Language Biomarkers of Brain Health

Rhoda Au, Ph.D.
2nd International Symposium on Language Resources and Intelligence
12-17-18
Global Burden of Dementia

The diagram illustrates the global burden of dementia, showing the projected numbers of people affected in different regions and years. By 2050, the world is expected to have 131.5 million people with dementia, primarily in Asia. Europe is expected to have 67.2 million, Africa 15.8 million, and the Americas 29.9 million. The number is projected to increase in each region by 2030, with Asia reaching 74.7 million, Europe 38.5 million, Africa 15.8 million, and the Americas 15.8 million.
Can We Prevent the Tsunami Wave of Dementia?
So Far the Answer is No

Why Alzheimer's Drugs Keep Failing

Alzheimer’s Disease: What Stands Between Us and a Cure?

Eli Lilly’s Experimental Alzheimer’s Drug Fails in Large Trial

Merck's Alzheimer's Drug Fails; Are Biogen, Roche, AstraZeneca Next?
Current Situation

Adapted by Drs. Dana Penney and Randall Davis from Fig 3 of Criteria For Preclinical Alzheimer’s Disease, Alzheimer’s Association report (2010), which in turn cites Jack C R, et al., Hypothetical model of dynamic biomarkers of the Alzheimer’s pathological cascade, The Lancet, 9:1, Jan 2010, pp 119-128.
The Impact

Leveraging Framingham
Framingham Cognitive Aging
Preclinical-Incident Dementia

Gen 1
1948

1976-78 Baseline NP
1981 MMSE
1984 NINDS-ADRDA AD Diagnostic Criteria
2016

Gen 2
1971

1991 MMSE
1999 NP/MRI
2016

Mild Cognitive Impairment
Cognitive Impairment Detection Challenge
FHS Two-Method Approach

The Traditional Approach

Boston Process Approach
- Preserve traditional measures
- Expand test beyond single score
- Focus on path to final response
  - Value of incorrect responses
## Logical Memory Error Measures

<table>
<thead>
<tr>
<th>Traditional scores</th>
<th>Qualitative Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Confabulations – related</td>
</tr>
<tr>
<td></td>
<td>Intrusions – related</td>
</tr>
<tr>
<td></td>
<td>Confabulations – unrelated</td>
</tr>
<tr>
<td></td>
<td>Intrusions – unrelated</td>
</tr>
</tbody>
</table>
Framingham Cognitive Aging Today
Preclinical-Incident Disease

Gen 1
1948

1976-78 Baseline NP
1981 MMSE
1984 NINDS-ADRDA AD Diagnostic Criteria

1999 NP/MRI
2016 Repeat NP/MRI

Gen 2
1971

1991 MMSE
1999 NP/MRI
2016 Repeat NP/MRI
Framingham Cognitive Aging Today
Asymptomatic-Preclinical-Incident Disease

Gen 1
1948
- 1976 Baseline NP
- 1981 MMSE
- 1984 NINDS-ADRDA AD Diagnostic Criteria
- 2016

Gen 2
1971
- 1991 MMSE
- 1999 NP/MRI
- 2016

Gen 3
2002
- 2008 NP/MRI w/ DTI
- 2020

Repeat NP/MRI
What is Preclinical?

- 65+ years old
  - Measures differentiate cognitively intact vs. clinically demented
- 45-65 years old
  - Measures differentiate cognitively intact vs. pre-clinical
- <45 years old
  - Measures differentiate cognitively intact vs. cognitively intact
BPA on Steroids
Digital Voice

Not Demented: 2009

Mild Cognitive Impairment: 2015
Digital Voice Markers
Cognitive Impairment

Sample Demographics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>% Men</th>
<th>Ave. Age (s.d.)</th>
<th>Education^A</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Dementia (ND)</td>
<td>35</td>
<td>60.0</td>
<td>83±8</td>
<td>2.0</td>
</tr>
<tr>
<td>Cognitively Impaired-ND (CIND)</td>
<td>58</td>
<td>40.4</td>
<td>83±8</td>
<td>1.5</td>
</tr>
<tr>
<td>Dementia*</td>
<td>107</td>
<td>37.5</td>
<td>83±6</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Audio Features

A. **Acoustic Features**
   - e.g., pitch, Harmonic-to-Noise ratio, jitter

B. **Quantitative Audio Features**
   - e.g., #fillers, speed, pauses length, answer delay

C. **Linguistic Features**
   - e.g., part of speech tags, tree depth, frequency

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

**Classification Performance (AUC)**

MANUALLY TRANSCRIBED DATASET
Limited to 127 participants:

- Dementia (n=45, mean age=83±7)
- ND (n=82; mean age=65±17)

<table>
<thead>
<tr>
<th></th>
<th>Dementia vs. ND</th>
<th>Dementia vs. CIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>0.91 [0.85, 0.97]</td>
<td>0.81 [0.73, 0.72]</td>
</tr>
<tr>
<td>Linguistic Features</td>
<td>0.9 [0.82, 0.98]</td>
<td>0.66 [0.58, 0.78]</td>
</tr>
<tr>
<td>Audio + Linguistic</td>
<td>0.89 [0.82, 0.97]</td>
<td>0.73 [0.65, 0.82]</td>
</tr>
<tr>
<td>Audio Features Only</td>
<td>0.76 [0.66, 0.86]</td>
<td>0.72 [0.63, 0.81]</td>
</tr>
<tr>
<td>Health &amp; Demographic</td>
<td>0.82 [0.71, 0.93]</td>
<td></td>
</tr>
</tbody>
</table>

Mean AUC Scores and 95th CI
FDA Digital Health Initiative

The FDA has implemented the *Digital Health Innovation Action Plan*

Promote the development and use of digital health technologies

Source: https://www.fda.gov/MedicalDevices/DigitalHealth/default.htm
Digital health technologies may act as susceptibility/risk biomarkers.

Digital Health Initiative

• FDA Digital Health Criteria:
  - There is a valid association between the digital device output and the targeted clinical condition.
  - The digital device processes input data to generate accurate, reliable, and precise output data.
  - The output data achieves the intended clinical care purpose in the target population.
Digital Biomarkers

DIGITAL BIOMARKERS

Background
- Behavioral activity markers offer an alternative tool for detecting normal cognitive aging to MCI transitions.

Methods
- Average talk time/day collected through small wearable digital recording devices.
- Data examined by speech detection algorithms.

Results
- MCI subjects used more words during conversations and exhibited longer daily talking time than normal subjects.
- MCI subjects exhibit subtle language processing deficits that are sensitive to transitions to MCI.

Digital Voice Biomarkers

- **Normal older adult**
  - Low amyloid and tau

- **Normal older adult**
  - High amyloid and tau

- **AD patient**
  - High amyloid and tau

- **Tau**
  - Brain scans showing differences in amyloid and tau levels

- **Amyloid**
  - Brain scans showing differences in amyloid levels
What Else?
Cognitive Impairment

Cognitive Health
Sustaining brain health across life span
What We Need to Do
Precision Brain Health

Data Science/AI
## What It Will Take

### Paradigmatic Shift

<table>
<thead>
<tr>
<th><strong>Dimension</strong></th>
<th><strong>Current Mindset</strong></th>
<th><strong>From</strong></th>
<th><strong>To</strong></th>
<th><strong>New Mindset</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Investment Focus</strong></td>
<td>Finding the cure for Alzheimer’s disease and related disorders</td>
<td></td>
<td></td>
<td>Finding how to prevent brain related disorders</td>
</tr>
<tr>
<td><strong>Research Focus</strong></td>
<td>Siloed, a-priori hypothesis-driven research to find treatments for Alzheimer’s diseases</td>
<td></td>
<td></td>
<td>Understanding lifelong health and experiences of sub-populations to optimize their brain health</td>
</tr>
<tr>
<td><strong>Research Scope</strong></td>
<td>Isolating single or small set of factors associated with disease</td>
<td></td>
<td></td>
<td>Deploying scalable, multi-factorial solutions to diverse populations with and without disease risk</td>
</tr>
<tr>
<td><strong>Care Focus</strong></td>
<td>Diagnosing and treating dementia</td>
<td></td>
<td></td>
<td>Monitoring and sustaining brain health</td>
</tr>
<tr>
<td><strong>Healthcare Delivery</strong></td>
<td>Stove-piped 烟道式 Healthcare Institutions</td>
<td></td>
<td></td>
<td>Healthcare Ecosystem</td>
</tr>
<tr>
<td><strong>Data Capture</strong></td>
<td>In clinic, sporadic time point, highly invasive, often self reported data collection</td>
<td></td>
<td></td>
<td>Continuous, passive capture of brain related data with rich demographic metadata</td>
</tr>
<tr>
<td><strong>Health Information Management</strong></td>
<td>Raw, unusable cognitive and related data stored in disparate sources.</td>
<td></td>
<td></td>
<td>Networked, secure, seamless, integrated, sharable, cloud based platform made widely available</td>
</tr>
</tbody>
</table>
What We Want to Do Next
Build Plug and Play System
Where We Want to Do It
Test Platform Robustness Across Diverse Study Sites

UsAgainstAlzheimer’s

CONCUSSION Legacy Foundation

Alzheimer's Disease Research Centers

NIH National Institute on Aging
What We Have Learned

- **Secure**
- **Flexible & Agnostic**
- **Device Grade**
- **Open Source API**
- **Level of Use**

![Diagram showing interrelated concepts](image)

- **Balance Criteria**
- **Defined stakeholder value proposition**
- **Scalable**
- **Strong data infrastructure**
- **Easy data use/sharing**
Ambient Technology
Sustainable Remote Monitoring
Where Next?
Inferred Cognition

Sleep

Physical Activity

Medication

Social Activity

Eating Behavior

Cognition

Sleep

Physical Activity

Medication

Social Activity

Eating Behavior

Inferred Cognition

Cohens Veterans Bioscience

Sage

SAVONIX

mymeds

AnthroTronix

Kryptowire

Shimmer
What Ultimate Success Will Look Like
Open Science, Scalable, Integrated Research & Care Platform
Impact of Success
Think Different, Be Different
When a fast shot forces 15-year-old Chris Chen to the floor, it seems like the point is over.
Thanks.