14	25 $\pm$ 21	29 $\pm$ 24	0.67	0.10, 0.91	0.88	0.61, 0.97	1.12	0.26
MG2	41 $\pm$ 30	41 $\pm$ 38	26 $\pm$ 21	38 $\pm$ 38	0.84	0.46, 0.96	0.63	0.14, 0.87	0.97	0.33
MG3	38 $\pm$ 21	28 $\pm$ 14	36 $\pm$ 27	35 $\pm$ 29	0.64	0.04, 0.90	0.88	0.64, 0.96	1.27	0.20
Latency										
VG, horizontal	179 $\pm$ 31	185 $\pm$ 36	166 $\pm$ 13	166 $\pm$ 19	0.87	0.55, 0.97	0.71	0.28, 0.90	0.93	0.35
VG, vertical	184 $\pm$ 30	180 $\pm$ 29	177 $\pm$ 26	182 $\pm$ 18	0.89	0.62, 0.97	0.77	0.34, 0.93	0.82	0.41
AS	328 $\pm$ 43	313 $\pm$ 61	275 $\pm$ 21	257 $\pm$ 26	0.83	0.39, 0.96	0.68	0.19, 0.90	0.70	0.49
MG1	349 $\pm$ 56	321 $\pm$ 63	290 $\pm$ 34	280 $\pm$ 42	0.74	0.24, 0.93	0.62	0.06, 0.89	0.44	0.66
MG3	318 $\pm$ 60	313 $\pm$ 37	281 $\pm$ 46	271 $\pm$ 41	0.55	0, 0.87	0.58	0, 0.87	0.09	0.93
Gain										
VG, horizontal	0.96 $\pm$ 0.09	1.0 $\pm$ 0.07	0.94 $\pm$ 0.08	0.93 $\pm$ 0.12	0.20†	0, 0.72	0.59	0.07, 0.86	0.99	0.32
VG, vertical†	0.88 $\pm$ 0.12	0.93 $\pm$ 0.15	0.90 $\pm$ 0.17	0.91 $\pm$ 0.09						
AS	0.73 $\pm$ 0.36	0.83 $\pm$ 0.31	1.03 $\pm$ 0.29	1.02 $\pm$ 0.35	0.82	0.38, 0.96	0.83	0.50, 0.95	0.04	0.97
MG1	0.87 $\pm$ 0.17	0.96 $\pm$ 0.09	0.87 $\pm$ 0.10	0.85 $\pm$ 0.13	0.76	0.26, 0.94	0.52	0, 0.85	0.82	0.41
MG3†	0.88 $\pm$ 0.15	0.91 $\pm$ 0.11	0.86 $\pm$ 0.17	0.86 $\pm$ 0.11						
Velocity										
VG, horizontal	384 $\pm$ 69	405 $\pm$ 36	404 $\pm$ 54	411 $\pm$ 71	0.43	0, 0.83	0.44	0, 0.79	0.01	0.99
VG, vertical†	388 $\pm$ 90	407 $\pm$ 54	393 $\pm$ 124	392 $\pm$ 98						
AS†	338 $\pm$ 90	348 $\pm$ 43	348 $\pm$ 44	336 $\pm$ 53						
MG1†	313 $\pm$ 65	360 $\pm$ 45	332 $\pm$ 66	357 $\pm$ 95	0.57	0, 0.88	0.58	0, 0.87	0.03	0.97
MG3†	298 $\pm$ 81	330 $\pm$ 59	361 $\pm$ 99	331 $\pm$ 56						

* Mean $\pm$ SD.
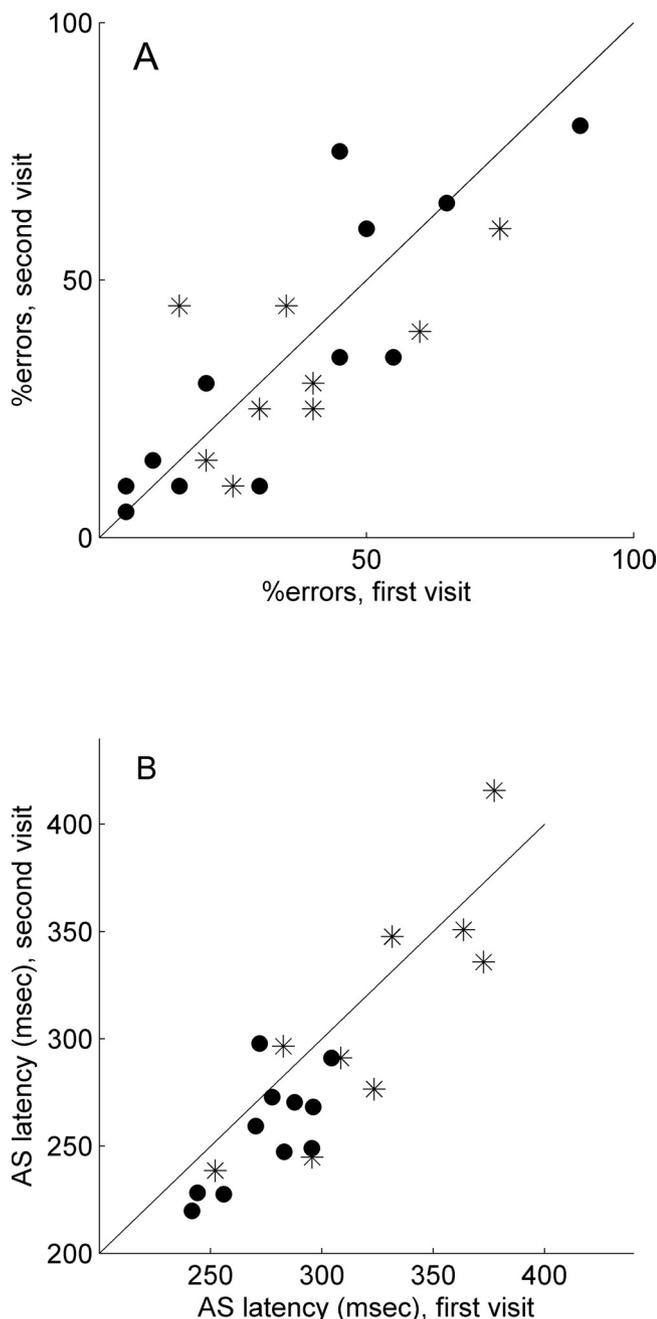† ICC was not significant.

FIGURE 1.   (A) Percentage of errors in the MG3 task in the first visit versus second visit. (B) Latency of the AS task in the first visit versus second visit. *Circles:* NC; *stars:* PDHD.

Pearson Correlation

When computing the Pearson correlation coefficient, we found the same pattern as with the ICC for the measures of percentage of errors, latency, velocity, and gain across all five tasks. As expected, the Pearson correlation coefficient was typically slightly higher or equal to the ICC.

Between-Session Practice Effect

Comparisons of saccadic measures between the first and second visits showed that there were minor, but significant, improvements in the AS task. The percentage of errors and latency decreased ($P = 0.01$) in the second visit.

DISCUSSION

Our findings suggest that saccadic performance is stable across a 1-month interval in a wide range of testing paradigms in both NC and PDHD individuals. The 1-month test–retest reliability was excellent for the percentage of trials with errors and for latency of volitional saccades. Those measures are found to be abnormal in the early stages of HD and may be potential biomarkers for prediagnostic HD.[8,9] Between visits, minor practice effects were observed only for the AS task, as indicated by the reduced percentage of errors and decreased latency. Overall, our results are in good agreement with those reported previously in other samples, including healthy control subjects, patients with schizophrenia, and their relatives.[21–26]

Our results clearly demonstrate that individuals with prediagnostic HD had excellent reliability for the latency and percentage of trials with errors across a wide range of tasks. However, because those with advanced HD demonstrated a wide range of oculomotor control abnormalities, reliability for these study measures must be confirmed separately for this group.

In conclusion, our results imply a temporal stability of saccadic measures in persons with prediagnostic HD. Excellent reliability of saccadic latency and percentage of errors in both AS and MG tasks suggests that these measures could serve as potential biomarkers to evaluate the efficacy of neuroprotective agents in slowing or delaying the progression of HD.

References

1. Amado I, Landgraf S, Bourdel MC, Leonardi S, Krebs MO. Predictive saccades are impaired in biological nonpsychotic siblings of schizophrenia patients. *J Psychiatry Neurosci.* 2008;33:17–22.
2. Calkins ME, Curtis CE, Iacono WG, Grove WM. Antisaccade performance is impaired in medically and psychiatrically healthy biological relatives of schizophrenia patients. *Schizophr Res.* 2004;71:167–178.
3. Crawford TJ, Sharma T, Puri BK, Murray RM, Berridge DM, Lewis SW. Saccadic eye movements in families multiply affected with schizophrenia: the Maudsley Family Study. *Am J Psychiatry.* 1998;155:1703–1710.
4. Curtis CE, Calkins ME, Grove WM, Feil KJ, Iacono WG. Saccadic disinhibition in patients with acute and remitted schizophrenia and their first-degree biological relatives. *Am J Psychiatry.* 2001;158:100–106.
5. Curtis CE, Calkins ME, Iacono WG. Saccadic disinhibition in schizophrenia patients and their first-degree biological relatives: a parametric study of the effects of increasing inhibitory load. *Exp Brain Res.* 2001;137:228–236.
6. McDowell JE, Myles-Worsley M, Coon H, Byerley W, Clementz BA. Measuring liability for schizophrenia using optimized antisaccade stimulus parameters. *Psychophysiology.* 1999;36:138–141.
7. Nieman D, Becker H, van de Fliert R, et al. Antisaccade task performance in patients at ultra high risk for developing psychosis. *Schizophr Res.* 2007;95:54–60.

0.55–0.89, Table 2). Figure 1B illustrates the linear relationship between AS latency at the first and second visits.

The ICC showed moderate to high agreement for gain in AS and MG1 tasks (ICC = 0.52–0.83). The ICC for velocity of the saccades also demonstrated moderate agreement (ICC = 0.43–0.58). For most saccadic measures in Table 2, the ICCs were significantly different from zero.

The equality test showed no significant differences between the ICC in the NC and PDHD groups for percentage of errors, latency, velocity, and gain. There were also no significant differences between the ICC in the men and women for any saccadic measures, suggesting that the two sexes showed similar test–retest reliability.

8. Blekher T, Johnson SA, Marshall J, et al. Saccades in presymptomatic and early stages of Huntington disease. *Neurology.* 2006;67: 394–399.

9. Golding CV, Danchaivijitr C, Hodgson TL, Tabrizi SJ, Kennard C. Identification of an oculomotor biomarker of preclinical Huntington disease. *Neurology.* 2006;67:485–487.

10. Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell.* 1993;72: 971–983.

11. Vonsattel JP, Myers RH, Stevens TJ, Ferrante RJ, Bird ED, Richardson EP. Neuropathological classification of Huntington's disease. *J Neuropathol Exp Neurol.* 1985;44:559–577.

12. Aylward EH, Li Q, Stine OC, et al. Longitudinal change in basal ganglia volume in patients with Huntington's disease. *Neurology.* 1997;48:394–399.

13. Aylward EH, Codori AM, Rosenblatt A, et al. Rate of caudate atrophy in presymptomatic and symptomatic stages of Huntington's disease. *Mov Disord.* 2000;15:552–560.

14. Rosas HD, Liu AK, Hersch S, et al. Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology.* 2002;58:695–701.

15. Beglinger LJ, Nopoulos PC, Jorge RE, et al. White matter volume and cognitive dysfunction in early Huntington's disease. *Cogn Behav Neurol.* 2005;18:102–107.

16. Paulsen JS, Magnotta VA, Mikos AE, et al. Brain structure in preclinical Huntington's disease. *Biol Psychiatry.* 2006;59:57–63.

17. Rosas HD, Koroshetz WJ, Chen YI, et al. Evidence for more widespread cerebral pathology in early HD: an MRI-based morphometric analysis. *Neurology.* 2003;60:1615–1620.

18. Reading SA, Yassa MA, Bakker A, et al. Regional white matter change in pre-symptomatic Huntington's disease: a diffusion tensor imaging study. *Psychiatry Res.* 2005;140:55–62.

19. Hicks SL, Robert MP, Golding CV, Tabrizi SJ, Kennard C. Oculomotor deficits indicate the progression of Huntington's disease. *Prog Brain Res.* 2008;171:555–558.

20. Kirkwood SC, Siemers E, Hodes ME, Conneally PM, Christian JC, Foroud T. Subtle changes among presymptomatic carriers of the Huntington's disease gene. *J Neurol Neurosurg Psychiatry.* 2000; 6:773–779.

21. Calkins ME, Iacono WG, Curtis CE. Smooth pursuit and antisaccade performance evidence trait stability in schizophrenia patients and their relatives. *Int J Psychophysiol.* 2003;49:139–146.

22. Ettinger U, Kumari V, Crawford TJ, Davis RE, Sharma T, Corr PJ. Reliability of smooth pursuit, fixation, and saccadic eye movements. *Psychophysiology.* 2003;40:620–628.

23. Klein C, Berg P. Four-week test-retest stability of individual differences in the saccadic CNV, two saccadic task parameters, and selected neuropsychological tests. *Psychophysiology.* 2001;38: 704–711.

24. Klein C, Fischer B. Instrumental and test-retest reliability of saccadic measures. *Biol Psychol.* 2005;68:201–213.

25. Roy-Byrne P, Radant A, Wingerson D, Cowley DS. Human oculomotor function: reliability and diurnal variation. *Biol Psychiatry.* 1995;3:92–97.

26. Gooding DC, Mohapatra L, Shea HB. Temporal stability of saccadic task performance in schizophrenia and bipolar patients. *Psychol Med.* 2004;34:921–932.

27. Huntington Study Group. Unified Huntington's disease rating scale: Reliability and consistency. *Mov Disord.* 1996;11:136–142.

28. Blekher TM, Yee RD, Kirkwood SC, et al. Oculomotor control in asymptomatic and recently diagnosed individuals with the genetic marker for Huntington's disease. *Vision Res.* 2004;44:2729–2736.

29. Bahill AT, Clark MR, Stark L. The main sequence, a tool for studying human eye movements. *Math Biosci.* 1975;24:191–204.

30. Bartko JJ. Measurement and reliability: statistical thinking considerations. *Schizophr Bull.* 1991;17:483–489.

31. Konishi S. Normalizing and variance stabilizing transformations for intraclass correlations. *Ann Inst Stat Math.* 1985;37:87–94.

32. Konishi S, Gupta AK. Testing the quality of several intraclass correlation coefficients. *J Stat Plan Infer.* 1989;21:91–105.