



OPEN ACCESS

EDITED BY

Ana Capuano,
Rush University, United States

REVIEWED BY

Maude Wagner,
Rush University, United States
Jeannette Simino,
University of Mississippi Medical Center,
United States

*CORRESPONDENCE

Honghuang Lin
✉ honghuang.lin@umassmed.edu

RECEIVED 30 April 2023

ACCEPTED 27 October 2023

PUBLISHED 23 November 2023

CITATION

Ding H, Hamel AP, Karjadi C, Ang TFA, Lu S, Thomas RJ, Au R and Lin H (2023) Association between acoustic features and brain volumes: the Framingham Heart Study. *Front. Dement.* 2:1214940. doi: 10.3389/frdem.2023.1214940

COPYRIGHT

© 2023 Ding, Hamel, Karjadi, Ang, Lu, Thomas, Au and Lin. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Association between acoustic features and brain volumes: the Framingham Heart Study

Huitong Ding^{1,2}, Alexander P. Hamel³, Cody Karjadi², Ting F. A. Ang^{1,2,4,5}, Sophia Lu⁵, Robert J. Thomas⁶, Rhoda Au^{1,2,4,5,7} and Honghuang Lin^{3*}

¹Department of Anatomy and Neurobiology, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States, ²The Framingham Heart Study, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States, ³Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States, ⁴Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States, ⁵Slope Epidemiology Center, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States, ⁶Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, United States, ⁷Departments of Neurology and Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States

Introduction: Although brain magnetic resonance imaging (MRI) is a valuable tool for investigating structural changes in the brain associated with neurodegeneration, the development of non-invasive and cost-effective alternative methods for detecting early cognitive impairment is crucial. The human voice has been increasingly used as an indicator for effectively detecting cognitive disorders, but it remains unclear whether acoustic features are associated with structural neuroimaging.

Methods: This study aims to investigate the association between acoustic features and brain volume and compare the predictive power of each for mild cognitive impairment (MCI) in a large community-based population. The study included participants from the Framingham Heart Study (FHS) who had at least one voice recording and an MRI scan. Sixty-five acoustic features were extracted with the OpenSMILE software (v2.1.3) from each voice recording. Nine MRI measures were derived according to the FHS MRI protocol. We examined the associations between acoustic features and MRI measures using linear regression models adjusted for age, sex, and education. Acoustic composite scores were generated by combining acoustic features significantly associated with MRI measures. The MCI prediction ability of acoustic composite scores and MRI measures were compared by building random forest models and calculating the mean area under the receiver operating characteristic curve (AUC) of 10-fold cross-validation.

Results: The study included 4,293 participants (age 57 ± 13 years, 53.9% women). During 9.3 ± 3.7 years follow-up, 106 participants were diagnosed with MCI. Seven MRI measures were significantly associated with more than 20 acoustic features after adjusting for multiple testing. The acoustic composite scores can improve the AUC for MCI prediction to 0.794, compared to 0.759 achieved by MRI measures.

Discussion: We found multiple acoustic features were associated with MRI measures, suggesting the potential for using acoustic features as easily accessible digital biomarkers for the early diagnosis of MCI.

KEYWORDS

mild cognitive impairment, digital voice, brain volume, association, prediction

1 Introduction

Mild Cognitive Impairment (MCI) represents a stage of cognitive impairment, during which cognitive decline does not significantly affect daily functioning (Gauthier et al., 2006). Individuals with MCI may experience difficulty with executive function and remembering events (Themistocleous et al., 2018). Currently, there are no definitive disease-modifying treatments available (Sang et al., 2022). However, it is widely agreed that early detection is critical. Interventions aimed at reducing modifiable risk factors such as blood pressure control and optimal physical exercise have the potential to delay, attenuate, or even prevent disease onset and/or progression (Livingston et al., 2020; Rosenberg et al., 2020). Therefore, detecting MCI is vital so that interventions targeting the neurodegenerative process, such as clinical trials, may be initiated to help uncover potential treatment plans (Morrison et al., 2022).

Brain magnetic resonance imaging (MRI) is a useful tool for investigating structural changes in the brain that are associated with neurodegeneration, including MCI (Ries et al., 2008). Multiple MRI measures are found to be associated with the pathology and progression of cognitive impairment (Chen and Herskovits, 2010; Del Sole et al., 2016; Graham and Sharp, 2019; Zhu et al., 2021). By detecting subtle changes in brain volume, MRI can help identify individuals who are at greater risk of developing MCI (Fennema-Notestine et al., 2009). However, the cost of MRI and the need for easy serial testing limits its adoption in low-resource clinical settings or settings where imaging technologies may be limited. In the United States, for instance, MRI scans have an average cost of \$1,325, with prices varying from \$375 to \$2,850 (Prudenzi et al., 2019). Therefore, it is important to develop alternative methods for detecting early cognitive impairment using non-invasive and cost-effective techniques which measure specific brain outputs and which can ideally be captured relatively passively and be automated.

Communication through vocalization is a key human characteristic, and engages a number of complex brain networks. The human voice is an easily accessible and non-invasive method of collecting data that has gained interest as a potential tool for detecting cognitive decline (Ding et al., 2022). Speech production is a highly complex cognitive task (Seraji-Bzorgzad et al., 2019), and recording speech is easily achievable with the availability of recording devices. Vocal output is modified by numerous conditions including as examples affect, alertness/sleepiness, dyspnea, and structural or functional abnormalities from the cortex to the vocal-articulatory complex. Language deficits have been found to occur in the prodromal stages of cognitive impairment (Cuetos et al., 2007), which may occur years before clinical diagnosis (Taler and Phillips, 2008; Deramecourt et al., 2010), potentially making voice-based assessment a promising indicator for MCI. Meanwhile, recent advancements in speech feature extraction technology enable the quantification of voice signal properties from multiple dimensions, enabling a comprehensive description of specific pathologies through voice features. Previous research has demonstrated the association of acoustic features with neuropsychological tests and MCI (Ding et al., 2022). Moreover, linguistic changes have been associated with specific brain regions, such as atrophy in the hippocampus (Ramos-Escobar et al., 2022), temporoparietal regions (Grossman et al., 1997),

and speech motor control networks (Kearney and Guenther, 2019). However, the relationship between acoustic features and MRI measures remains understudied. Investigating the association between these two modalities can provide a deeper understanding of neurodegeneration, complementing the structural information provided by MRI with the functional information conveyed by voice features. Furthermore, leveraging voice-based biomarkers as a screening method can provide a more economical alternative for MCI screening, making it a valuable complement to MRI-based assessments.

The objective of this study is to investigate the association between acoustic features and MRI measures in the Framingham Heart Study (FHS). We further explore the potential to incorporate acoustic features in the prediction of incident MCI.

2 Materials and methods

2.1 Sample selection

The FHS is a community-based prospective cohort study that has been conducted since 1948, with details on the FHS cohorts previously reported in publications (Wolf, 2012; Mahmood et al., 2014; Tsao and Vasani, 2015). Cognitive testing was introduced as part of the FHS in 1976, and in 1999, it became routine to recruit participants for standardized neuropsychological (NP) assessments, that also included a concomitant MRI scan. For the current study, we included participants who had at least one voice-recorded NP assessment and a contemporary MRI scan within 1 year from 2005 to 2017. We excluded those whose voice recording was less than 10 min in length ($n = 8$), and those with missing education information ($n = 8$). To evaluate the added predictability of the acoustic composite score for incident MCI, we also excluded participants who were below 60 years old at the time of voice recording ($n = 2,459$), those with prevalent MCI or dementia ($n = 145$), and those who were flagged as potential MCI but had not gone through dementia review ($n = 142$). All procedures and protocols of the FHS were approved by the Institutional Review Board of the Boston University Medical Campus, and written informed consent was obtained from all participants.

2.2 Voice recordings

Since 2005, the FHS has digitally recorded all verbal interactions between the tester and the participant during administration of NP tests as well as the participant's spoken responses to neuropsychological test questions. A sliding window approach was used to divide each recording into 20-ms segments with a shifting size of 10 ms (Luz et al., 2021; Dumpala et al., 2022). These segments were then analyzed using OpenSMILE software (v2.1.3) (Eyben et al., 2010) to extract a set of 65 low-level descriptor (LLD) features (Schuller et al., 2016), which include pitch, voice quality, loudness, signal energy, waveform, auditory, FFT spectrum, spectral, and cepstral. For each recording, the mean of each LLD feature was computed to capture its high-level statistical features. Then, normalization was performed by

subtracting the mean and dividing by the standard deviation. These features have demonstrated great performance across different tasks, such as speech processing, music information retrieval, and emotion recognition (Tahon and Devillers, 2015). A summary of these acoustic features is provided in Supplementary Table 1 and the previous publication (Weninger et al., 2013).

2.3 MRI data collection and preprocess

The FHS MRI protocol has been described previously (Thomas et al., 2021). Briefly, participants were imaged using a Siemens 1.5T field strength machine (Siemens Medical) with a 3-dimensional T1- and T2-weighted coronal spoiled gradient-recalled echo sequence. All images were centrally processed at University of California Davis Medical Center with standardized brain structural MRI segmentation procedures (Rajapakse et al., 1996; Fletcher et al., 2012). An expectation-maximization algorithm was used to perform segmentation of gray matter, white matter, and cerebrospinal fluid following skull stripping and intensity inhomogeneity correction. Segmentation of the hippocampus was performed utilizing a standard atlas hippocampal segmentation algorithm (Vercauteren et al., 2007; Boccardi et al., 2014, 2015; Bocchetta et al., 2015). Established procedures were utilized to perform segmentation of white matter hyperintensity (WMH) (Rajapakse et al., 1996; Fletcher et al., 2012). Total cerebral cranial volume (TCV) was determined by outlining the intracranial vault lying above the tentorium and was used for correcting head size (Smith et al., 2008; Aljabar et al., 2009; DeStefano et al., 2009; Jefferson et al., 2010; Spartano et al., 2019).

This study included the following MRI measures: total cerebral brain volume (TCBV), cerebral white matter volume (CWMV), cerebral gray matter volume (CGMV), hippocampal volume (HV), cortical gray matter (CGM), segmented frontal lobe gray matter volume (FLGMV), segmented parietal lobe gray matter volume (PLGMV), segmented temporal lobe gray matter volume (TLGMV), and segmented occipital lobe gray matter volume (OLGMV). All MRI measures were represented as the percentage of these volumes over the TCV to correct for head size difference (DeCarli et al., 2005).

2.4 Ascertainment of mild cognitive impairment

The cognitive ascertainment procedures utilized in the FHS have been thoroughly described (Seshadri et al., 1997). NP tests are the principal measures used to evaluate the cognitive status of FHS participants. For those who showed signs of possible cognitive impairment, NP tests were administered on average every 1–2 years. If cognitive decline was detected, a clinical review was conducted by a panel consisting of at least one neurologist and one neuropsychologist. The review panel diagnosed MCI based on if a participant showed evidence of cognitive performance decline in at least one cognitive domain, showed no evidence of functional decline, and did not meet criteria for dementia (Yuan et al., 2021). To measure the extent of impairment, a severity rating

was provided that is similar in objective as the Clinical Dementia Rating scale (Hughes et al., 1982). The primary outcome of this study was incident MCI, which is defined as individuals who were cognitively intact at the time of voice recording but later diagnosed with MCI.

2.5 Statistical analyses

This study used the Wilcoxon rank-sum test for continuous variables (Haynes, 2013) and the Chi-squared test for categorical variables (McHugh, 2013) to compare the difference in demographics and MRI measures between incident MCI and normal control (NC) groups. Linear regression models were further used to assess the association between each acoustic feature and MRI measures (Pinheiro and Bates, 2000). To adjust for multiple comparisons, given the total number of acoustic features tested against each MRI measure, we employed the Bonferroni correction method (Armstrong, 2014), and the corrected significance threshold was defined as $P = 0.05/65 \approx 7.7E-04$ given that 65 acoustic features were considered.

A set of acoustic composite scores was generated for MRI measures as a weighted combination of acoustic features that were found to be significantly associated with the MRI measure. The weight assigned to each acoustic feature in the composite score was established through the training of a linear regression model. For a given participant i , their acoustic composite score of an MRI measure was calculated using the following formula:

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

Here, m refers to the count of acoustic features that exhibit a significant association with the MRI measure. The estimate of effect size for acoustic feature j obtained from the linear regression model is represented by α_j , while V_{ij} denotes the normalized acoustics feature j for participant i . All models were adjusted for age, sex, and education.

Random forest models were then developed to assess the model performance in terms of the area under the receiver operating characteristics curve (AUC). Three models were compared: a baseline model using age, sex, and education as predictors; a second model using age, sex, education, and 9 MRI measures; and a third using age, sex, education, and acoustic composite scores as predictors. The mean AUC of 10-fold cross-validation was calculated for each model. We further conducted a sensitivity analysis to evaluate the stability of the prediction performance by constructing two additional models: one using only MRI measures and another using only the acoustic composite score. All statistical analyses were conducted using Python (version 3.9.7).

3 Results

Our study included 4,293 participants of FHS (mean baseline age 57 ± 13 years; 53.9% women; 57.1% self-reported college educated or higher). The details of sample characteristics are shown in Table 1.

TABLE 1 Sample characteristics.

Variable	Association analysis	Prediction analysis		P-value***
	Total (n = 4,293)	Incident MCI (n = 106)	Referents** (n = 1,441)	
Age (years), mean (SD)	57 (13)	76 (8)	68 (7)	<0.001
Gender, n (%)				0.523
Women	2,314 (53.9)	53 (50.0)	774 (53.7)	
Men	1,979 (46.1)	53 (50.0)	667 (46.3)	
Education, n (%)				0.001
No high school	89 (2.1)	6 (5.7)	38 (2.6)	<0.001
High school	767 (17.9)	33 (31.1)	313 (21.7)	<0.001
Some college	986 (23.0)	32 (30.2)	341 (23.7)	<0.001
College and higher	2,451 (57.1)	35 (33.0)	749 (52.0)	<0.001
MRI measures*, median (IQR)				
Total cerebral brain volume (%)	77.99 (75.82–79.56)	73.63 (72.00–75.41)	76.01 (74.48–77.62)	<0.001
Cerebral white matter volume (%)	36.92 (35.31–38.35)	34.42 (32.93–36.09)	36.11 (34.49–37.67)	<0.001
Cerebral gray matter volume (%)	40.71 (39.15–42.09)	38.46 (37.44–39.39)	39.54 (38.39–40.74)	<0.001
Hippocampal volume (%)	0.54 (0.51–0.57)	0.53 (0.50–0.56)	0.54 (0.51–0.58)	<0.001
Cortical gray matter volume (%)	37.27 (35.77–38.57)	35.10 (34.26–36.03)	36.15 (35.03–37.30)	<0.001
Segmented frontal lobe gray matter volume (%)	14.42 (13.75–15.05)	13.55 (13.19–14.02)	13.91 (13.37–14.49)	<0.001
Segmented parietal lobe gray matter volume (%)	7.98 (7.63–8.33)	7.58 (7.26–7.93)	7.83 (7.47–8.15)	<0.001
Segmented temporal lobe gray matter volume (%)	9.94 (9.54–10.33)	9.39 (9.09–9.84)	9.73 (9.35–10.08)	<0.001
Segmented occipital lobe gray matter volume (%)	4.87 (4.54–5.18)	4.57 (4.25–4.75)	4.70 (4.41–5.02)	<0.001

*All MRI measures were corrected for head size by calculating the percentage of the volumes over the total cerebral cranial volume above the tentorium.

**The referents are the participants who remained cognitively intact throughout the follow-up period.

***The P-value was calculated by the Student's *t*-test for age, Wilcoxon rank-sum test for MRI measures due to skewed distribution, and Chi-squared test for categorical variables.

The distribution metrics for each acoustic feature, encompassing min, 25% quantile, median, 75% quantile, and max, are outlined in [Supplementary Table 2](#). Their interrelationships are shown in a correlation heatmap found in [Supplementary Figure 1](#). We examined the association of acoustic features with MRI measures. As shown in [Tables 2, 3](#), seven MRI measures (CWMV, CGMV, HV, CGM, PLGMV, TLGMV, and OLG MV) were significantly associated with over 20 acoustic features after Bonferroni correction ($P < 7.7E-04$). Cerebral gray matter volume was significantly associated with 47 acoustic features. The acoustic feature, voicingFinalUnclipped, which represents the voicing probability of the final fundamental frequency candidate, was the most significantly associated feature with 4 MRI gray matter measures (CGMV, CGM, TLGMV, and OLG MV). A larger voicingFinalUnclipped, for example, was strongly associated with a smaller segmented occipital lobe gray matter volume (OLGMV) ($P = 3.57E-22$). The feature, pcm_fftMag_spectralKurtosis, which quantifies the spectral shape or distribution of audio signal energy, was most significantly associated with total cerebral brain volume (TCBV). Similarly, the feature, audSpec_Rfilt, which captures crucial aspects of the spectral content and structure of audio signals as perceived by the human auditory system, was most significantly associated with cerebral white matter volume (CWMV). Additionally, pcm_fftMag_spectralSkewness was the

most significant acoustic feature associated with hippocampal volume. It represents the shape or distribution of the signal's energy across different frequency bands. A comprehensive overview of the associations between acoustic features and MRI measures is shown in [Supplementary Tables 3–11](#). In the sensitivity analysis, we further included 98 participants with prevalent stroke to examine the association between acoustic features and MRI measures. As shown in [Supplementary Table 12](#), similar acoustic features were found to associate with MRI measures. We also excluded the participants who were younger than 60 years and examined the association between acoustic features and MRI measures ([Supplementary Table 12](#)). About half of the associations remained significant. In addition, we found 4 associations were only observed in old people, suggesting potential distinct patterns between acoustic features and neuroimaging features in old people.

We further built a composite score from these significant acoustic features for each MRI measure. As expected, these composite scores were all significantly associated with each corresponding MRI measure. We further evaluated the added predictive power of 9 acoustic composite scores for incident MCI. The analysis was limited to 1,547 participants who were at least 60 years at the time of voice recordings. Among them, 106 were diagnosed with MCI during an average of 9.3 ± 3.7 years of follow-up. For the referent group, the baseline median Mini-Mental

TABLE 2 The most significant acoustic feature for each MRI measure.

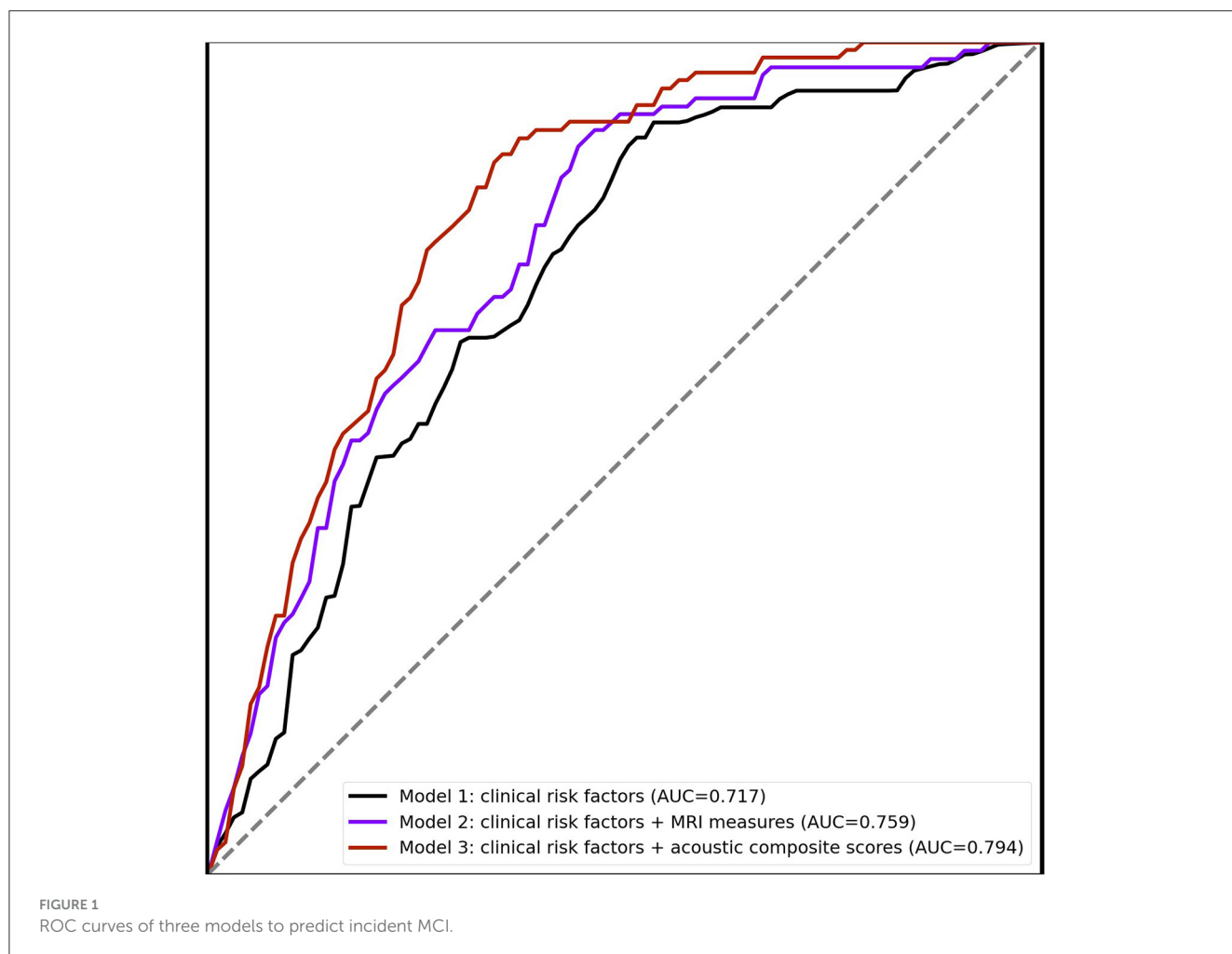
MRI measures	Number of significant acoustic features	Association of most significant acoustic feature				
		Most significant feature	Description	Effect size	Standard error	P-value*
Total cerebral brain volume (TCBV)	10	pcm_fftMag_spectralKurtosis_sma	Magnitude of spectral kurtosis	0.1132	0.0149	3.30E-14
Cerebral white matter volume (CWMV)	44	audSpec_Rfilt_sma[5]	RASTA-style filtered auditory spectrum, band 6	0.1099	0.0154	1.27E-12
Cerebral gray matter volume (CGMV)	47	voicingFinalUnclipped_sma	The voicing probability of the final fundamental frequency candidate	-0.2121	0.0193	1.06E-27
Hippocampal volume (HV)	36	pcm_fftMag_spectralSkewness_sma	Magnitude of spectral skewness	0.1473	0.0197	8.75E-14
Cortical gray matter volume (CGM)	36	voicingFinalUnclipped_sma	The voicing probability of the final fundamental frequency candidate	-0.1763	0.0201	2.13E-18
Segmented frontal lobe gray matter volume (FLGMV)	10	F0final_sma	The fundamental frequency computed from the Cepstrum	-0.0555	0.0132	2.52E-05
Segmented parietal lobe gray matter volume (PLGMV)	22	audSpec_Rfilt_sma[5]	RASTA-style filtered auditory spectrum, band 6	-0.0873	0.0152	9.65E-09
Segmented temporal lobe gray matter volume (TLGMV)	31	voicingFinalUnclipped_sma	The voicing probability of the final fundamental frequency candidate	-0.1803	0.0232	8.51E-15
Segmented occipital lobe gray matter volume (OLGMV)	27	voicingFinalUnclipped_sma	The voicing probability of the final fundamental frequency candidate	-0.2336	0.0240	3.57E-22

*Linear regression models were used to assess the association between each acoustic feature and MRI measures adjusted for age, sex, and education. Significant associations were claimed if $P < 0.05/65 \approx 7.7E-04$.

TABLE 3 The associations of acoustic composite scores with MRI measures.

MRI measures	Association with acoustic composite score		
	Effect size	Standard error	P-value*
Total cerebral brain volume (TCBV)	0.1183	0.0102	1.33E-30
Cerebral white matter volume (CWMV)	0.1733	0.0152	8.30E-30
Cerebral gray matter volume (CGMV)	0.2233	0.0117	3.22E-78
Hippocampal volume (HV)	0.1710	0.0166	1.15e-24
Cortical gray matter volume (CGM)	0.1888	0.0122	2.03E-52
Segmented frontal lobe gray matter volume (FLGMV)	0.0788	0.0133	3.46E-09
Segmented parietal lobe gray matter volume (PLGMV)	0.1293	0.0149	6.99E-18
Segmented temporal lobe gray matter volume (TLGMV)	0.1827	0.0140	5.63E-38
Segmented occipital lobe gray matter volume (OLGMV)	0.2292	0.0143	4.05E-56

*Linear regression models were used to assess the association between each acoustic composite score and MRI measures adjusted for age, sex, and education.



State Examination (MMSE) score is 29 with an interquartile range (IQR) of 2. For the MCI group, the baseline median MMSE score is 29 with an IQR of 3. We built three prediction models based on random forest. Figure 1 shows that the AUC of MCI prediction can be improved from 0.717 (Model 1) to 0.759 (Model 2) by including 9 MRI measures with risk factors. The model

with clinical risk factors and acoustic composite scores (Model 3) can further improve performance of MCI prediction to AUC 0.794. In the sensitivity analysis, we also built models solely based on MRI measures or acoustic composite scores, which reached an AUC of 0.721 and 0.687, respectively. The AUC values for predicting incident MCI, based on clinical risk factors combined

with each distinct acoustic composite score, are presented in [Supplementary Table 13](#). The composite score derived from the segmented temporal lobe gray matter volume exhibited the highest predictive performance for MCI with an AUC of 0.808. We further include APOE genotype, diabetes, and hypertension as additional clinical risk factors in the three models. The model with acoustic composite scores continued to show the best performance of MCI prediction (AUC 0.795) ([Supplementary Figure 2](#)).

4 Discussion

Although MRI measures have been used as an important biomarker of neurodegeneration, approximately 70% of the global population has limited or no access to MRI technology ([Liu et al., 2021](#)). Therefore, it would be interesting to explore human voice as a non-invasive and cost-effective alternative to detect early cognitive impairment. We examined the relationship between acoustic features and MRI measures on a large community-based cohort, and found significant associations between many acoustic features and gray and white matter volumetric MRI measures. The performance of the model with only acoustic composite scores and clinical risk factors reached an AUC of 0.794 to predict incident MCI. Compared to the burden and cost of conducting MRI scan, the prediction model based on acoustic features is a more cost-effective solution. These results suggest the feasibility of using voice as a potential biomarker for cognitive health screening.

Speech production is a complex process that involves several brain regions. The primary motor cortex, located in frontal lobe, controls the movements of the articulators, such as the lips and tongue ([Simonyan and Horwitz, 2011](#)). Consistently, this study found that 7 acoustic features were associated with segmented frontal lobe gray matter volume. Previous studies have also shown that the gray matter volume of the right and left temporal lobes play an important role in language processing and speech production ([Pihlajamäki et al., 2000](#); [Hickok and Poeppel, 2004](#); [Price, 2010](#)). Notably, this study found that multiple gray matter regions were associated with acoustic features, suggesting a more comprehensive connection between gray matter volume and speech production. The most significantly associated acoustic feature with multiple MRI measures was voicingFinalUnclipped, which quantifies the sound quality of an individual's speech. This feature can provide information about the timing and coordination of vocal cord movement during speech production. Considering speech production involves multiple brain regions working together in a coordinated manner, these results may be useful for intriguing hypotheses about speech mechanism for future validation.

Our results extend the current body of evidence supporting the predictive ability of human voice for incident MCI. The added predictive ability of acoustic features was evaluated by constructing random forest models with baseline features and acoustic composite scores. These acoustic composite scores were created to provide a consolidated reflection of multiple acoustic features, potentially offering a more comprehensive insight into the underlying neurobiological alterations represented by MRI measure. The utilization of composite scores presents several

advantages. It allows for the reduction of dimensionality, mitigating the risk of overfitting, especially in cases where multiple correlated features are present. Moreover, by condensing information from various features into a single composite score, we can achieve a more robust and generalized representation of the data, enhancing the interpretability of the results, especially in the context of population-based estimates. The model with baseline features and nine acoustic composite scores achieved an AUC of 0.794 for incident MCI prediction. However, the models relying solely on MRI measures or acoustic composite scores showed inferior performance, suggesting that clinical risk factors play a vital role in the prediction models. The ability to monitor acoustic features remotely offers a more convenient way to assess cognitive health. Moreover, the easy acquisition of voice in daily life makes it an ideal tool for long-term monitoring of cognitive status. However, there is a lack of research about the relationship between acoustic features and brain structure. Given the rich information from human voice and the cost-effectiveness of voice recording, our study suggests that acoustic features might serve as a new data modality to detect nuanced changes in cognition.

Strengths of this study include that the association between acoustic features and MRI measures was examined in participants from a community-based cohort with a diverse range of ages and health conditions. 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 and prospectively reveals a temporal relationship between acoustic features and MCI. The use of acoustic features as a biomarker for cognitive impairment could provide a valuable tool for clinicians to screen patients for cognitive decline, especially in settings where imaging technologies such as MRI are not readily available. Moreover, the utilization of acoustic features via remote/digital technology, such as a smartphone application that participants can speak into, enables clinicians to detect MCI outside of clinical settings and effectively reduce the cost of detection. Beyond the clinical settings, it provided the ability to use remote/digital technology (i.e., a smartphone app that a patient speaks into) to help clinicians detect MCI and effectively lower the cost of detection. Additionally, such an approach could be used to track the progression of cognitive decline over time and potentially monitor the effectiveness of treatments.

This study also has several limitations. First, it is important to note that despite a rigorous adjudication process for MCI diagnoses, there remains the possibility of misclassifications. Second, voice recordings were collected in a well-controlled environment; therefore it is unclear whether the results would hold based on voice from daily communications. Third, due to the cross-sectional nature of association analysis, we could not get the causality relationship between voice and brain structure. Affective state and sleepiness/alertness are other factors which can intuitively impact voice characteristics, and may impact analysis positively or negatively—as such modulation may be transient or may alternatively amplify MCI-related change. Another limitation of this study is that the observed associations between acoustic features and MRI measures could be influenced by the normal aging process. This is because the regression analyses were

performed across the entire cohort, and the included brain regions predominantly reflect global atrophy rather than specific acoustic processes. Besides, a limitation in comparing the methods is that the acoustic composite scores were formulated based on MRI measures, rather than being ascertained independently from MRI data. Finally, FHS participants were mostly of European ancestry and English speakers; therefore, the applicability of our findings to populations of other ethnicities and languages needs to be examined. It should be expected that different languages and dialects, or heavily accented vocal outputs will pose tractable challenges. External validation is imperative to substantiate our findings before they can be broadly applied or generalized.

In summary, we examined the association of acoustic features with MRI measures in a large community-based cohort. While more research is needed to fully understand the relationship between acoustic features with MRI measures, this study provides evidence that acoustic features might be used as potential biomarkers to assess future MCI risk.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the datasets analyzed for this study could be requested through a formal research application to the Framingham Heart Study. Requests to access these datasets should be directed to <https://www.framinghamheartstudy.org/fhs-for-researchers/>.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of the Boston University Medical Campus. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

HD and HL contributed to the study design. HD performed the data analysis and drafted the manuscript. HD, HL, RA, RT, and AH contributed to the manuscript preparation. All authors critically reviewed the manuscript and have approved the final manuscript.

Funding

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, U01AG068221, and from the National Institute of Neurological Disorders and Stroke, NS017950. It was also supported by Defense Advanced Research Projects Agency contract (FA8750-16-C-0299); 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 funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

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

Conflict of interest

RA is a scientific advisor to Signant Health and a consultant to Biogen and the Davos Alzheimer's Collaborative. RT: (1) received royalties through the Beth Israel Deaconess Medical Center for a patent (ECG-spectrogram) licensed to MyCardio, LLC; (2) Consults in the general area of Sleep Medicine for GLG Councils and Guidepoint Global.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frdem.2023.1214940/full#supplementary-material>

References

- Aljabar, P., Heckemann, R. A., Hammers, A., Hajnal, J. V., and Rueckert, D. (2009). Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage* 46, 726–738. doi: 10.1016/j.neuroimage.2009.02.018
- Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalm. Physiol. Opt.* 34, 502–508. doi: 10.1111/opo.12131
- Boccardi, M., Bocchetta, M., Apostolova, L. G., Barnes, J., Bartzokis, G., Corbetta, G., et al. (2015). Delphi definition of the EADC-ADNI harmonized protocol for hippocampal segmentation on magnetic resonance. *Alzheimers Dement.* 11, 126–138. doi: 10.1016/j.jalz.2014.02.009
- Boccardi, M., Bocchetta, M., Apostolova, L. G., Preboske, G., Robitaille, N., Pasqualetti, P., et al. (2014). Establishing magnetic resonance images orientation for the EADC-ADNI manual hippocampal segmentation protocol. *J. Neuroimaging* 24, 509–514. doi: 10.1111/jon.12065
- Bocchetta, M., Boccardi, M., Ganzola, R., Apostolova, L. G., Preboske, G., Wolf, D., et al. (2015). Harmonized benchmark labels of the hippocampus on magnetic resonance: the EADC-ADNI project. *Alzheimers Dement.* 11, 151.e5–160.e5. doi: 10.1016/j.jalz.2013.12.019
- Chen, R., and Herskovits, E. H. (2010). Machine-learning techniques for building a diagnostic model for very mild dementia. *Neuroimage* 52, 234–244. doi: 10.1016/j.neuroimage.2010.03.084
- Cuetos, F., Arango-Lasprilla, J. C., Uribe, C., Valencia, C., and Lopera, F. (2007). Linguistic changes in verbal expression: a preclinical marker of Alzheimer's disease. *J. Int. Neuropsychol. Soc.* 13, 433–439. doi: 10.1017/S1355617707070609
- DeCarli, C., Massaro, J., Harvey, D., Hald, J., Tullberg, M., Au, R., et al. (2005). Measures of brain morphology and infarction in the framingham heart study: establishing what is normal. *Neurobiol. Aging* 26, 491–510. doi: 10.1016/j.neurobiolaging.2004.05.004
- Del Sole, A., Malaspina, S., and Biasina, A. M. (2016). Magnetic resonance imaging and positron emission tomography in the diagnosis of neurodegenerative dementias. *Funct. Neurol.* 31, 205. doi: 10.11138/FNeur/2016.31.4.205
- Deramecourt, V., Lebert, F., Debachy, B., Mackowiak-Cordoliani, M., Bombois, S., Kerdraon, O., et al. (2010). Prediction of pathology in primary progressive language and speech disorders. *Neurology* 74, 42–49. doi: 10.1212/WNL.0b013e3181c7198e
- DeStefano, A. L., Seshadri, S., Beiser, A., Atwood, L. D., Massaro, J. M., Au, R., et al. (2009). Bivariate heritability of total and regional brain volumes: the Framingham Study. *Alzheimer Dis. Assoc. Disord.* 23, 218. doi: 10.1097/WAD.0b013e31819cadd8
- Ding, H., Mandapati, A., Karjadi, C., Ang, T. F. A., Lu, S., Miao, X., et al. (2022). Association between acoustic features and neuropsychological test performance in the Framingham Heart Study: observational study. *J. Med. Internet Res.* 24, e42886. doi: 10.2196/42886
- Dumpala, S. H., Rodriguez, S., Rempel, S., Sajjadian, M., Uher, R., and Oore, S. (2022). "Detecting depression with a temporal context of speaker embeddings," in *Proc AAAI SAS*.
- Eyben, F., Wöllmer, M., and Schuller, B. (eds.). (2010). "Opensmile: the munich versatile and fast open-source audio feature extractor," in *Proceedings of the 18th ACM international conference on Multimedia*. doi: 10.1145/1873951.1874246
- Fennema-Notestine, C., Hagler, D. J. Jr., McEvoy, L. K., Fleisher, A. S., Wu, E. H., Karow, D. S., et al. (2009). Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Hum. Brain Mapp.* 30, 3238–3253. doi: 10.1002/hbm.20744
- Fletcher, E., Singh, B., Harvey, D., Carmichael, O., and DeCarli, C. (eds.). (2012). "Adaptive image segmentation for robust measurement of longitudinal brain tissue change," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE)*. doi: 10.1109/EMBC.2012.6347195
- Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., et al. (2006). Mild cognitive impairment. *Lancet* 367, 1262–1270. doi: 10.1016/S0140-6736(06)68542-5
- Graham, N. S., and Sharp, D. J. (2019). Understanding neurodegeneration after traumatic brain injury: from mechanisms to clinical trials in dementia. *J. Neurol. Neurosurg. Psychiatry* 90, 1221–1233. doi: 10.1136/jnnp-2017-317557
- Grossman, M., Payer, F., Onishi, K., White-Devine, T., Morrison, D., D'Esposito, M., et al. (1997). Constraints on the cerebral basis for semantic processing from neuroimaging studies of Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 63, 152–158. doi: 10.1136/jnnp.63.2.152
- Haynes, W. (2013). "Wilcoxon rank sum test," in *Encyclopedia of Systems Biology*, eds W. Dubitzky, O. Wolkenhauer, K.-H. Cho, and H. Yokota (New York, NY: Springer), 2354–2355. doi: 10.1007/978-1-4419-9863-7_1185
- Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011
- Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., and Martin, R. L. (1982). A new clinical scale for the staging of dementia. *Brit. J. Psychiatry* 140, 566–572. doi: 10.1192/bjp.140.6.566
- Jefferson, A. L., Himali, J. J., Beiser, A. S., Au, R., Massaro, J. M., Seshadri, S., et al. (2010). Cardiac index is associated with brain aging: the Framingham Heart Study. *Circulation* 122, 690–697. doi: 10.1161/CIRCULATIONAHA.109.905091
- Kearney, E., and Guenther, F. H. (2019). Articulating: the neural mechanisms of speech production. *Lang. Cogn. Neurosci.* 34, 1214–1229. doi: 10.1080/23273798.2019.1589541
- Liu, Y., Leong, A. T., Zhao, Y., Xiao, L., Mak, H. K., Tsang, A. C. O., et al. (2021). A low-cost and shielding-free ultra-low-field brain MRI scanner. *Nat. Commun.* 12, 7238. doi: 10.1038/s41467-021-27317-1
- Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., et al. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 396, 413–446. doi: 10.1016/S0140-6736(20)30367-6
- Luz, S., Haider, F., de la Fuente Garcia, S., Fromm, D., and MacWhinney, B. (2021). Alzheimer's dementia recognition through spontaneous speech. *Front. Comput. Sci.* 3, 780169. doi: 10.3389/fcomp.2021.780169
- Mahmood, S. S., Levy, D., Vasan, R. S., and Wang, T. J. (2014). The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet* 383, 999–1008. doi: 10.1016/S0140-6736(13)61752-3
- McHugh, M. L. (2013). The chi-square test of independence. *Biochem. Med.* 23, 143–149. doi: 10.11613/BM.2013.018
- Morrison, M. S., Aparicio, H. J., Blennow, K., Zetterberg, H., Ashton, N. J., Karikari, T. K., et al. (2022). Antemortem plasma phosphorylated tau (181) predicts Alzheimer's disease neuropathology and regional tau at autopsy. *Brain* 145, 3546–3557. doi: 10.1093/brain/awac175
- Pihlajamäki, M., Tanila, H., Hänninen, T., Könönen, M., Laakso, M., Partanen, K., et al. (2000). Verbal fluency activates the left medial temporal lobe: a functional magnetic resonance imaging study. *Ann. Neurol.* 47, 470–476. doi: 10.1002/1531-8249(200004)47:4<470::AID-ANA10>3.0.CO;2-M
- Pinheiro, J. C., and Bates, D. M. (2000). "Linear mixed-effects models: basic concepts and examples," in *Mixed-Effects Models in Sand S-PLUS. Statistics and Computing* (New York, NY: Springer). doi: 10.1007/978-1-4419-0318-1_1
- Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N. Y. Acad. Sci.* 1191, 62–88. doi: 10.1111/j.1749-6632.2010.05444.x
- Prudenzi, A., Fioravanti, A., Petriconi, L., and Caracciolo, V. (eds.). (2019). "Power Quality problems in hospital: a case study," in *2019 IEEE Milan PowerTech (IEEE)*. doi: 10.1109/PTC.2019.8810682
- Rajapakse, J. C., Giedd, J. N., DeCarli, C., Snell, J. W., McLaughlin, A., Vauss, Y. C., et al. (1996). A technique for single-channel MR brain tissue segmentation: application to a pediatric sample. *Magnet. Reson. Imaging* 14, 1053–1065. doi: 10.1016/S0730-725X(96)00113-0
- Ramos-Escobar, N., Mercier, M., Trébuchon-Fonséca, A., Rodriguez-Fornells, A., François, C., and Schön, D. (2022). Hippocampal and auditory contributions to speech segmentation. *Cortex* 150, 1–11. doi: 10.1016/j.cortex.2022.01.017
- Ries, M. L., Carlsson, C. M., Rowley, H. A., Sager, M. A., Gleason, C. E., Asthana, S., et al. (2008). Magnetic resonance imaging characterization of brain structure and function in mild cognitive impairment: a review. *J. Am. Geriatr. Soc.* 56, 920–934. doi: 10.1111/j.1532-5415.2008.01684.x
- Rosenberg, A., Mangialasche, F., Ngandu, T., Solomon, A., and Kivipelto, M. (2020). Multidomain interventions to prevent cognitive impairment, Alzheimer's disease, and dementia: from FINGER to World-Wide FINGERS. *J. Prev. Alzheimers Dis.* 7, 29–36. doi: 10.14283/jpad.2019.41
- Sang, Z., Wang, K., Dong, J., and Tang, L. (2022). Alzheimer's disease: updated multi-targets therapeutics are in clinical and in progress. *Eur. J. Med. Chem.* 2022, 114464. doi: 10.1016/j.ejmech.2022.114464
- Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J. K., Baird, A., et al. (eds.). (2016). "The interspeech 2016 computational paralinguistics challenge: deception, sincerity and native language," in *17th Annual Conference of the International Speech Communication Association (Interspeech 2016)*. doi: 10.21437/Interspeech.2016-129
- Seraji-Borzgadz, N., Paulson, H., and Heidebrink, J. (2019). Neurologic examination in the elderly. *Handb. Clin. Neurol.* 167, 73–88. doi: 10.1016/B978-0-12-804766-8.00005-4
- Seshadri, S., Wolf, P. A., Beiser, A., Au, R., McNulty, K., White, R., et al. (1997). Lifetime risk of dementia and Alzheimer's disease: the impact of mortality on risk estimates in the Framingham Study. *Neurology* 49, 1498–1504. doi: 10.1212/WNL.49.6.1498

- Simonyan, K., and Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist* 17, 197–208. doi: 10.1177/1073858410386727
- Smith, E. E., Egorova, S., Blacker, D., Killiany, R. J., Muzikansky, A., Dickerson, B. C., et al. (2008). Magnetic resonance imaging white matter hyperintensities and brain volume in the prediction of mild cognitive impairment and dementia. *Arch. Neurol.* 65, 94–100. doi: 10.1001/archneuro.2007.23
- Spartano, N. L., Davis-Plourde, K. L., Himali, J. J., Andersson, C., Pase, M. P., Maillard, P., et al. (2019). Association of accelerometer-measured light-intensity physical activity with brain volume: the framingham heart study. *JAMA Netw. Open* 2, e192745. doi: 10.1001/jamanetworkopen.2019.2745
- Tahon, M., and Devillers, L. (2015). Towards a small set of robust acoustic features for emotion recognition: challenges. *IEEE/ACM Trans. Audio Speech Lang. Process.* 24, 16–28. doi: 10.1109/TASLP.2015.2487051
- Taler, V., and Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *J. Clin. Exp. Neuropsychol.* 30, 501–556. doi: 10.1080/13803390701550128
- Themistocleous, C., Eckerström, M., and Kokkinakis, D. (2018). Identification of mild cognitive impairment from speech in Swedish using deep sequential neural networks. *Front. Neurol.* 9, 975. doi: 10.3389/fneur.2018.00975
- Thomas, R. J., Kim, H., Maillard, P., DeCarli, C. S., Heckman, E. J., Karjadi, C., et al. (2021). Digital sleep measures and white matter health in the Framingham Heart Study. *Explorat. Med.* 2, 253. doi: 10.37349/emed.2021.00045
- Tsao, C. W., and Vasan, R. S. (2015). Cohort Profile: the Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int. J. Epidemiol.* 44, 1800–1813. doi: 10.1093/ije/dyv337
- Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (eds.). (2007). “Non-parametric diffeomorphic image registration with the demons algorithm,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer). doi: 10.1007/978-3-540-75759-7_39
- Weninger, F., Eyben, F., Schuller, B. W., Mortillaro, M., and Scherer, K. R. (2013). On the acoustics of emotion in audio: what speech, music, and sound have in common. *Front. Psychol.* 4, 292. doi: 10.3389/fpsyg.2013.00292
- Wolf, P. A. (2012). Contributions of the Framingham Heart Study to stroke and dementia epidemiologic research at 60 years. *Arch. Neurol.* 69, 567–571. doi: 10.1001/archneuro.2011.977
- Yuan, J., Libon, D. J., Karjadi, C., Ang, A. F., Devine, S., Auerbach, S. H., et al. (2021). Association between the digital clock drawing test and neuropsychological test performance: large community-based prospective cohort (Framingham heart study). *J. Med. Internet Res.* 23, e27407. doi: 10.2196/27407
- Zhu, W., Sun, L., Huang, J., Han, L., and Zhang, D. (2021). Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI. *IEEE Trans. Med. Imaging* 40, 2354–2366. doi: 10.1109/TMI.2021.3077079