

Original Paper

Association Between Acoustic Features and Neuropsychological Test Performance in the Framingham Heart Study: Observational Study

Huitong Ding¹, PhD; Amiya Mandapati^{2,3}, BA; Cody Karjadi^{1,4}, MSc; Ting Fang Alvin Ang^{1,4,5,6}, MPH, MD; Sophia Lu⁵, BS; Xiao Miao⁷, MD; James Glass⁸, PhD; Rhoda Au^{1,4,5,6,9}, MBA, PhD; Honghuang Lin¹⁰, PhD

¹Department of Anatomy and Neurobiology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States

²Department of Religious Studies, Brown University, Providence, RI, United States

³The Warren Alpert Medical School, Brown University, Providence, RI, United States

⁴The Framingham Heart Study, Chobanian & Avedisian School of Medicine, Boston University, Framingham, MA, United States

⁵Slone Epidemiology Center, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States

⁶Department of Epidemiology, School of Public Health, Boston University, Boston, MA, United States

⁷Innovation Research Institute of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

⁸Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, United States

⁹Department of Neurology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States

¹⁰Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

Corresponding Author:

Honghuang Lin, PhD

Department of Medicine

University of Massachusetts Chan Medical School

55 N Lake Ave

Worcester, MA, 01655

United States

Phone: 1 7744554881

Email: honghuang.lin@umassmed.edu

Abstract

Background: Human voice has increasingly been recognized as an effective indicator for the detection of cognitive disorders. However, the association of acoustic features with specific cognitive functions and mild cognitive impairment (MCI) has yet to be evaluated in a large community-based population.

Objective: This study aimed to investigate the association between acoustic features and neuropsychological (NP) tests across multiple cognitive domains and evaluate the added predictive power of acoustic composite scores for the classification of MCI.

Methods: This study included participants without dementia from the Framingham Heart Study, a large community-based cohort with longitudinal surveillance for incident dementia. For each participant, 65 low-level acoustic descriptors were derived from voice recordings of NP test administration. The associations between individual acoustic descriptors and 18 NP tests were assessed with linear mixed-effect models adjusted for age, sex, and education. Acoustic composite scores were then built by combining acoustic features significantly associated with NP tests. The added prediction power of acoustic composite scores for prevalent and incident MCI was also evaluated.

Results: The study included 7874 voice recordings from 4950 participants (age: mean 62, SD 14 years; 4336/7874, 55.07% women), of whom 453 were diagnosed with MCI. In all, 8 NP tests were associated with more than 15 acoustic features after adjusting for multiple testing. Additionally, 4 of the acoustic composite scores were significantly associated with prevalent MCI and 7 were associated with incident MCI. The acoustic composite scores can increase the area under the curve of the baseline model for MCI prediction from 0.712 to 0.755.

Conclusions: Multiple acoustic features are significantly associated with NP test performance and MCI, which can potentially be used as digital biomarkers for early cognitive impairment monitoring.

(*J Med Internet Res* 2022;24(12):e42886) doi: [10.2196/42886](https://doi.org/10.2196/42886)

KEYWORDS

mild cognitive impairment; digital voice; neuropsychological test; association; prediction

Introduction

Alzheimer disease (AD) is a chronic neurodegenerative disease characterized behaviorally by memory loss, language impairment, motor problems, loss of executive function, and emotional distress, which can progress to severe levels. There are currently no definitive disease-modifying treatment methods [1], but general consensus is that early detection is critical. Interventions through the reduction of modifiable risk factors may serve to delay, attenuate, or even prevent disease onset and progression [2,3]. Mild cognitive impairment (MCI) is a prodromal stage of AD in which cognitive decline does not affect essential functions of daily life [4], but some individuals may have difficulty remembering events and situations, as well as problems with executive function [5]. The detection of MCI is critical to initiate current interventions that may slow down the neurodegenerative process [6] and participate in clinical trials that may lead to effective treatments.

At present, diagnosis relies largely on some combination of clinical examination [7], neuroimaging (eg, magnetic resonance imaging [8] and positron emission tomography [9]), and neuropsychological (NP) testing [10]. Fluid biomarkers are being developed as alternatives to expensive and burdensome imaging through the analysis of cerebrospinal fluid [11] and blood analysis [12]. Although substantial advancements have been made in developing pathological indicators of AD (eg, imaging and fluid biomarkers), surprisingly little has been done to develop better cognitive assessment methods beyond the traditional NP tests. The well-documented heterogeneity of cognition has made the accurate diagnosis of MCI elusive [13,14].

Producing speech is a cognitively complex task [15], and recording speech is relatively easy given the widespread accessibility to recording devices. Research has found that language deficits may occur in the prodromal stages [16] of cognitive impairment, which present years prior to clinical diagnosis [17,18], potentially making it an effective indicator for MCI. Meanwhile, the development of speech feature extraction technology offers the possibility of quantifying voice signal properties from multiple dimensions. It empowers the comprehensive description of specific pathologies by voice features. The lexical, acoustic, and syntactic features extracted from the human voice have been shown to be significantly associated with dementia [19,20]. Using voice-based biomarkers as a screening method presupposes an economic solution for the early diagnosis of MCI. Increasing evidence suggests that

the human voice could be used as a powerful resource to derive pathologically appropriate biomarkers for dementia. Multiple acoustic biomarkers have also been related to the future risk of dementia [21].

Applying the findings of earlier research to a general population, however, is difficult due to the small sample sizes and use of cognitive assessment protocols that are not sufficiently comprehensive. Further, voice analyses that include linguistic features are difficult to generalize to other languages. There remains a paucity of research determining the relationship between acoustic features and NP tests that span across multiple cognitive domains. In addition, a comprehensive characterization of acoustic features that are associated with incident MCI is warranted. The objective of this study was to investigate the association of acoustic features and different NP test scores across cognitive domains and how they compare in identifying prevalent and incident MCI in the Framingham Heart Study (FHS) community-based cohort.

Methods**Sample Selection**

The original sample included 9253 observations from 5189 participants who completed at least one NP assessment that was voice recorded. A subset of participants had multiple recordings over the course of the study period. Each digital voice recording and the corresponding NP tests were treated as 1 observation. Exclusion criteria included those observations with missing education information (n=492), prevalent dementia (n=313), flagged as potential MCI but have not gone through dementia review (n=551), and those whose voice recording was less than 10 minutes in length (n=23).

Ethics Approval

The Institutional Review Board of the Boston University Medical Campus approved the procedures and protocols of the Framingham Heart Study (FHS is H-32132). All participants provided written informed consent.

NP Assessment

The details of FHS NP test administration have been reported previously [22]. Multiple cognitive domains are measured by 18 different tests [23-27] including verbal memory, verbal fluency, visual memory, attention and concentration, executive function, abstract reasoning, visuospatial organization, and language, as is illustrated in Table 1.

Table 1. Cognitive domain and corresponding neuropsychological (NP) tests.

Cognitive domain	NP test
Verbal memory	<ul style="list-style-type: none"> • Logical Memory—Immediate Recall • Logical Memory—Delayed Recall • Logical Memory—Recognition • Paired Associate Learning—Immediate Recall • Paired Associate Learning—Delayed Recall • Paired Associate Learning—Recognition
Visual memory	<ul style="list-style-type: none"> • Visual Reproduction—Immediate Recall • Visual Reproduction—Delayed Recall • Visual Reproduction—Recognition
Attention and concentration	<ul style="list-style-type: none"> • Digit Span—Forward • Trail Making Test A
Executive function	<ul style="list-style-type: none"> • Digit Span—Backward • Trail Making Test B
Abstract reasoning	<ul style="list-style-type: none"> • Similarities
Language	<ul style="list-style-type: none"> • Boston Naming Test—30-item version
Visuoperceptual organization	<ul style="list-style-type: none"> • Hooper Visual Organization Test
Verbal fluency	<ul style="list-style-type: none"> • Controlled Oral Word Association Test • Category Naming Test—Animal

Voice Recordings

Since 2005, the FHS has been digitally recording all spoken responses during NP test administration, which encompasses the verbal interactions between the tester and the participant. This study included digital voice recordings obtained from September 2005 to March 2020. OpenSMILE software (version 2.1.3) [28] was used to extract an acoustic feature set [29], which contains 65 low-level descriptors (LLDs) from these recordings. This acoustic feature set covers a broad range of information of the voice recordings including pitch, voice quality, loudness, signal energy, waveform, auditory, fast Fourier transform spectrum, spectral, and cepstral, which has been described in detail in a prior study [30]. The feature set has also been used in many fields, such as speech processing, music information retrieval, and emotion recognition [31]. The description of these features is summarized in Table S1 in [Multimedia Appendix 1](#). More details of these features can be found in the previous publication [30]. There are some audio recordings with 1 channel (mono; n=4738), and the others were recorded with 2 channels (stereo; n=3136). For the recordings with 2 channels, we included the first channel in the analysis. Each recording was divided into segments of 20 milliseconds using a sliding window approach with a shifting size of 10 milliseconds. The LLD features were extracted from these segments. For each recording, we further computed the mean of each LLD feature to capture its high-level statistical features, which were then normalized.

Ascertainment of MCI

The cognitive status of FHS participants included assessments by NP tests. For those identified with possible cognitive impairment, NP tests were administered on average about every

1 to 2 years. When potential cognitive impairment decline was present, a clinical review was conducted by a panel with at least one neurologist and one neuropsychologist. MCI diagnosis was determined by the review panel, which required that the participant exhibit evidence of a decline in cognitive performance in 1 or more cognitive domains, have no records indicating functional decline, and do not meet the criteria for dementia [32]. The Clinical Dementia Rating scale [33] was used to quantify the severity of impairment. In all, 2 outcomes were considered in this study. The prevalent MCI cases were subjects who were diagnosed with MCI before or at the time when the voice was recorded. The incident MCI cases were all subjects who were cognitively intact at baseline but were diagnosed with MCI during the follow-up.

Statistical Analyses

To compare the difference between demographics and standard NP test scores in MCI and normal control groups, Wilcoxon rank sum test was used for continuous variables [34]. The chi-square test was used to compare differences in frequencies for categorical variables [35]. Log transformations were applied for NP tests with skewed distributions to normalize them. Normalized values of NP tests and acoustic features were used in the analysis. Linear mixed-effects models were used to quantify the association between each acoustic feature and NP tests [36].

A set of acoustic composite scores was generated by regressing each NP test against the group of acoustic features that were significantly associated with each NP test. The acoustic composite score is a weighted combination of acoustic features. The weight of each acoustic feature in the composite score was derived by training a linear mixed-effects effect model. For

participant i , the acoustic composite score of an NP test is defined as

$$acoustic_NP_i = \sum_{j=1}^m \alpha_j \times V_{ij} \quad (1)$$

where m is the number of acoustic features significantly associated with the NP test, α_j is the estimate of effect size for the acoustics feature j derived from the linear mixed-effects effect model, and V_{ij} is the normalized acoustics feature j for participant i . The association between normalized acoustic composite scores with corresponding NP tests was assessed by linear mixed-effects models.

The association of normalized acoustic composite scores with prevalent MCI was assessed by logistic regression models. Based on the regression coefficients, the odds ratios (ORs) and 95% CIs were estimated.

To determine the relationship between acoustic composite scores and incident MCI, participants whose age at the voice recording was <60 years ($n=2718$) and those with prevalent MCI ($n=222$) were excluded. The first observation of each participant was included in this analysis. The association between acoustic composite scores with incident MCI was quantified by Cox proportional hazards models (censored at the last date of contact or death) [37]. All models were adjusted for age, sex, and education. Bonferroni correction was used to adjust for multiple tests.

We further evaluated the added predictability of the acoustic composite score for incident MCI. Receiver operating characteristic (ROC) analysis was performed to estimate the area under the curve (AUC) using a random forest model. A baseline model was constructed using age, sex, and education as predictors. A second model was constructed using these predictors and additional acoustic composite scores that were found to be significantly related to specific NP tests. The mean AUC of 10-fold cross-validation was computed for each model for comparison. We also performed a secondary analysis by including NP tests and clinical risk factors in the prediction of incident MCI. The statistical analyses were performed using Python software (version 3.9.7; Python Software Foundation).

Results

Our study included 7874 observations from 4950 participants of FHS (age: mean 62, SD 14 years; 4336/7874, 55.07% women; 4279/7874, 54.34% self-reported college-level education or higher). Most participants (2657/4950, 53.68%) had 1 voice recording. Some participants (1775/4950, 35.86%) had 2 recordings, and the remaining participants (518/4950, 10.46%) had 3 or more recordings. Among these observations, 453 of these observations were diagnosed with MCI. The details of sample characteristics are shown in Table 2.

We examined the association of acoustic features with NP tests. As shown in Table 3, eight NP tests (Visual Reproduction—Immediate Recall [VRi], Visual Reproduction—Delayed Recall [VRd], Digit Span—Forward, Digit Span—Backward, Similarities [SIM], Boston Naming Test—30-item version, Controlled Oral Word Association Test

[FAS], and Category Naming Test—Animal) were associated with more than 15 acoustic features. The *mfcc_sma* [2] was the most significant acoustic feature with 3 NP tests (Boston Naming Test—30-item version, FAS, and Category Naming Test—Animal) after Bonferroni correction ($P < 7.7 \times 10^{-4}$) that represents Mel-frequency cepstral coefficient (MFCC) 2. The details of associations between acoustic features and NP tests are fully depicted in Table S2 in Multimedia Appendix 1. We also summarized the acoustic features that were significantly associated with NP tests across cognitive domains in Table S3 in Multimedia Appendix 1. It shows that visual memory was associated with 49 acoustic features. Each cognitive domain had an average of 28 associated acoustic features. In the sensitivity analysis, besides age, sex, and education, we further included employment as an additional covariate to examine the stability of the association between acoustic features and NP tests. As shown in Table S4 in Multimedia Appendix 1, similar acoustic features were found to be associated with NP tests. In addition, we also examined the correlation between acoustic features and NP tests collected at the same time and a later time. For each NP test conducted at the first exam, we compared its correlation with acoustic features collected at the first exam and the second exam. As shown in Table S5 in Multimedia Appendix 1, only moderate changes were observed between the 2 exams.

Acoustic composite scores were also generated using the significant acoustic features for each NP test. As shown in Table 4, all these scores were significantly associated with their corresponding NP tests.

We then performed association analysis of acoustic composite scores with prevalent MCI. Table 5 shows that 4 acoustic composite scores (*acoustic_LMr*, *acoustic_TrailsB*, *acoustic_FAS*, and *acoustic_CNT_Animal*) were significantly associated with prevalent MCI (OR ranging from 0.69 to 1.23; $P < 3.1 \times 10^{-3}$). Lower acoustics composite scores (*acoustic_TrailsB*, *acoustic_FAS*, and *acoustic_CNT_Animal*) were associated with higher OR of MCI after adjusting for age, sex, and education ($P < 3.1 \times 10^{-3}$). The most significant acoustic composite score was for FAS Animal test ($P = 2.3 \times 10^{-7}$).

We further examined the association of acoustic composite scores with incident MCI by restricting the analysis to 2010 participants who were aged ≥ 60 years. Among them, 145 participants have incident MCI. As shown in Table 6, the acoustic composite scores for Logical Memory—Immediate Recall (LMi), VRi, VRd, Visual Reproduction—Recognition (VRr), SIM, Trail Making Test B (TrailsB), and Hooper Visual Organization Test (HVOT) tests were significantly associated with incident MCI ($P < 3.1 \times 10^{-3}$). Higher acoustic composite scores for VRi, VRd, SIM, and TrailsB tests were associated with higher MCI risk. The other 3 scores were negatively associated with MCI risk with hazard ratio lower than 1 after adjusting for age, sex, and education. We further built 2 Cox regression models for incident MCI to show the contribution of acoustic features. Model 1 includes age, sex, and education as predictors. Model 2 includes age, sex, education, and all significant associated acoustic composite scores with incident MCI. The change in Akaike information criterion [38] with the addition of acoustic composite scores to the model was

calculated. We observed a smaller Akaike information criterion for model 2, suggesting that the model better fit the prediction.

The added predictive power of *acoustic_LMi*, *acoustic_VRi*, *acoustic_VRd*, *acoustic_VRr*, *acoustic_SIM*, *acoustic_TrailsB*, and *acoustic_HVOT* for incident MCI were evaluated by comparing the AUC of different models. Model 1 only included age, sex, and education as the predictors of incident MCI. Besides age, sex, and education, Model 2 included 7 composite scores that were significantly associated with incident MCI as the predictors. Model 3 included age, sex, education, and 18 NP tests as predictors. [Figure 1](#) shows that the AUC of MCI prediction can be improved from 0.712 (model 1) to 0.755 (model 2) by including acoustic composite scores of LMi, VRi,

VRd, VRr, SIM, TrailsB, and HVOT tests. As shown in [Figure S1](#) in [Multimedia Appendix 1](#), the model with NP tests reached AUC=0.761, which is comparable to the one including demographic factors and acoustic composite scores (DeLong test $P=.97$). However, both models showed significant improvement over model 1 that included only demographic factors (DeLong test $P=.03$ and $P=.03$ for model 2 and model 3, respectively). These results indicate that the acoustics composite scores have similar predictive power to traditional NP tests. Compared to the burden of conducting NP tests, the prediction model based on acoustic features relied minimally on NP expertise; these results suggest the feasibility of developing real-time cognitive screening tools.

Table 2. Baseline characteristics.

Variable	Total observation (N=7874)	MCI ^a (n=453)	NC ^b (n=7421)	P value ^c
Age (years), mean (SD)	62 (14)	81 (8)	61 (14)	<.001
Gender, n (%)				.84
Women	4336 (55.07)	252 (55.63)	4084 (55.03)	
Men	3538 (44.93)	201 (44.37)	3337 (44.97)	
Education, n (%)				<.001
No high school	202 (2.57)	53 (11.70)	149 (2.01)	
High school	1443 (18.33)	118 (26.05)	1295 (17.45)	
Some college	1950 (24.77)	134 (29.58)	1816 (24.47)	
College and higher	4279 (54.34)	148 (32.67)	4161 (56.07)	
NP^d test score, mean (SD)				
LMi ^e	12.35 (3.62)	8.53 (3.76)	12.58 (3.48)	<.001
LMd ^f	11.36 (3.83)	6.93 (4.11)	11.62 (3.65)	<.001
LMr ^g	9.52 (1.28)	8.59 (1.72)	9.57 (1.23)	<.001
VRi ^h	8.61 (2.91)	4.48 (2.23)	8.85 (2.76)	<.001
VRd ⁱ	7.91 (3.17)	3.11 (2.30)	8.19 (2.99)	<.001
VRr ^j	3.11 (1.01)	1.89 (1.06)	3.18 (0.96)	<.001
PASi ^k	14.45 (3.58)	10.02 (2.79)	14.71 (3.45)	<.001
PASd ^l	8.56 (1.47)	6.56 (1.60)	8.68 (1.38)	<.001
PASr ^m	9.82 (0.64)	8.83 (1.74)	9.88 (0.45)	<.001
DSr ⁿ	6.71 (1.31)	6.06 (1.20)	6.75 (1.30)	<.001
DSb ^o	4.92 (1.30)	4.12 (1.01)	4.97 (1.30)	<.001
SIM ^p	16.83 (3.61)	12.63 (4.30)	17.08 (3.40)	<.001
BNT30 ^q	27.22 (2.81)	23.66 (4.14)	27.43 (2.56)	<.001
TrailsA ^r	0.42 (0.15)	0.66 (0.21)	0.40 (0.14)	<.001
TrailsB ^s	0.85 (0.34)	1.54 (0.50)	0.82 (0.29)	<.001
HVOT ^t	3.26 (0.15)	3.06 (0.22)	3.27 (0.13)	<.001
FAS ^u	39.85 (12.52)	28.76 (11.68)	40.50 (12.26)	<.001
CNT_Animal ^v	19.48 (5.68)	12.22 (4.37)	19.91 (5.46)	<.001

^aMCI: mild cognitive impairment.

^bNC: normal control.

^cSignificant associations were claimed if $P < .05/18 \approx .002$.

^dNP: neuropsychological.

^eLMi: Logical Memory—Immediate Recall.

^fLMd: Logical Memory—Delayed Recall.

^gLMr: Logical Memory—Recognition.

^hVRi: Visual Reproduction—Immediate Recall.

ⁱVRd: Visual Reproduction—Delayed Recall.

^jVRr: Visual Reproduction—Recognition.

^kPASi: Paired Associate Learning—Immediate Recall.

^lPASd: Paired Associate Learning—Delayed Recall.

^mPASr: Paired Associate Learning—Recognition.

ⁿDSf: Digit Span—Forward.

^oDSb: Digit Span—Backward.

^pSIM: Similarities.

^qBNT30: Boston Naming Test—30-item version.

^rTrailsA: Trail Making Test A.

^sTrailsB: Trail Making Test B.

^tHVOT: Hooper Visual Organization Test.

^uFAS: Controlled Oral Word Association Test.

^vCNT_Animal: Category Naming Test—Animal.

Table 3. The most significant acoustic feature for each neuropsychological (NP) test.

NP test	Significant acoustic features, n	The most significant acoustic feature	Effect size	SE	<i>P</i> value ^a
LMi ^b	7	audSpec_Rfilt_sma [25]	0.0490	0.0095	2.7×10^{-7}
LMd ^c	3	audSpec_Rfilt_sma [25]	0.0402	0.0094	1.9×10^{-5}
LMr ^d	3	audSpec_Rfilt_sma [23]	0.0397	0.0108	2.3×10^{-4}
VRi ^e	49	mfcc_sma [11]	0.1409	0.0082	8.4×10^{-66}
VRd ^f	43	mfcc_sma [11]	0.1137	0.0082	3.7×10^{-44}
VRr ^g	10	pcm_fftMag_spectralRollOff75.0_sma	-0.0358	0.0095	1.7×10^{-4}
PASi ^h	0	N/A ⁱ	N/A	N/A	N/A
PASd ^j	0	N/A	N/A	N/A	N/A
PASr ^k	7	audSpec_Rfilt_sma [1]	-0.0709	0.0112	2.3×10^{-10}
DSf ^l	44	audSpec_Rfilt_sma [6]	0.0898	0.0107	4.8×10^{-17}
DSb ^m	30	audSpec_Rfilt_sma [5]	0.0624	0.0110	1.2×10^{-8}
SIM ⁿ	24	pcm_fftMag_spectralRollOff75.0_sma	-0.0530	0.0084	2.4×10^{-10}
BNT30 ^o	23	mfcc_sma [2]	0.0433	0.0069	3.2×10^{-10}
TrailsA ^p	15	pcm_fftMag_spectralSkewness_sma	-0.0363	0.0075	1.4×10^{-6}
TrailsB ^q	1	pcm_fftMag_spectralSkewness_sma	-0.0269	0.0074	3.1×10^{-4}
HVOT ^r	5	F0final_sma	-0.0472	0.0093	3.6×10^{-7}
FAS ^s	26	mfcc_sma [2]	0.0534	0.0073	3.6×10^{-13}
CNT_Animal ^t	34	mfcc_sma [2]	0.0715	0.0082	2.6×10^{-18}

^aSignificant associations were claimed if $P < .05/65 \approx 7.7 \times 10^{-4}$.

^bLMi: Logical Memory—Immediate Recall.

^cLMd: Logical Memory—Delayed Recall.

^dLMr: Logical Memory—Recognition.

^eVRi: Visual Reproduction—Immediate Recall.

^fVRd: Visual Reproduction—Delayed Recall.

^gVRr: Visual Reproduction—Recognition.

^hPASi: Paired Associate Learning—Immediate Recall.

ⁱN/A: not applicable.

^jPASd: Paired Associate Learning—Delayed Recall.

^kPASr: Paired Associate Learning—Recognition.

^lDSf: Digit Span—Forward.

^mDSb: Digit Span—Backward.

ⁿSIM: Similarities.

^oBNT30: Boston Naming Test—30-item version.

^pTrailsA: Trail Making Test A.

^qTrailsB: Trail Making Test B.

^rHVOT: Hooper Visual Organization Test.

^sFAS: Controlled Oral Word Association Test.

^tCNT_Animal: Category Naming Test—Animal.

Table 4. Association between acoustic composite scores and corresponding neuropsychological tests.

Acoustic composite score	Effect size	SE	<i>P</i> value ^a
acoustic_LMi ^b	0.0579	0.0094	6.6×10^{-10}
acoustic_LMd ^c	0.0310	0.0095	1.1×10^{-3}
acoustic_LMr ^d	0.0358	0.0105	6.8×10^{-4}
acoustic_VRi ^e	0.1510	0.0086	3.3×10^{-69}
acoustic_VRd ^f	0.1079	0.0086	6.5×10^{-36}
acoustic_VRr ^g	-0.0291	0.0098	3.0×10^{-3}
acoustic_PASr ^h	0.0841	0.0114	1.3×10^{-13}
acoustic_DSf ⁱ	0.1298	0.0097	1.8×10^{-40}
acoustic_DSB ^j	0.0553	0.0102	6.2×10^{-8}
acoustic_SIM ^k	0.0719	0.0089	5.1×10^{-16}
acoustic_BNT30 ^l	0.0458	0.0071	1.4×10^{-10}
acoustic_TrailsA ^m	0.0408	0.0088	3.0×10^{-6}
acoustic_TrailsB ⁿ	-0.0269	0.0075	3.1×10^{-4}
acoustic_HVOT ^o	0.0284	0.0090	1.7×10^{-3}
acoustic_FAS ^p	0.0827	0.0079	1.4×10^{-25}
acoustic_CNT_Animal ^q	0.0529	0.0098	6.5×10^{-8}

^aSignificant associations were claimed if $P < .05/16 \approx 3.1 \times 10^{-3}$.

^bLMi: Logical Memory—Immediate Recall.

^cLMd: Logical Memory—Delayed Recall.

^dLMr: Logical Memory—Recognition.

^eVRi: Visual Reproduction—Immediate Recall.

^fVRd: Visual Reproduction—Delayed Recall.

^gVRr: Visual Reproduction—Recognition.

^hPASr: Paired Associate Learning—Recognition.

ⁱDSf: Digit Span—Forward.

^jDSb: Digit Span—Backward.

^kSIM: Similarities.

^lBNT30: Boston Naming Test—30-item version.

^mTrailsA: Trail Making Test A.

ⁿTrailsB: Trail Making Test B.

^oHVOT: Hooper Visual Organization Test.

^pFAS: Controlled Oral Word Association Test.

^qCNT_Animal: Category Naming Test—Animal.

Table 5. Association between acoustic composite scores and prevalent mild cognitive impairment.

Acoustic composite score	Odds ratio (95% CI)	<i>P</i> value ^a
acoustic_LMi ^b	1.09 (0.94-1.26)	2.6×10^{-1}
acoustic_LMd ^c	1.14 (0.99-1.31)	7.4×10^{-2}
acoustic_LMr ^d	1.23 (1.08-1.40)	1.6×10^{-3}
acoustic_VRi ^e	1.05 (0.92-1.19)	4.7×10^{-1}
acoustic_VRd ^f	1.07 (0.94-1.21)	3.2×10^{-1}
acoustic_VRr ^g	0.94 (0.80-1.10)	4.6×10^{-1}
acoustic_PASr ^h	0.9 (0.81-0.99)	3.6×10^{-2}
acoustic_DSf ⁱ	1.17 (1.04-1.32)	1.1×10^{-2}
acoustic_DSB ^j	0.94 (0.83-1.07)	3.5×10^{-1}
acoustic_SIM ^k	0.94 (0.84-1.06)	3.1×10^{-1}
acoustic_BNT30 ^l	0.92 (0.82-1.04)	2.0×10^{-1}
acoustic_TrailsA ^m	1.12 (0.98-1.28)	9.6×10^{-2}
acoustic_TrailsB ⁿ	0.69 (0.59-0.81)	1.0×10^{-5}
acoustic_HVOT ^o	0.91 (0.81-1.03)	1.4×10^{-1}
acoustic_FAS ^p	0.72 (0.64-0.81)	3.9×10^{-8}
acoustic_CNT_Animal ^q	0.70 (0.61-0.80)	2.3×10^{-7}

^aSignificant associations were claimed if $P < .05/16 \approx 3.1 \times 10^{-3}$.

^bLMi: Logical Memory—Immediate Recall.

^cLMd: Logical Memory—Delayed Recall.

^dLMr: Logical Memory—Recognition.

^eVRi: Visual Reproduction—Immediate Recall.

^fVRd: Visual Reproduction—Delayed Recall.

^gVRr: Visual Reproduction—Recognition.

^hPASr: Paired Associate Learning—Recognition.

ⁱDSf: Digit Span—Forward.

^jDSb: Digit Span—Backward.

^kSIM: Similarities.

^lBNT30: Boston Naming Test—30-item version.

^mTrailsA: Trail Making Test A.

ⁿTrailsB: Trail Making Test B.

^oHVOT: Hooper Visual Organization Test.

^pFAS: Controlled Oral Word Association Test.

^qCNT_Animal: Category Naming Test—Animal.

Table 6. Association between acoustic composite scores and incident mild cognitive impairment.

Acoustic composite score	Hazard ratio (95% CI)	<i>P</i> value ^a
acoustic_LMi ^b	0.60 (0.47-0.77)	5.1×10^{-5}
acoustic_LMd ^c	0.76 (0.59-0.97)	2.9×10^{-2}
acoustic_LMr ^d	0.74 (0.61-0.91)	3.9×10^{-3}
acoustic_VRi ^e	1.28 (1.10-1.48)	1.1×10^{-3}
acoustic_VRd ^f	1.25 (1.08-1.44)	2.4×10^{-3}
acoustic_VRr ^g	0.44 (0.33-0.59)	6.0×10^{-8}
acoustic_PASr ^h	1.11 (0.95-1.30)	2.0×10^{-1}
acoustic_DSf ⁱ	1.11 (0.96-1.29)	1.6×10^{-1}
acoustic_DSb ^j	1.09 (0.93-1.27)	2.9×10^{-1}
acoustic_SIM ^k	1.37 (1.16-1.61)	1.7×10^{-4}
acoustic_BNT30 ^l	1.23 (1.06-1.43)	6.4×10^{-3}
acoustic_TrailsA ^m	0.75 (0.61-0.93)	7.9×10^{-3}
acoustic_TrailsB ⁿ	2.03 (1.58-2.60)	2.5×10^{-8}
acoustic_HVOT ^o	0.78 (0.67-0.91)	1.7×10^{-3}
acoustic_FAS ^p	0.87 (0.76-1.01)	6.1×10^{-2}
acoustic_CNT_Animal ^q	0.85 (0.70-1.02)	8.6×10^{-2}

^aSignificant associations were claimed if $P < .05/16 \approx 3.1 \times 10^{-3}$.

^bLMi: Logical Memory—Immediate Recall.

^cLMd: Logical Memory—Delayed Recall.

^dLMr: Logical Memory—Recognition.

^eVRi: Visual Reproduction—Immediate Recall.

^fVRd: Visual Reproduction—Delayed Recall.

^gVRr: Visual Reproduction—Recognition.

^hPASr: Paired Associate Learning—Recognition.

ⁱDSf: Digit Span—Forward.

^jDSb: Digit Span—Backward.

^kSIM: Similarities.

^lBNT30: Boston Naming Test—30-item version.

^mTrailsA: Trail Making Test A.

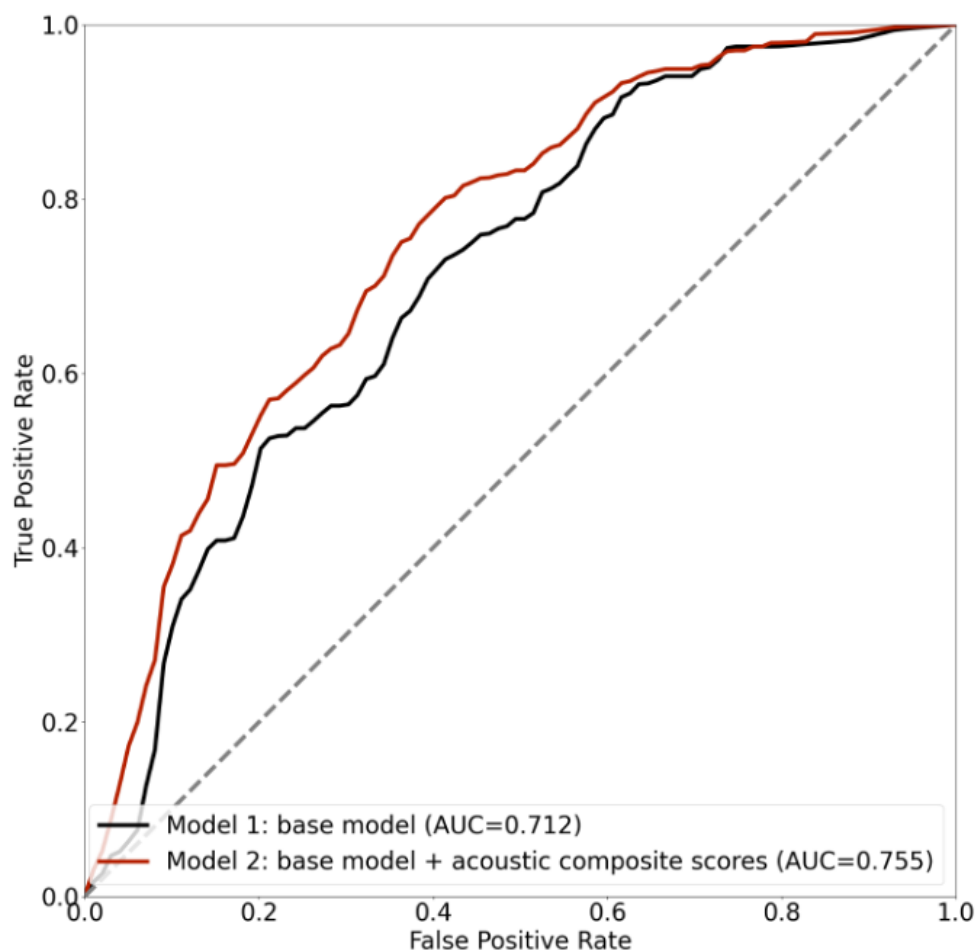
ⁿTrailsB: Trail Making Test B.

^oHVOT: Hooper Visual Organization Test.

^pFAS: Controlled Oral Word Association Test.

^qCNT_Animal: Category Naming Test—Animal.

Figure 1. The receiver operating characteristic curves of 2 models to predict incident mild cognitive impairment. AUC: area under the curve.



Discussion

Principal Findings

Relating acoustic features with NP test performance is potentially a novel way for screening at the preclinical stages of AD and other dementias. This paper clarifies the relationship between comprehensive acoustic features and NP test performance on large cohort data. Representations relative spectra-style filtered auditory spectrum (spectral), MFCC (cepstral), and magnitude of spectral features (spectral) are 3 categories of acoustic features that were significantly associated with NP test performance. Representations relative spectra-style filtered auditory spectrum is a filtered representation of an audio signal that is robust to additive and convolutional noise [39]. MFCC is a standardized technique for audio feature extraction [40]. It helps in reducing the frequency information of the input speech signal into coefficients, which represent audio based on the perception of human auditory systems. Prior studies have detected changes of these features in people with neurodegenerative processes [41-43]. The acoustic composite score generated for each NP test was a linear combination of LLD features, which are clinically easily interpretable. As stated in the results above, 4 acoustic composite scores were significantly associated with prevalent MCI, and 7 were also found to be significantly associated with incident MCI. Furthermore, the score corresponding to TrailsB test is

significantly associated with both prevalent MCI and incident MCI.

Results could expand current evidence regarding the predictive ability of digital voice on MCI that are critical to monitor early cognitive decline. The added predictive ability of acoustic features was evaluated by constructing random forest models with baseline features and additional acoustic composite scores. The model with baseline features and 7 acoustic composite scores corresponding to LMi, VRi, VRd, VRr, SIM, TrailsB, and HVOT tests could achieve an AUC of 0.755 for incident MCI prediction. Monitoring acoustic features outside of the clinical settings offers a more convenient way to aid in the assessment of cognitive health than traditional methods. Increasing evidence suggests that the human voice can be a predictor of cognitive decline before a clinical diagnosis of AD is made [44]. It has been used to screen for MCI [45], dementia [46], and other neurodegenerative diseases such as Parkinson [47] and Huntington disease [48] because of its ease of administration and clinical assessment capability. Moreover, the easy acquisition of voice in daily life makes it an ideal measure for long-term monitoring of cognitive status. However, there is a lack of research about the relationship between acoustic features and NP tests that reflect multiple cognitive functions. Our study could provide some construct validity for this point. In this study, we recorded voice for NP tests that require verbal responses. Although some NP tests do not require verbal responses, these tests might tap some cognitive domains

similar to those that require verbal responses. We therefore included these tests as well to capture potential application of acoustic characteristics to assess different cognitive domains. Each NP test might require multiple cognitive domains to complete, which might be shared with other NP tests with subtle differences. Given the rich information from human voice, our study suggests that acoustic features might serve as a new data modality to test this nuance.

Notably, the association between acoustic features and a standard epidemiologic NP test procedure was examined based on participants from a community-based cohort with a diverse range of ages and health conditions. The large volume of voice data provides a more robust representation of participants. Each voice recording lasts, on average, around an hour and contains a wealth of information. The longitudinal collection of data provides a great opportunity to assess the cognitive health of participants throughout the entire course of the disease and prospectively reveals a temporal relationship between acoustic features and MCI. It is worth noting that 4 of the acoustic composite scores (*acoustic_LMr*, *acoustic_TrailsB*, *acoustic_FAS*, and *acoustic_CNT_Animal*) were significantly associated with prevalent MCI, but 7 acoustic composite scores (*acoustic_LMi*, *acoustic_VRi*, *acoustic_VRd*, *acoustic_VRr*, *acoustic_SIM*, *acoustic_TrailsB*, and *acoustic_HVOT*) were associated with incident MCI. It seems that the voice characteristics differentiating prevalent MCI cases from patients who are still cognitively intact are different from the voice characteristics that are predictive of future risk of cognitive impairment. Future research is needed to further investigate the potential mechanisms that underlie these features and help to account for the MCI prevalence and incidence difference. Further, this study found differences in acoustic features between TrailsA and TrailsB, which provides confirmatory evidence that acoustic features are differential for different cognitive domains. TrailsA, as a measure of simple attention compared to the more complex executive functions measured by TrailsB [49], would be expected to have different acoustic features that would be aligned with motor control and perceptual complexity

[50] in the latter and not the former. These differential results suggest that acoustic features might provide a way to detect such subtle differences across cognitive domains. The patterns of acoustic features that are accurately representative of the comprehensive range of cognitive domains will be further explored in future studies.

This study also has some limitations. First, the use of NP tests to diagnose MCI may have led to some circularity and an overestimation of the diagnosis performance [32]. Second, despite that diagnoses are arrived at through a careful adjudication process, there may be some misclassification of MCI. Third, although the FHS collected the voice recordings in a well-controlled environment, there might still be some other factors affecting the quality of voice that were not taken into account. Fourth, this study did not consider linguistic features, which has been shown to be effective in predicting cognitive status. Although we recognize that the inclusion of linguistic features might further improve the prediction of incident MCI, we chose to focus on acoustic features because they are much more generalizable to a broader population, including potentially to other languages. Linguistic features are much more likely to be biased by language, culture, and education. Finally, FHS participants were mostly of European ancestry and English speakers; therefore, the applicability of our findings to populations of another race and ethnicity needs to be examined.

Conclusion

We examined the association of acoustic features with specific cognitive functions—prevalent and incident MCI—in a large community-based population. Overall, this study's establishment of a relationship between MCI risk and human voice features provides foundational evidence for an alternative cognitive assessment approach that is cost-effective and easy to administer for detecting cognition-related disorders. Multiple acoustic features were significantly associated with NP test performance and MCI and could be potentially used as a digital biomarker for early cognitive impairment monitoring.

Acknowledgments

We acknowledge the Framingham Heart Study (FHS) participants for their dedication. This study would not be possible without them. We also thank the researchers in the FHS for their efforts over the years in the examination of subjects.

This work was supported by the National Heart, Lung, and Blood Institute (contract N01-HC-25195) and by grants from the National Institute on Aging (AG-008122, AG-16495, AG-062109, AG-049810, AG-068753, AG054156, and U01AG068221) and the National Institute of Neurological Disorders and Stroke (NS017950). It was also supported by Defense Advanced Research Projects Agency (contract FA8750-16-C-0299) and Pfizer, Inc. This work was also supported by the grants from the Alzheimer's Association (AARG-NTF-20-643020) and the American Heart Association (20SFRN35360180). The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Institutes of Health or the US Department of Health and Human Services. The funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The derived acoustic features could be requested through a formal research application to the Framingham Heart Study [51].

Conflicts of Interest

RA is a scientific advisor to Signant Health and a consultant to Biogen and the Davos Alzheimer's Collaborative.

Multimedia Appendix 1

Supplemental tables and figure.

[\[DOCX File, 288 KB-Multimedia Appendix 1\]](#)

References

1. Sang Z, Wang K, Dong J, Tang L. Alzheimer's disease: updated multi-targets therapeutics are in clinical and in progress. *Eur J Med Chem* 2022 Aug 05;238:114464. [doi: [10.1016/j.ejmech.2022.114464](https://doi.org/10.1016/j.ejmech.2022.114464)] [Medline: [35635955](https://pubmed.ncbi.nlm.nih.gov/35635955/)]
2. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 2020 Aug 08;396(10248):413-446 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30367-6](https://doi.org/10.1016/S0140-6736(20)30367-6)] [Medline: [32738937](https://pubmed.ncbi.nlm.nih.gov/32738937/)]
3. Rosenberg A, Mangialasche F, Ngandu T, Solomon A, Kivipelto M. Multidomain interventions to prevent cognitive impairment, Alzheimer's disease, and dementia: from FINGER to World-Wide FINGERS. *J Prev Alzheimers Dis* 2020 Oct 10;7(1):29-36 [FREE Full text] [doi: [10.14283/jpad.2019.41](https://doi.org/10.14283/jpad.2019.41)] [Medline: [32010923](https://pubmed.ncbi.nlm.nih.gov/32010923/)]
4. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, International Psychogeriatric Association Expert Conference on mild cognitive impairment. Mild cognitive impairment. *Lancet* 2006 Apr 15;367(9518):1262-1270. [doi: [10.1016/S0140-6736\(06\)68542-5](https://doi.org/10.1016/S0140-6736(06)68542-5)] [Medline: [16631882](https://pubmed.ncbi.nlm.nih.gov/16631882/)]
5. Themistocleous C, Eckerström M, Kokkinakis D. Identification of mild cognitive impairment from speech in Swedish using deep sequential neural networks. *Front Neurol* 2018 Nov 15;9:975 [FREE Full text] [doi: [10.3389/fneur.2018.00975](https://doi.org/10.3389/fneur.2018.00975)] [Medline: [30498472](https://pubmed.ncbi.nlm.nih.gov/30498472/)]
6. Morrison MS, Aparicio HJ, Blennow K, Zetterberg H, Ashton NJ, Karikari TK, et al. Ante-mortem plasma phosphorylated tau (181) predicts Alzheimer's disease neuropathology and regional tau at autopsy. *Brain* 2022 Oct 21;145(10):3546-3557. [doi: [10.1093/brain/awac175](https://doi.org/10.1093/brain/awac175)] [Medline: [35554506](https://pubmed.ncbi.nlm.nih.gov/35554506/)]
7. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007 Aug;6(8):734-746. [doi: [10.1016/S1474-4422\(07\)70178-3](https://doi.org/10.1016/S1474-4422(07)70178-3)] [Medline: [17616482](https://pubmed.ncbi.nlm.nih.gov/17616482/)]
8. Chincarini A, Bosco P, Calvini P, Gemme G, Esposito M, Olivieri C, Alzheimer's Disease Neuroimaging Initiative. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. *Neuroimage* 2011 Sep 15;58(2):469-480. [doi: [10.1016/j.neuroimage.2011.05.083](https://doi.org/10.1016/j.neuroimage.2011.05.083)] [Medline: [21718788](https://pubmed.ncbi.nlm.nih.gov/21718788/)]
9. Chételat G, Arbizu J, Barthel H, Garibotto V, Law I, Morbelli S, et al. Amyloid-PET and F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias. *Lancet Neurol* 2020 Nov;19(11):951-962. [doi: [10.1016/S1474-4422\(20\)30314-8](https://doi.org/10.1016/S1474-4422(20)30314-8)] [Medline: [33098804](https://pubmed.ncbi.nlm.nih.gov/33098804/)]
10. Ang TF, An N, Ding H, Devine S, Auerbach SH, Massaro J, et al. Using data science to diagnose and characterize heterogeneity of Alzheimer's disease. *Alzheimers Dement (N Y)* 2019 Jun 27;5(1):264-271 [FREE Full text] [doi: [10.1016/j.trci.2019.05.002](https://doi.org/10.1016/j.trci.2019.05.002)] [Medline: [31304232](https://pubmed.ncbi.nlm.nih.gov/31304232/)]
11. Blennow K, Hampel H, Weiner M, Zetterberg H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* 2010 Mar 16;6(3):131-144. [doi: [10.1038/nrneurol.2010.4](https://doi.org/10.1038/nrneurol.2010.4)] [Medline: [20157306](https://pubmed.ncbi.nlm.nih.gov/20157306/)]
12. Henriksen K, O'Bryant SE, Hampel H, Trojanowski JQ, Montine TJ, Jeromin A, Blood-Based Biomarker Interest Group. The future of blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement* 2014 Jan 15;10(1):115-131 [FREE Full text] [doi: [10.1016/j.jalz.2013.01.013](https://doi.org/10.1016/j.jalz.2013.01.013)] [Medline: [23850333](https://pubmed.ncbi.nlm.nih.gov/23850333/)]
13. Ganguli M, Snitz BE, Saxton JA, Chang CH, Lee C, Vander Bilt J, et al. Outcomes of mild cognitive impairment by definition: a population study. *Arch Neurol* 2011 Jun 13;68(6):761-767 [FREE Full text] [doi: [10.1001/archneurol.2011.101](https://doi.org/10.1001/archneurol.2011.101)] [Medline: [21670400](https://pubmed.ncbi.nlm.nih.gov/21670400/)]
14. Tabatabaei-Jafari H, Shaw ME, Cherbuin N. Cerebral atrophy in mild cognitive impairment: a systematic review with meta-analysis. *Alzheimers Dement (Amst)* 2015 Dec 11;1(4):487-504 [FREE Full text] [doi: [10.1016/j.dadm.2015.11.002](https://doi.org/10.1016/j.dadm.2015.11.002)] [Medline: [27239527](https://pubmed.ncbi.nlm.nih.gov/27239527/)]
15. Seraji-Bzorgzad N, Paulson H, Heidebrink J. Neurologic examination in the elderly. *Handb Clin Neurol* 2019;167:73-88 [FREE Full text] [doi: [10.1016/B978-0-12-804766-8.00005-4](https://doi.org/10.1016/B978-0-12-804766-8.00005-4)] [Medline: [31753158](https://pubmed.ncbi.nlm.nih.gov/31753158/)]
16. Cuetos F, Arango-Lasprilla JC, Uribe C, Valencia C, Lopera F. Linguistic changes in verbal expression: a preclinical marker of Alzheimer's disease. *J Int Neuropsychol Soc* 2007 May;13(3):433-439. [doi: [10.1017/S1355617707070609](https://doi.org/10.1017/S1355617707070609)] [Medline: [17445292](https://pubmed.ncbi.nlm.nih.gov/17445292/)]
17. Deramecourt V, Lebert F, Debachy B, Mackowiak-Cordoliani MA, Bombois S, Kerdraon O, et al. Prediction of pathology in primary progressive language and speech disorders. *Neurology* 2010 Jan 05;74(1):42-49. [doi: [10.1212/WNL.0b013e3181c7198e](https://doi.org/10.1212/WNL.0b013e3181c7198e)] [Medline: [19940270](https://pubmed.ncbi.nlm.nih.gov/19940270/)]
18. Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *J Clin Exp Neuropsychol* 2008 Jul 05;30(5):501-556. [doi: [10.1080/13803390701550128](https://doi.org/10.1080/13803390701550128)] [Medline: [18569251](https://pubmed.ncbi.nlm.nih.gov/18569251/)]
19. Le X, Lancashire I, Hirst G, Jokel R. Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. *Literary and Linguistic Computing* 2011 May 24;26(4):435-461. [doi: [10.1093/lc/fqr013](https://doi.org/10.1093/lc/fqr013)]

20. Beltrami D, Gagliardi G, Rossini Favretti R, Ghidoni E, Tamburini F, Calzà L. Speech analysis by natural language processing techniques: a possible tool for very early detection of cognitive decline? *Front Aging Neurosci* 2018 Nov 13;10:369 [FREE Full text] [doi: [10.3389/fnagi.2018.00369](https://doi.org/10.3389/fnagi.2018.00369)] [Medline: [30483116](https://pubmed.ncbi.nlm.nih.gov/30483116/)]
21. Lin H, Karjadi C, Ang TFA, Prajakta J, McManus C, Alhanai TW, et al. Identification of digital voice biomarkers for cognitive health. *Explor Med* 2020 Dec 31;1:406-417 [FREE Full text] [doi: [10.37349/emed.2020.00028](https://doi.org/10.37349/emed.2020.00028)] [Medline: [33665648](https://pubmed.ncbi.nlm.nih.gov/33665648/)]
22. Au R, Seshadri S, Wolf PA, Elias MF, Elias PK, Sullivan L, et al. New norms for a new generation: cognitive performance in the framingham offspring cohort. *Exp Aging Res* 2004 Oct;30(4):333-358. [doi: [10.1080/03610730490484380](https://doi.org/10.1080/03610730490484380)] [Medline: [15371099](https://pubmed.ncbi.nlm.nih.gov/15371099/)]
23. Wechsler D, Stone CP. Wechsler Memory Scale (WMS). New York, NY: The Psychological Corporation; 1948.
24. Benton AL, Hamsher KD, Sivan AB. Multilingual Aphasia Examination. Iowa City, IA: University of Iowa; 1976.
25. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia, PA: Lea & Febi; 1983.
26. Spreen O, Strauss E. A Compendium of Neuropsychological Tests: Administration, Norms and Commentary. New York, NY: NY Oxford University Press; 1991.
27. Hooper HE. Hooper Visual Organization Test Manual. Los Angeles, CA: Western Psychological Services (WPS); 1958.
28. Eyben F, Wöllmer M, Schuller B. Opensmile: the munich versatile and fast open-source audio feature extractor. 2010 Oct 25 Presented at: MM '10: the 18th ACM international conference on Multimedia; October 25-29, 2010; Firenze, Italy p. 1459-1462. [doi: [10.1145/1873951.1874246](https://doi.org/10.1145/1873951.1874246)]
29. Schuller B, Steidl S, Batliner A, Hirschberg J, Burgoon JK, Baird A, et al. 2016 Presented at: Interspeech 2016; September 8-12, 2016; San Francisco, CA p. 2001-2005. [doi: [10.21437/interspeech.2016-129](https://doi.org/10.21437/interspeech.2016-129)]
30. Weninger F, Eyben F, Schuller BW, Mortillaro M, Scherer KR. On the acoustics of emotion in audio: what speech, music, and sound have in common. *Front Psychol* 2013 May 27;4:292 [FREE Full text] [doi: [10.3389/fpsyg.2013.00292](https://doi.org/10.3389/fpsyg.2013.00292)] [Medline: [23750144](https://pubmed.ncbi.nlm.nih.gov/23750144/)]
31. Tahon M, Devillers L. Towards a small set of robust acoustic features for emotion recognition: challenges. *IEEE/ACM Trans Audio Speech Lang Process* 2016 Jan;24(1):16-28. [doi: [10.1109/taslp.2015.2487051](https://doi.org/10.1109/taslp.2015.2487051)]
32. Yuan J, Libon DJ, Karjadi C, Ang AFA, Devine S, Auerbach SH, et al. Association between the digital clock drawing test and neuropsychological test performance: large community-based prospective cohort (Framingham Heart Study). *J Med Internet Res* 2021 Jun 08;23(6):e27407 [FREE Full text] [doi: [10.2196/27407](https://doi.org/10.2196/27407)] [Medline: [34100766](https://pubmed.ncbi.nlm.nih.gov/34100766/)]
33. Hughes CP, Berg L, Danziger W, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982 Jun 29;140(6):566-572. [doi: [10.1192/bjp.140.6.566](https://doi.org/10.1192/bjp.140.6.566)] [Medline: [7104545](https://pubmed.ncbi.nlm.nih.gov/7104545/)]
34. Haynes W. Wilcoxon rank sum test. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H, editors. *Encyclopedia of Systems Biology*. New York, NY: Springer; 2013:2354-2355.
35. McHugh ML. The chi-square test of independence. *Biochem Med (Zagreb)* 2013;23(2):143-149 [FREE Full text] [doi: [10.11613/bm.2013.018](https://doi.org/10.11613/bm.2013.018)] [Medline: [23894860](https://pubmed.ncbi.nlm.nih.gov/23894860/)]
36. Pinheiro JC, Bates DM. Linear mixed-effects models: basic concepts and examples. In: *Mixed-Effects Models in S and S-Plus*. New York, NY: Springer; 2000:3-56.
37. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *J Am Stat Assoc* 1989 Dec;84(408):1074-1078. [doi: [10.1080/01621459.1989.10478874](https://doi.org/10.1080/01621459.1989.10478874)]
38. Burnham KP, Anderson DR. Multimodel Inference. *Sociological Methods & Research* 2016 Jun 30;33(2):261-304. [doi: [10.1177/0049124104268644](https://doi.org/10.1177/0049124104268644)]
39. Hermansky H, Morgan N. RASTA processing of speech. *IEEE Trans Speech Audio Process* 1994;2(4):578-589. [doi: [10.1109/89.326616](https://doi.org/10.1109/89.326616)]
40. Rabiner L, Schafer R. *Theory and Applications of Digital Speech Processing*. Hoboken, NJ: Prentice Hall Press; 2010.
41. Al-Hameed S, Benaissa M, Christensen H, Mirheidari B, Blackburn D, Reuber M. A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints. *PLoS One* 2019 May 24;14(5):e0217388 [FREE Full text] [doi: [10.1371/journal.pone.0217388](https://doi.org/10.1371/journal.pone.0217388)] [Medline: [31125389](https://pubmed.ncbi.nlm.nih.gov/31125389/)]
42. Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJG. Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: a systematic review article. *Front Psychol* 2021 Mar 23;12:620251 [FREE Full text] [doi: [10.3389/fpsyg.2021.620251](https://doi.org/10.3389/fpsyg.2021.620251)] [Medline: [33833713](https://pubmed.ncbi.nlm.nih.gov/33833713/)]
43. Li J, Yu J, Ye Z, Wong S, Mak M, Mak B, et al. A comparative study of acoustic and linguistic features classification for Alzheimer's disease detection. 2021 May 13 Presented at: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); June 6-11, 2021; Toronto, ON p. 6423-6427. [doi: [10.1109/icassp39728.2021.9414147](https://doi.org/10.1109/icassp39728.2021.9414147)]
44. Gonnerman LM, Aronoff JM, Almor A, Kempler D, Andersen ES. From Beetle to bug: progression of error types in naming in Alzheimer's disease. *Proceedings of the Annual Meeting of the Cognitive Science Society* 2004;26(26):1563 [FREE Full text]
45. Nagumo R, Zhang Y, Ogawa Y, Hosokawa M, Abe K, Ukeda T, et al. Automatic detection of cognitive impairments through acoustic analysis of speech. *Curr Alzheimer Res* 2020 Mar 20;17(1):60-68 [FREE Full text] [doi: [10.2174/1567205017666200213094513](https://doi.org/10.2174/1567205017666200213094513)] [Medline: [32053074](https://pubmed.ncbi.nlm.nih.gov/32053074/)]

46. Meilán JJG, Martínez-Sánchez F, Carro J, López DE, Millian-Morell L, Arana JM. Speech in Alzheimer's disease: can temporal and acoustic parameters discriminate dementia? *Dement Geriatr Cogn Disord* 2014 Jan 30;37(5-6):327-334. [doi: [10.1159/000356726](https://doi.org/10.1159/000356726)] [Medline: [24481220](https://pubmed.ncbi.nlm.nih.gov/24481220/)]
47. Mittal V, Sharma RK. Machine learning approach for classification of Parkinson disease using acoustic features. *J Reliable Intell Environ* 2021 Apr 26;7(3):233-239. [doi: [10.1007/s40860-021-00141-6](https://doi.org/10.1007/s40860-021-00141-6)]
48. Perez M, Jin W, Le D, Carlozzi N, Dayalu P, Roberts A, et al. Classification of Huntington disease using acoustic and lexical features. *Interspeech* 2018;2018:1898-1902 [FREE Full text] [doi: [10.21437/interspeech.2018-2029](https://doi.org/10.21437/interspeech.2018-2029)] [Medline: [33241056](https://pubmed.ncbi.nlm.nih.gov/33241056/)]
49. Crowe SF. The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. *J Clin Psychol* 1998 Aug;54(5):585-591. [doi: [10.1002/\(sici\)1097-4679\(199808\)54:5<585::aid-jclp4>3.0.co;2-k](https://doi.org/10.1002/(sici)1097-4679(199808)54:5<585::aid-jclp4>3.0.co;2-k)] [Medline: [9696108](https://pubmed.ncbi.nlm.nih.gov/9696108/)]
50. Arbuthnott K, Frank J. Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *J Clin Exp Neuropsychol* 2000 Aug;22(4):518-528. [doi: [10.1076/1380-3395\(200008\)22:4:1-0:FT518](https://doi.org/10.1076/1380-3395(200008)22:4:1-0:FT518)] [Medline: [10923061](https://pubmed.ncbi.nlm.nih.gov/10923061/)]
51. For researchers. Framingham Heart Study. URL: <https://www.framinghamheartstudy.org/fhs-for-researchers/>

Abbreviations

AD: Alzheimer disease

AUC: area under the curve

FAS: Controlled Oral Word Association Test

FHS: Framingham Heart Study

HVOT: Hooper Visual Organization Test

LLD: low-level descriptor

LMi: Logical Memory—Immediate Recall

MCI: mild cognitive impairment

MFCC: Mel-frequency cepstral coefficient **VRi:** Visual Reproduction—Immediate Recall

NP: neuropsychological

OR: odds ratio

ROC: receiver operating characteristic

SIM: Similarities

TrailsB: Trail Making Test B

VRd: Visual Reproduction—Delayed Recall

VRr: Visual Reproduction—Recognition

Edited by G Eysenbach, T Leung; submitted 22.09.22; peer-reviewed by I van der Linde, J Chan; comments to author 25.10.22; revised version received 29.11.22; accepted 03.12.22; published 22.12.22

Please cite as:

Ding H, Mandapati A, Karjadi C, Ang TFA, Lu S, Miao X, Glass J, Au R, Lin H

Association Between Acoustic Features and Neuropsychological Test Performance in the Framingham Heart Study: Observational Study

J Med Internet Res 2022;24(12):e42886

URL: <https://www.jmir.org/2022/12/e42886>

doi: [10.2196/42886](https://doi.org/10.2196/42886)

PMID:

©Huitong Ding, Amiya Mandapati, Cody Karjadi, Ting Fang Alvin Ang, Sophia Lu, Xiao Miao, James Glass, Rhoda Au, Honghuang Lin. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 22.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.