

Automated and objective analysis of speech in premanifest and early-stage Huntington's disease

Adam P. Vogel PhD^{1,2,3}, Cheuk S. J. Chan PhD¹, Geoffrey W. Stuart PhD^{3,4,5}, Paul Maruff PhD^{6,7}, Yenni Lie MBBS⁸, Julie C. Stout PhD^{9,10}

1 Centre for Neuroscience of Speech, The University of Melbourne, Australia

2 Division of Translational Genomics of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Germany & Center for Neurology, University Hospital Tübingen, Germany

3 Redenlab Inc., Melbourne, Australia

4. Department of Cognitive Science, Macquarie University, Sydney Australia

5. School of Psychological Sciences, The University of Melbourne, Australia

6 Cogstate Inc., Melbourne Australia

7 Florey Institute of Neuroscience and Mental Health

8 Statewide Progressive Neurological Disease Service, Calvary Health Care Bethlehem, Melbourne Australia

9 School of Psychological Sciences, Monash University, Melbourne Australia

10 Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Victoria, Australia

ORCID:

Adam Vogel <https://orcid.org/0000-0002-3505-2631>

Julie Stout <https://orcid.org/0000-0001-9637-8590>

Geoffrey Stuart <https://orcid.org/0000-0001-7142-9822>

Yenni Lie <https://orcid.org/0000-0002-6758-6435>

Corresponding author: Adam Vogel PhD

550 Swanston Street, Parkville, VIC Australia 3010

vogela@unimelb.edu.au

Word count: 2750

Running title: Motor speech in premanifest HD

Financial disclosures and conflicts of interest:

A. P. Vogel is Chief Science Officer of Redenlab Inc.

C. S. J. Chan reports no disclosures.

G. W. Stuart is Director of Statistics for Redenlab Inc.

Y. Lie reports no disclosures.

P. Maruff is Chief Innovation Officer of Cogstate Inc.

J. Stout is Director of Zindamatrix Pty Ltd.

Funding Sources for Study:

A. P. Vogel received salaried support from the National Health and Medical Research Council, Australia (#1082910) and institutional support from The University of Melbourne.

J. Stout is funded by a National Health and Medical Research Council, Australia, Investigator Grant (#1173472).

Abstract

Background: Clinical markers that show change in performance in people with Huntington’s disease (HD) during the presymptomatic and prodromal stages remain a target of investigation in clinical medicine. Alongside genetic and neuroimaging initiatives, digital speech analytics has shown promise as a sensitive clinical marker of premanifest HD.

Objective: To investigate the sensitivity of digital speech measures for detecting subtle cognitive-linguistic and fine motor features in people carrying the expanded HD gene, with and without symptoms.

Methods: Speech data were acquired from 110 participants (55 people with the expanded HD gene including 16 presymptomatic HD; 16 prodromal HD; 14 early-stage HD; 9 mid-stage HD; and 55 matched healthy controls). Objective digital speech measures were derived from speech tasks that fit along a continuum of motor and cognitive complexity. Acoustic features quantified speakers’ articulatory agility, voice quality and speech-timing. Subjects also completed the tests of cognition and upper limb motor function.

Results: Some presymptomatic HD (furthest from disease onset) differed to healthy controls on timing measures derived from the syllable repetition and monologue. Prodromal HD presented with reduced articulatory agility, reduced speech rate and longer and variable pauses. Speech agility correlated with poorer performance on the upper limb motor test.

Conclusion: Tasks with a mix of cognitive and motor demands differentiated prodromal HD from their matched control groups. Motor speech tasks alone did not differentiate groups until participants were relatively closer to disease onset or symptomatic. Data demonstrated how ubiquitous behaviors like speech, when analyzed objectively, provide insight into disease related change.

Introduction

Huntington's disease (HD) is an autosomal dominantly inherited neurodegenerative disease caused by an unstable expansion in the number of CAG trinucleotide repeats in the 1T15 huntingtin (*HTT*) gene^{1,2}. Its dominant inheritance means that prediagnostic genetic testing can be used to identify people at risk of developing HD prior to reaching diagnostic certainty³. Investigation of both biological and symptomatic changes in people carrying the repeat expansion for *HTT* provides a basis for development of brain behavior models of early HD. During this stage, termed the premanifest stage⁴, changes may be observed across cognitive, behavioral, fine motor and/or speech functions in HD CAG expansion, potentially years or decades before the emergence of diagnostic motor symptoms⁵⁻⁷. A clinical diagnosis of HD is only made when a person carrying the expanded HD gene presents with unequivocal motor symptoms⁴.

The genetic predictability of HD also provides an opportunity for introducing preventative or curative genetic and pharmaceutical therapies⁸. One of the anticipated challenges of clinical trials in this premanifest stage is the need for sensitive and reliable markers of disease expression and response to treatment⁹⁻¹¹. Once people are symptomatic (manifest), HD impacts cognitive, linguistic, and motor function^{7, 12-15}. Qualitatively similar but quantitatively less severe changes in some of these symptoms have been reported in the prodromal period¹⁶⁻¹⁹. There are, however, conflicting reports on the nature of speech changes occurring prior to an HD diagnosis^{7, 20}. Further, the relationship of fine motor function to speech in premanifest HD remains largely unknown.

Here we explored the use of speech analytics to provide continuous, quantitative data on disease status in HD. The aim was to investigate whether speech differed to healthy controls in premanifest (presymptomatic, prodromal) or manifest (early, mid) HD groups and to explore the relationship between these features and changes in fine motor control.

Methods

Participants

Fifty-five participants carrying the expanded huntingtin (*HTT*) gene and 55 age and sex matched controls were recruited from the ENROLL-HD registry at Calvary Health Care Bethlehem Hospital and a Monash University HD research volunteer database in Melbourne, Australia. All participants in the HD groups had ≥ 38 CAG repeats in one of the HD alleles. Participants received a standard neurological exam by a neurologist using the Unified Huntington's Disease Rating Scale (UHDRS) yielding a Total Motor Score and a Diagnostic

Confidence Level ^{4, 21}. Potential HD participants were excluded if they presented with any comorbid neurological diseases that could affect performance on study measures (e.g., stroke, multiple sclerosis); clinical symptoms other than those caused by HD; a history of communication disorder (e.g., stuttering); lack premorbid competency in English; a history of alcohol or drug abuse that required medical intervention; or a history of learning disability and/or intellectual impairment. Eligibility for inclusion as a healthy control required having no family history of HD, or other major neurological disease, and had unremarkable cognition, speech, or language function. English was the first language for all participants.

Disease burden scores (DBS) were calculated for all participants with the expanded HD gene (DBS = age*[CAG repeat -35.5]) ²². Years to disease onset was estimated for participants with premanifest HD ²³. Participants were allocated to one of four groups, based on proposed criteria from Ross et al., 2019: premanifest (presymptomatic, prodromal) or manifest (early, mid) groups (Table 1). Participants were included in the presymptomatic group if they were allocated a DCL of 0 or 1 and presented with an absence of motor and cognitive symptoms related to HD. Participants were allocated to the prodromal group if they presented with subtle motor signs, with or without minor/major neurocognitive changes, which corresponded with a DCL of 2 or 3 (Ross et al., 2019). Participants with a DCL of 4 were classified using the UHDRS Total Functional Capacity Rating Scale (TFC) scores as early HD (TFC \geq 7) or mid-stage HD (TFC 4-6) ²⁴. There were four control groups, each matched to the sex and age of the corresponding experimental group (see Table 2 for clinical and demographic information and Supplementary Materials).

INSERT TABLE 1 ABOUT HERE

Table 1: Participant group descriptions and inclusion criteria of HD groups

INSERT TABLE 2 ABOUT HERE

Table 2: Demographic, clinical, and genetic characteristics of participants with the expanded HD gene (Mean +/- SD unless specified)

Materials and Stimuli

We recorded speech in a quiet room using a standard laptop computer, a cardioid condenser (head-mounted) microphone (AKG520, Harman International Industries), coupled

with an audio interface (Quad-Capture USB 2.0, Roland, Tokyo, Japan). Speech tasks included (i) sustained vowel /a:/, (ii) syllable repetition (DDK) (/papa/, /pata/), (iii) saying days of the week, (iv) reading a phonetically balanced passage from “The North Wind and The Sun”, and (v) an unprepared monologue for approximately one minute. Tasks varied along a spectrum of cognitive-motor complexity^{10, 25, 26}. All speech tasks except the monologue were repeated to minimize the impact of unfamiliarity^{27, 28}.

Acoustic analysis of speech

Objective measures of timing and voice quality were derived from samples using acoustic analysis. Timing metrics relating to syllabic rate, silence length and their variability were derived from monologue, reading, days of the week and syllable repetition tasks as described previously^{10, 29}. Voice quality measures focused on (i) vocal fold vibration patterns, (ii) incomplete vocal fold closure, and (iii) placement of the articulators during vowel phonation were derived from the vowel³⁰. Detailed descriptions of each acoustic measure employed in this study are described in Supplementary materials Table S1.

Cognitive and upper limb fine motor assessments

The Montreal Cognitive Assessment (MoCA)³¹ and the Cogstate Brief Battery (CBB) were used to measure cognitive function. The MoCA is used as a cognitive screener with healthy controls scoring $\geq 26/30$. The CBB battery assesses processing speed, attention/vigilance, working memory, and learning³², and has demonstrated sensitivity in other neurodegenerative conditions³³. The Purdue pegboard test (PPT) was administered to all participants to assess fine motor performance (Model 32020, Lafayette Instrument Co., Lafayette, Indiana)³⁴. The PPT requires participants to insert metal pegs into a vertical row of holes using their dominant and non-dominant hands separately, proceeded by using both hands simultaneously and alternating hands.

Statistical analysis

Statistical analyses were performed using Stata statistical software, version 17. All acoustic data extracted from premanifest, and manifest HD groups were compared to control data from this study. Due to violations of normality, with the HD group data distributions tending to have longer tails (see Figure 2 for examples), a nonparametric dominance statistic was used to compare groups³⁵. This statistic quantifies the probability that a randomly chosen participant from the HD group will have a worse score on a given measure than a randomly chosen participant from the control group (assessed across all possible pairs), with a value of 0.5 representing the null hypothesis. Significance tests used exact probability

estimation. Confidence intervals were constructed using bootstrap resampling using bias-corrected and accelerated estimation³⁶ as recommended here³⁷. False discovery rate adjustment, assuming positive regression dependency, was used to adjust for multiple comparisons³⁸. To explore multivariate contrasts between groups, a small subset of measures that showed the best discrimination between HD and control groups was selected for further analysis. To avoid overfitting and to deal with correlations between the measures, elasticnet regression³⁹ was employed to construct composite scores for multivariate discrimination using logistic regression. Pearson correlation coefficients were also used to explore the relationships between clinical variables and speech outcomes.

Multivariate Analysis

To evaluate how speech and motor measures might combine to discriminate HD gene carriers from controls in the pre-symptomatic and prodromal groups, we first selected key measures from the univariate analysis shown in Figure 1. This was to prevent overfitting when performing multivariate analysis, given the limited sample sizes. Two composite scores were developed using elasticnet logistic regression. The first included univariate measures of speech and motor for analysis, and the second included speech measures only. Motor performance, independent of speech, was assessed using Purdue pegboard right/left/both measures.

For the prodromal group, the elasticnet-derived composite score for speech measures was derived from the standardized scores from several speech tasks. The general composite score consisted of these same speech measures, plus performance on the Purdue pegboard test. These equally weighted composite scores were highly correlated with the exactly weighted scores from elasticnet regression (speech: $r=0.976$; general: $r=.994$). In the case of the prodromal group, the exact weights from elasticnet regression were far from equal, and so equal weights were not used. The speech composite score consisted of the measures derived from several tasks. The general composite score was derived from of the measures Purdue pegboard, and speech tasks. These composite measures were also highly correlated with exact regression scores (speech: $r=.989$; general: $r=.957$).

Results

Measures of speech timing and voice quality differed between HD participants and matched study controls, with the size of the effect increasing with disease severity (Figure 1). The largest differences between HD and controls were observed in manifest cohorts (premanifest < early < mid). Acoustic measures of timing derived from the monologue and

syllable repetition tasks differentiated premanifest participants from matched controls. Voice quality was different between manifest and matched study participants, but not premanifest stages. As shown in Figure 2, the composite measures improved discrimination between all HD (premanifest and manifest) participants and matched study controls relative to the single measures.

INSERT FIGURE 1 ABOUT HERE

Figure 1: Standardized comparison of speech in premanifest and manifest HD to matched controls

Speech in premanifest HD

No statistically significant differences (q-values $<.05$, representing the false discovery rate associated with each measure) were observed between presymptomatic and study control participants. Timing measures derived from the monologue and syllable repetition tasks revealed some presymptomatic participants did not fit within the healthy control range, appearing as outliers on several speech features (see Figure 2 and Table 3). The prodromal group differed to study controls on timing measures drawn from the monologue and reading tasks as well as for the Purdue motor task.

Speech in manifest HD

Speech in Early manifest HD was characterized by a slower and more inconsistent syllable repetition rate, increased irregularity in vocal fold vibration patterns (RPDE: $q=0.001$) during vowel phonation, and speech-timing deficits in complex speech tasks (reading: articulation rate: $q = 0.01$; percentage of silence: $q=0.007$ | monologue: mean pause duration: $q=0.001$; pause duration variability: $q=0.028$). Participants in the mid-stage manifest HD group presented with near global deficits across voice quality (jitter %: $q=0.007$), breath support (maximum phonation time: $q=0.044$) and measures of timing across complex tasks.

INSERT FIGURE 2 ABOUT HERE

Figure 2: Standardized participant performance across core speech tasks compared to matched controls in presymptomatic and prodromal HD

Premanifest outlier performance across assessment domains

Figure 2 shows premanifest participants appearing as outliers relative to the control distributions across assessment domains. We examined whether these outlier participants were represented across different tests (Table 3). There appears to be a dissociation between the speech measures and the purely motor measure (Purdue). In the Presymptomatic group, four participants had deficits across individual speech measures and but not for the Purdue pegboard. One shows a specific deficit on the Purdue pegboard only. For the Prodromal group, most outliers presented with both speech and Purdue pegboard tasks. On participant (42) is a “multivariate outlier” (not an outlier on any individual measure) and two participants were outliers on individual speech measures but not the composite.

INSERT TABLE 3 ABOUT HERE

Discussion

An automated speech analysis was used to separate premanifest and manifest HD from age and sex matched participants. Timing metrics derived from complex connected speech tasks identified healthy adults carrying the HD CAG expansion but not yet presenting with overt diagnostic signs. In some cases, abnormal speech timing was accompanied by pure motor deficits, with speech measures outperforming the pegboard, representing 85% of outlier instances in the presymptomatic group. Relative to matched healthy adults, participants in the premanifest groups demonstrated speech timing deficits during cognitively loaded tasks like reading and unprepared monologues, as well as the syllable repetition task, but not during recitation tasks like days of the week or sustained vowel tasks. The overall size of deficit relative to controls increased in the manifest groups, with the largest impact of dysarthria exhibited in the mid stage manifest HD group.

Two-thirds of presymptomatic HD participants produced speech at a similar rate to healthy adults, with the remaining third exhibiting longer pauses and slower syllabic rate than anticipated. A larger proportion of the prodromal group produced speech outside the normal range on at least one speech measure. Like presymptomatic HD, but to a larger extent, the prodromal HD group produced slower speech, with longer pauses across the monologue and syllable repetition tasks. While speech deficits in prodromal HD have been reported^{7, 16-18, 40}, our study is the first to successfully capture them in premanifest HD using an automated speech analytics pipeline.

The purely manual motor (Purdue) and speech measures both contributed important

information to separating groups despite being moderately associated. At an individual level, several premanifest participants showed the emergence of HD motor signs in one or both measures suggesting that a single participant level analysis adds critical insights to the group level analyses. At a group level the shift in the set of scores was not uniform. Yet, some participants with the HD CAG expansion performed well outside the normal range in the presymptomatic and prodromal phases.

Speech measures such as articulation rate and mean pause duration on the monologue task best classified premanifest individuals from controls. The unprepared monologue is a complex communication task requiring rapid formulation of novel language and contemporaneous motor planning across multiple speech systems (i.e., respiration, phonation, prosody, resonance, articulation)²⁵. Unlike the recitation (days of the week) and syllable repetition tasks, which require relatively little linguistic planning, the monologue draws on cognitive domains like working memory, executive function, and lexical-semantic access. These additional demands on cognitive-linguistic performance perhaps contribute to the sensitivity of the monologue to early disease, similar to other cognitive/motor disorders like Fronto-temporal dementia²⁶ and spinocerebellar ataxia¹⁰.

A composite of speech measures derived from the monologue and other tasks added to overall discriminability between groups. Like a listener's ability to perceive an overall deficit based on a speaker's communication in conversation, composite measures can encompass features across multiple speech subsystems, capturing performance beyond isolated and potentially esoteric acoustic features^{41, 42}. Unlike listener ratings however, digital acoustic analysis of speech is objective and is repeatable irrespective of clinical experience or training, making it better suited to monitoring change resulting from disease progression or treatment response²⁸.

Voice quality did not reliably separate HD groups from controls until individuals became symptomatic (i.e., early HD), similar to some earlier work⁴³ but contrary to other earlier reports¹⁷. These conflicting outcomes are likely the result of different recording and analysis paradigms, or cohort characteristics of the premanifest HD population. Earlier studies would also not have differentiated between presymptomatic and prodromal stages when describing premanifest HD, given that this is new terminology in the HD field.

Limitations and future work

Despite being the first study to explore speech in two tightly phenotyped premanifest groups alongside two symptomatic cohorts of HD, our sample size is modest. Replication and

extension of our methodology could explore the use of home-based assessments and consortia-based data collection to increase cohort size and statistical power. The progression of HD and evolving speech phenotype across each disease stage suggests that composite measures need to be tailored to individual characteristics and severity. We have proposed a composite feature best suited to describing premanifest participants however these properties may change given that all speech systems are impacted in early to mid HD. Future research could also align structural and functional brain imaging, neurofilament light (NfL) concentrations⁴⁴, associated behavioral data (e.g., neuropsychological) and the role of medications (e.g., antipsychotic) with detailed speech outcomes to better describe the underlying mechanisms preceding or causing the HD dysarthria profile.

Conclusion

Automated digital speech tests have potential to discriminate premanifest HD from controls at the single-person level. These changes appear to worsen as disease progresses. Digital markers that draw on multisystem behaviors (e.g., respiration, muscular control, cognitive-linguistic performance) to produce composite measures appear to outperform single feature metrics for differentiating disease group in HD. The ease of elicitation and administration of speech tests in a clinical trial setting compared to other modes of assessment (e.g., pegboard tests) highlights their potential for use beyond disease characterization.

Authors' Roles

Design: APV, CSJC, JS

Execution: APV, CSJC

Analysis: APV, CSJC, GWS, YL, PM, JS

Writing: APV, CSJC

Editing of final version of the manuscript: APV, CSJC, GWS, YL, PM, JS

Statistical Analysis was conducted by Geoffrey W. Stuart PhD, Redenlab Inc, Victoria, Australia

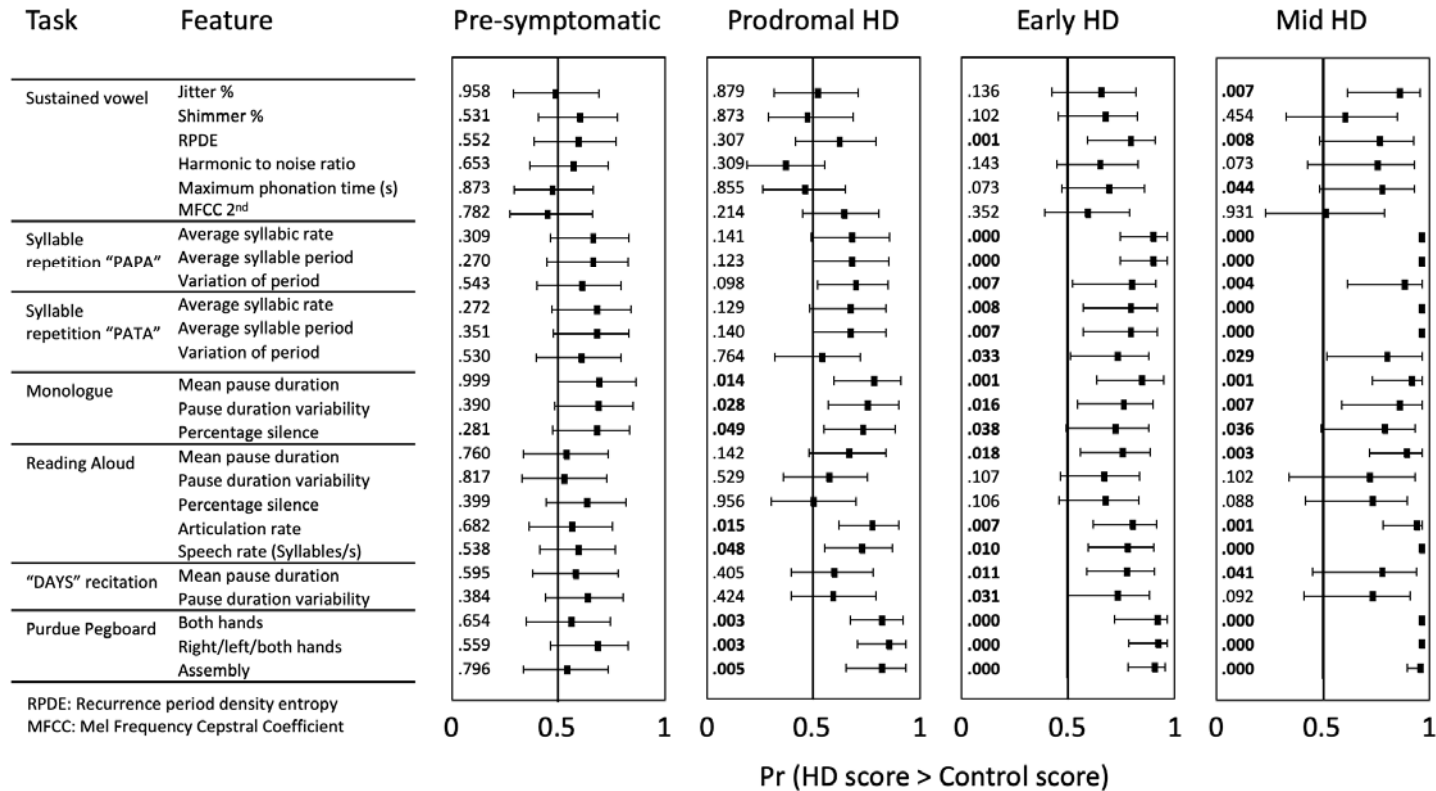
References

1. The 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.
2. Kremer B, Goldberg P, Andrew SE, et al. A worldwide study of the Huntington's disease mutation: The sensitivity and specificity of measuring CAG repeats. *The New England Journal of Medicine* 1994;330(20):1401-1406.
3. Myers RH. Huntington's Disease Genetics. *The Journal of the American Society for Experimental NeuroTherapeutics* 2004;1:255-262.
4. Ross CA, Reilmann R, Cardoso F, et al. Movement Disorder Society Task Force Viewpoint: Huntington's Disease Diagnostic Categories. *Movement disorders clinical practice* 2019;6(7):541-546.
5. Paulsen JS, Miller AC, Hayes T, Shaw E. Cognitive and behavioral changes in Huntington disease before diagnosis. *Handb Clin Neurol* 2017;144:69-91.
6. Harrington DL, Rubinov M, Durgerian S, et al. Network topology and functional connectivity disturbances precede the onset of Huntington's disease. *Brain* 2015;138(Pt 8):2332-2346.
7. Chan JCS, Stout JC, Vogel AP. Speech in prodromal and symptomatic Huntington's disease as a model of measuring onset and progression in dominantly inherited neurodegenerative diseases. *Neuroscience & Biobehavioral Reviews* 2019;107:450-460.
8. Ross CA, Aylward EH, Wild EJ, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nature Reviews Neurology* 2014;10(4):204-216.
9. Killoran A, Biglan K. Biomarkers for Huntington's disease: A brief overview. *Journal of Rare Diseases Research & Treatment* 2016;1(2):46-50.
10. Vogel AP, Magee M, Torres-Vega R, et al. Features of speech and swallowing dysfunction in pre-ataxic spinocerebellar ataxia type 2. *Neurology* 2020;95(2):e194-e205.
11. Weir DW, Sturrock A, Leavitt BR. Development of biomarkers for Huntington's disease. *The Lancet Neurology* 2011;10(6):573-590.
12. Ruz J, Klempř J, Baborová E, et al. Objective acoustic quantification of phonatory dysfunction in Huntington's disease. *PLoS One* 2013;8(6):e65881-e65881.
13. Hartelius L, Carlstedt A, Ytterberg M, Lillvik M, Laakso K. Speech disorders in mild and moderate Huntington disease: results of dysarthria assessments of 19 individuals. *J Med Speech Lang Pathol* 2003;11(1):1-14.
14. Hinzen W, Rosselló J, Morey C, et al. A systematic linguistic profile of spontaneous narrative speech in pre-symptomatic and early stage Huntington's disease. *Cortex* 2018;100:71-83.
15. Diehl SK, Mefferd AS, Lin Y-C, et al. Motor speech patterns in Huntington disease. *Neurology* 2019;93(22):e2042-e2052.
16. Skodda S, Grönheit W, Lukas C, et al. Two different phenomena in basic motor speech performance in premanifest Huntington disease. *Neurology* 2016.
17. Ruz J, Saft C, Schlegel U, Hoffman R, Skodda S. Phonatory Dysfunction as a Preclinical Symptom of Huntington Disease. *PLoS One* 2014;9(11):e113412.
18. Vogel AP, Shirbin C, Churchyard AJ, Stout JC. Speech acoustic markers of early stage and prodromal Huntington's disease: A marker of disease onset? *Neuropsychologia* 2012;50(14):3273-3278.
19. Kaploun L. Acoustic Analysis of Voice/speech Characteristics in Nonsymptomatic Gene Carriers of Huntington's Disease: Does the Speech/voice Differ from Normal Controls? : Columbia University; 2008.
20. Vogel AP, Churchyard A, Stout J, Saft C. Two different phenomena in basic motor speech performance in premanifest Huntington disease. *Neurology* 2016;87(21):2283.

21. Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group. *Movement Disorders* 1996;11(2):136-142.
22. Penney JB, Vonsattel J-P, Macdonald ME, Gusella JF, Myers RH. CAG repeat number governs the development rate of pathology in Huntington's disease. *Annals of Neurology* 1997;41(5):689-692.
23. Paulsen JS, Langbehn DR, Stout JC, et al. Detection of Huntington's disease decades before diagnosis: the Predict-HD study. *J Neurol Neurosurg Psychiatry* 2008;79(8):874-880.
24. Group HS. Unified Huntington's Disease Rating Scale (UHDRS). Available from: www.huntington-study-group.org/UHDRS/tabid/67/Default.aspx [Last accessed 5 June 2014] 2016.
25. Vogel AP, Fletcher J, Maruff P. The impact of task automaticity on speech in noise. *Speech Communication* 2014;65:1-8.
26. Vogel AP, Poole ML, Pemberton H, et al. Motor speech signature of behavioral variant frontotemporal dementia: Refining the phenotype. *Neurology* 2017;89(8):837-844.
27. Vogel AP, Fletcher J, Snyder PJ, Fredrickson A, Maruff P. Reliability, stability, and sensitivity to change and impairment in acoustic measures of timing and frequency. *Journal of Voice* 2011;25(2):137-149.
28. Vogel AP, Maruff P. Monitoring change requires a rethink of assessment practices in voice and speech. *Logopedics Phoniatrics Vocology* 2014;39(2):56-61.
29. Vogel AP, Fletcher J, Maruff P. Acoustic analysis of the effects of sustained wakefulness on speech. *Journal of the Acoustical Society of America* 2010;128(6):3747-3756.
30. Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson's Disease. *IEEE Transactions on Biomedical Engineering* 2012;59(5):1264-1271.
31. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society* 2005;53(4):695-699.
32. Maruff P, Thomas E, Cysique L, et al. Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Archives of Clinical Neuropsychology* 2009;24(2):165-178.
33. Lim YY, Pietrzak RH, Bourgeat P, et al. Relationships Between Performance on the Cogstate Brief Battery, Neurodegeneration, and A β Accumulation in Cognitively Normal Older Adults and Adults with MCI. *Archives of Clinical Neuropsychology* 2014;30(1):49-58.
34. Tiffin J, Asher EJ. The Purdue Pegboard: norms and studies of reliability and validity. *J Appl Psychol* 1948;32(3):234-247.
35. Delaney HD, Vargha A. Comparing several robust tests of stochastic equality with ordinally scaled variables and small to moderate sized samples. *Psychological Methods* 2002;7(4):485-503.
36. Hall P. Theoretical Comparison of Bootstrap Confidence Intervals. *The Annals of Statistics* 1988;16(3):927-953.
37. Ruscio J, Mullen T. Confidence Intervals for the Probability of Superiority Effect Size Measure and the Area Under a Receiver Operating Characteristic Curve. *Multivariate Behavioral Research* 2012;47(2):201-223.
38. Yekutieli D, Benjamini Y. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference* 1999;82(1):171-196.
39. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2005;67(2):301-320.

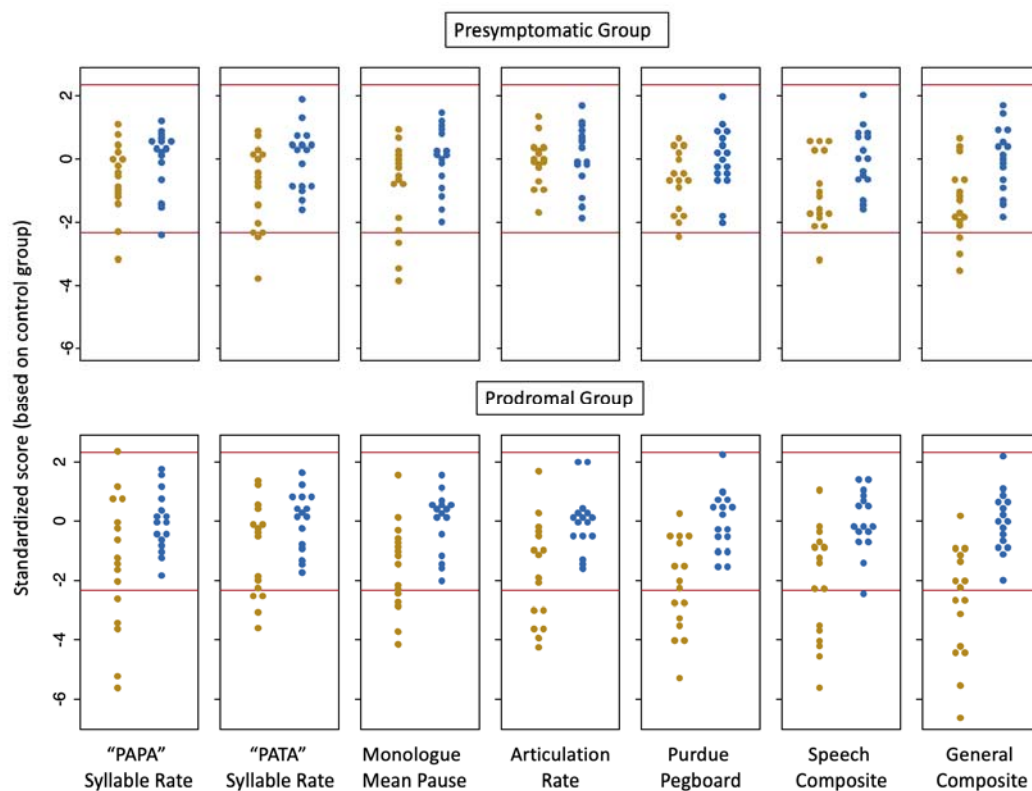
40. Vogel AP, Shirbin C, Komiti A, Churchyard AJ, Stout JC. Speech as a marker of early stage Huntington's Disease onset. *Clinical Genetics*; 2011: Blackwell Publishing Ltd. p. 60.
41. Chan JCS, Stout JC, Shirbin CA, Vogel AP. Listener Detection of Objectively Validated Acoustic Features of Speech in Huntington's Disease. *Journal of Huntington's Disease* 2022;11:71-79.
42. Noffs G, Boonstra FMC, Perera T, et al. Speech metrics, general disability, brain imaging and quality of life in multiple sclerosis. *European Journal of Neurology* 2021;28(1):259-268.
43. Riad R, Titeux H, Lemoine L, et al. Vocal markers from sustained phonation in Huntington's Disease. arXiv preprint arXiv:200605365 2020.
44. Rodrigues FB, Byrne LM, Tortelli R, et al. Mutant huntingtin and neurofilament light have distinct longitudinal dynamics in Huntington's disease. *Science Translational Medicine* 2020;12(574):eabc2888.

Figure 1: Standardized comparison of speech in premanifest and manifest HD to matched controls



Note: Scores of 0.5 (horizontal axis) indicate complete overlap of groups. Scores closer to 1 indicate separation of groups in the direction of increased deficit. Numerical values (q-values) in bold indicate tests with a false discovery rate <.05. Error bars represent 95% confidence intervals.

Figure 2: Standardized participant performance across core speech tasks compared to matched controls in presymptomatic and prodromal HD



Each dot represents a single participant. Red lines represent the upper and lower 1% confidence bounds of the control distributions (Z greater than 2.326). Participant scores outside of these boundaries means their performance sits beyond what is expected for age and sex matched healthy adults. All matched controls fit within the normal range and are represented by blue dots. HD participants are represented by yellow dots.

Table 1: Participant group descriptions and inclusion criteria of HD groups

| | HD groups | Criteria | Group description |
|-------------------|-------------------|-------------------------|--|
| Premanifest phase | Presymptomatic HD | DCL = 0 to 1 | <ul style="list-style-type: none">• No detectable cognitive changes• Possible nonspecific minor motor findings that represent the individual's baseline |
| | Prodromal HD | DCL = 2 or 3 | <ul style="list-style-type: none">• Subtle motor decline from premorbid level of function, with or without minor or major neurocognitive changes |
| Manifest phase | Early HD | DCL = 4 TFC = 7 – 13 | <ul style="list-style-type: none">• Clinical motor symptoms of HD• Detectable change in TFC |
| | Mid-stage HD | DCL = 4 TFC = 4 – 6 | <ul style="list-style-type: none">• Clinical motor symptoms of HD• Extensive detectable changes in TFC |

Note: DCL = diagnostic confidence level; TFC = Total Functional Capacity

Table 2: Demographic, clinical, and genetic characteristics of participants with the expanded HD gene (Mean +/- SD unless specified)

| | Presymptomatic (n=16) | Prodromal (n=16) | Early (n=14) | Mid (n=9) |
|-------------------------------|----------------------------------|-----------------------------|-------------------------|----------------------|
| Age (years) | 41.1 (13.8) | 44.0 (10.6) | 56.1 (10.7) | 48.7 (13.7) |
| Sex (female) | 12 (75%) | 10 (62.5%) | 2(14.3%) | 4 (44.4%) |
| DCL | 0.13 (0.34) | 1.56 (0.73) | 4.0 (0.0) | 4.0 (0.0) |
| DBS ^a | 240.0 (84.6) | 305.4 (94.2) | 418.7(116.9) | 362.9(129.6) |
| Range of DBS | 94.5 – 364 | 143 – 480 | 324.5 – 714 | 150 – 532.5 |
| UHDRS- TMS ^b | 0.3 (1.0) | 4.3 (5.5) | 18.1 (7.6) | 32.7 (11.2) |
| TFC scores ^c | 13.0 (0.0) | 12.6 (1.0) | 10.0 (2.5) | 6.4 (1.5) |
| Range of TFC | 13-13 | 10-13 | 3-13 | 4-9 |
| CAG repeats | 41.8 (2.6) | 42.6 (2.1) | 43.1 (4.0) | 44.8 (5.3) |
| Estimated YTO ^d | 18.8 (10.7) | 12.7 (8.1) | n/a | n/a |
| Disease duration | n/a | n/a | 3.1 (1.5) | 6.3 (2.9) |

Note: DCL = diagnostic confidence level; DBS = disease burden scores; SD = standard deviation; UHDRS-TMS = The United Huntington’s Disease Rating Scale – Total motor score; TFC = Total Functional Capacity; YTO = estimated years to disease onset; ^d Estimated years to disease onset was calculated using probability formulas ²³.

Table 3: Premanifest HD CAG expansion carriers who are outliers on any selected measure, or on a composite score that includes the Purdue Pegboard test (General) or only speech measures (Speech).

| Px ID | Monologue mean pause duration | Articulation Rate | “PAPA” Rate | “PATA” Rate | Purdue Pegboard | Speech Composite | General Composite |
|-----------------------------|-------------------------------|-------------------|-------------|-------------|-----------------|------------------|-------------------|
| Presymptomatic Phase | | | | | | | |
| 2 | -3.41 | | -3.18 | | | | -2.95 |
| 30 | | | | -3.83 | | -3.27 | .348 |
| 46 | | | | | -2.44 | | -2.50 |
| 36 | | | | -2.53 | | | |
| 39 | | | | -2.38 | | | |
| 31 | -2.63 | | | | | | |
| 49 | -3.82 | | | | | | |
| Prodromal Phase | | | | | | | |
| 7 | | | | | -2.73 | | |
| 10 | | | -3.25 | | -3.24 | | |
| 11 | -2.40 | | | | -2.24 | | |
| 12 | | -2.86 | -5.07 | -2.26 | -3.98 | -3.95 | -4.69 |
| 21 | -2.86 | | | | | | |
| 25 | -3.62 | -3.49 | | | -3.98 | -2.53 | -3.44 |
| 28 | -2.97 | -2.97 | | | | -3.04 | -3.14 |
| 29 | -3.86 | -3.86 | -3.53 | -3.63 | -5.23 | -5.15 | -6.13 |
| 40 | -3.48 | -3.47 | -5.57 | -2.57 | -3.48 | -4.69 | -5.20 |
| 43 | -4.16 | -4.16 | -2.64 | -3.00 | -2.74 | -4.12 | -4.47 |
| 42 | | | | | | -2.39 | -2.57 |
| 57 | | | | -2.58 | | | |