

# USE VOICE ANALYSIS FOR ALZHEIMER'S SCREENING

D. Stück<sup>1</sup>, A. Signorini<sup>1</sup>, T. Hanai<sup>2</sup>, M. Sandoval<sup>1</sup>, C. Lemke<sup>1</sup>, J. Glass<sup>2</sup>, S. Hardy<sup>3,4</sup>, M. Lavalée<sup>3,4</sup>, B. Wasserman<sup>3,4</sup>, T. Ang<sup>3,5</sup>, C. Nowak<sup>3,4</sup> and R. Au<sup>3,4,5,6</sup>

(1) Evidation Health, (2) Massachusetts Institute of Technology, (3) Framingham Heart Study, (4) Boston University School of Medicine, (5) Boston University School of Public Health, (6) Boston University Alzheimer's Disease Center

evidation

BOSTON UNIVERSITY



## Introduction

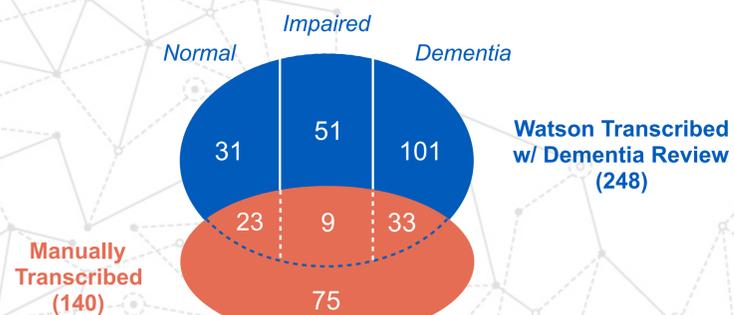
Alzheimer's Disease affects 10% people over 65 and is the 6th leading cause of death in the US. Earlier detection is needed but current screening methodologies are not always clinically practical. Here we show a non-invasive digital voice biomarker that can be used for early detection of dementia.

## Data Used

As a basis for our analysis we used data collected from the Framingham Heart Study (FHS) cohort during neuropsychological testing exams and the associated high-fidelity dementia status reports.

### AT EACH VISIT WERE GATHERED:

1. Demographic Information
2. Health Data (e.g., serum labs values, weight)
3. Cognitive Status
4. Recorded Audio

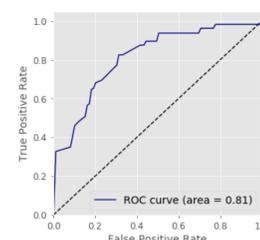


Venn Diagram FHS dataset cohorts used in analyses

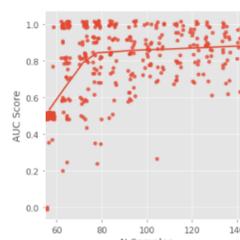
## Analysis

We used audio analysis, speech transcription, and language processing to automatically reduce voice segments from participants into a set of digital biomarkers for the early detection and diagnosis of dementia:

- **Input:** Acoustic, quantitative and linguistic features
- **Target:** Participant's cognitive status
- **Machine Learning:** Random Forest classifier
- **Performance Metric:** Area Under the Curve (AUC)
- **Validation Strategy:** 10-fold cross validation



Dementia vs. Normal  
ROC curve, Audio features

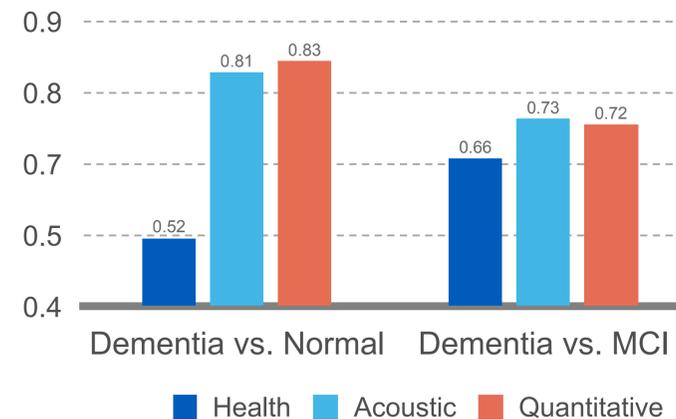


Dementia vs. Normal  
AUC scores, each fold w/ jitter

## Audio Features

- Acoustic Features**  
e.g., pitch, Harmonic-to-Noise ratio, jitter
- Quantitative Audio Features**  
e.g., #fillers, speed, pauses length, answer delay
- Linguistic Features**  
e.g., part of speech tags, tree depth, frequency

## Classification Performance (AUC)



## Results

In predicting dementia, **context agnostic features (A,B)** are 55% more accurate than health data [See Figure Above]. **Linguistic features (C)** have highest predicting power even in mixed age population [See Table Below].

Dementia vs. Normal (*)	
All	0.91 [0.85, 0.97]
Linguistic Features	0.9 [0.82, 0.98]
Audio + Linguistic	0.89 [0.82, 0.97]
Audio Features Only	0.76 [0.66, 0.86]
Health & Demographic	0.82 [0.71, 0.93]

Mean AUC Scores and 95<sup>th</sup> CI

### \* MANUALLY TRANSCRIBED DATASET

Limited to 140 participants with skewed age distribution: cohort with dementia (45) had an average age of 84 versus 67 for normal cohort.

## Conclusion

While these results are preliminary, simple acoustic and language features computed over speech segments show promise for the development of accurate digital biomarkers of cognitive impairment.

- **Strengths:** Audio features performed well even if original input were lossy old recordings.
- **Limitations:** Results obtained on limited dataset, manual transcriptions were necessary.
- **Future Developments:** Increase sample size, obtain better audio recordings, supplement data with other data sources.

## Acknowledgments

This work was supported by the Framingham Heart Study's National Heart, Lung, and Blood Institute contract (N01-HC-25195; HHSN2682015000011), by grants (R01-AG016495, R01-AG008122, R01-AG033040) from the National Institute on Aging, and by grant (R01-NS017950) from the National Institute of Neurological Disorders and Stroke.

## References

- **Evidation** <https://evidation.com/research>
- **MIT** <https://csail.mit.edu/research>
- **FHS** <https://framinghamheartstudy.org>

[asignorini@evidation.com](mailto:asignorini@evidation.com)