MRI Hippocampal Volume for Enrichment

Derek Hill¹,² & Jerry Novak³, Pat Cole⁴, Diane Stephenson⁵

1. IXICO plc
2. UCL, London, UK
3. Johnson and Johnson
4. Takeda
5. CAMD, Critical Path Institute
CAMD AD HV Biomarker
Team Members

Alzheimer’s Association—Maria Carrillo, Jim Hendrix
BioClinica—Chahin Pachai
Boehringer Ingelheim—Mark Gordon
Eli Lilly—Adam Schwarz, Peng Yu, Brian Willis
FDA—Marc Walton, Jim Kaiser
Fatebenefratelli—Giovanni Frisoni, Alberto Redolfi, Marina Boccardi
ICON—David Raunig
Ixico—Derek Hill, Katherine Gray, Robin Wolz
Janssen—Jerry Novak, Mahesh Samtani
Novartis—Paul Maguire, Richard Meibach
Orion—Kristiina Haasio
Pentara—Suzanne Hendrix
Pfizer—Kaori Ito, Rachel Schindler
Synarc—Joyce Suhy
Takeda—Pat Cole
UCDavis—Laurel Beckett
University of Trento, Italy—Jorge Jovicich
Mayo Clinic—Cliff Jack
UCSD—Jim Brewer
Dementia Research Centre, Institute of Neurology, London—Nick Fox
Disclosure

Derek Hill is a shareholder and director of IXICO plc

IXICO has technologies that are referred to in this presentation, and some results presented are obtained from IXICO technologies, but these technologies are not promoted in this presentation.
Overview

Why is HV relevant?
Where was the field in 2008?
Where are we now
Value of biomarker regulatory decisions to drug development
Advances in the field since 2011 EMA decision
Where do we need to be going in the future
How should the FDA and CAMD be working together to help?
Why is Hippocampal Volume Relevant?

- Atrophy begins, and is ultimately most severe, in the medial temporal lobe
- Pattern of atrophy in AD mirrors that of neurofibrillary pathology
- Hippocampal atrophy correlates closely with changes in cognition.
- In a meta-analysis shows 73% sensitivity and 81% specificity for predicting MCI progression to dementia
- 20 years of supportive literature
- MRI widely available
- Complements information from amyloid markers

*Hill et al. Alz & Dem. 2014*
Where was the field in 2008?

Protocols focused on mild-moderate AD
Clinical inclusion criteria predominated
Amyloid PET becoming used in protocols for target engagement
Where we are now?

Many prodromal (or earlier) protocols
Biomarker enrichment common
Enrichment approach is protocol dependent
Failures in mild-moderate AD clinical trials are encouraging studies in pre-dementia.

The challenge of selecting the patient population are consequently greater:

- Biomarkers for enrichment arguably essential.

Qualified biomarkers could replace protocol-specific enrichment:

- De-risking expensive studies
- Encouraging investment in Drug Development

EMA qualification of HCV in 2011 has not been followed by FDA.
EMA qualified low hippocampal volume in 2011 with a “user beware” caveat. FDA is requesting much more data.
Standardization: for structural MRI: a recipe exists

ADNI approach has been (approximately) applied in many global trials


The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI

Methods

Clifford R. Jack Jr., MD1,*, Matt A. Bernstein, PhD1, Nick C. Fox, MD2, Paul Thompson, PhD3, Gene Alexander, PhD4, Danielle Harvey, PhD5, Bret Borowski, RTR1, Paula J. Britson, BS1, Jennifer L. Whitwell, PhD1, Chadwick Ward, BA1, Anders M. Dale, PhD6, Joel P. Felmlee, PhD1, Jeffrey L. Gunter, PhD1, Derek L.G. Hill, PhD7, Ron Killiany, PhD8, Norbert Schuff, PhD9, Sabrina Fox-Bosetti, PhD9, Chen Lin, PhD1, Colin Studholme, PhD9, Charles S. DeCarli, MD10, Gunnar Krueger, PhD11, Heidi A. Ward, PhD1,12, Gregory J. Metzger, PhD13, Katherine T. Scott, PhD11, Richard Mallozzi, PhD12, Daniel Blzek, PhD12, Joshua Levy, PhD14, Josef P. Debbins, PhD4,12, Adam S. Fleisher, MD6, Marilyn Albert, PhD15,16, Robert Green, MD17, George Bartzokis, MD3, Gary Glover, PhD18, John Mugler, PhD19, Michael W. Weiner, MD9, and for the ADNI Study

Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease

Clifford R. Jack, Jr.,a,b, Frederik Barkhofb, Matt A. Bernstein,a, Marc Cantillonb, Patricia E. Cole,d, Charles DeCarli,2, Bruno Duboisf, Simon Duchesneg,h, Nick C. Foxi, Giovanni B. Frisoni,1i, Harald Hampelk, Derek L. G. Hillkm, Keith Johnson,1l, Jean-François Mangin,1m, Philip Scheltens,1n, Adam J. Schwarz,1, Reisa Sperling,1, Joyce Suhym, Paul M. Thompson1, Michael Weiner,1w, Norman L. Foster1

Alzheimer's & Dementia 7 (2011) 474–485

Perspectives
A more detailed understanding of variability

MRI Biomarker qualification is focused on single structure: hippocampus
  • A prognostic biomarker.

CAMD’s efforts to prepare data for the FDA is advancing understanding of variability
  • Instrument variability eg: test-re:test within and between field-strengths
  • Translating biomarker between clinical cohorts
Advances in the field since 2011 EMA decision

- Significant further standardization completed and research published, much generated under CAMD umbrella
- Harmonization of hippocampus
- Significant additional data on instrument variability
- Demonstration of practical benefit of enrichment on protocol design
- Data on transferability of biomarker to other cohorts
Harmonization of hippocampal definition

- Historically, many different SOPs existed for quantifying hippocampal volume, with high between-method variability
- Algorithm-dependent cut points are defined in a training dataset and then be applied to discriminate subjects in a new cohort (time to dementia or change in score)
- Frisoni EADC-AA hippocampus harmonization project has addressed this

Alzheimer's & Dementia, 2014 Sept
Instrument variability shown to be small compared to biological variability

- Test: re-test analysis on ADNI-1 data
- High robustness between back-to-back acquisitions on 1.5T and 3T as well as across both field strengths
- Very consistent results for LEAP and BMAS-HCV

<table>
<thead>
<tr>
<th></th>
<th>LEAP^</th>
<th>BMAS-HCV*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.5T vs 1.5T</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.994</td>
<td>0.995</td>
</tr>
<tr>
<td>$\Delta_{\text{signed}}$ [%]</td>
<td>-0.19±1.93</td>
<td>0.04±1.94</td>
</tr>
<tr>
<td>$\Delta_{\text{unsigned}}$ [%]</td>
<td>1.51±1.22</td>
<td>1.47±1.26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LEAP^</th>
<th>BMAS-HCV*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3T vs 3T</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.992</td>
<td>0.993</td>
</tr>
<tr>
<td>$\Delta_{\text{signed}}$ [%]</td>
<td>-0.21±2.50</td>
<td>0.29±2.17</td>
</tr>
<tr>
<td>$\Delta_{\text{unsigned}}$ [%]</td>
<td>1.52±1.37</td>
<td>1.65±1.43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LEAP^</th>
<th>BMAS-HCV*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.5T vs 3T</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.979</td>
<td>0.978</td>
</tr>
<tr>
<td>$\Delta_{\text{signed}}$ [%]</td>
<td>-1.17±3.07</td>
<td>-1.01±3.72</td>
</tr>
<tr>
<td>$\Delta_{\text{unsigned}}$ [%]</td>
<td>2.68±1.89</td>
<td>3.03±2.36</td>
</tr>
</tbody>
</table>

^LEAP results taken from Wolz et al, J. Alz. & Dement., 2014
*BMAS-HCV results taken from Roche et al, AAIC 2013, bioclinica.com/aaic2013-roche-poster
Practical benefits of HCV enrichment modelled

- Enriching AD clinical trials excludes many slow-progression subjects
- Using hippocampal volume as an inclusion criteria has a positive effect on effect size, the number of subjects needed to be screened as well as trial cost and time.
- Several “mature” algorithms perform equivalently
- Presented are results for enrichment with a cut point at the 25th percentile

<table>
<thead>
<tr>
<th></th>
<th>Unenriched</th>
<th>FreeSurfer</th>
<th>HMAPS</th>
<th>LEAP</th>
<th>NeuroQuant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two year change of MMSE&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-1.68±3.49</td>
<td>-2.39±3.67</td>
<td>-2.42±3.69</td>
<td>-2.67±3.77</td>
<td>-2.61±3.76</td>
</tr>
<tr>
<td>Effect size&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>-0.48 (-0.57 to -0.39)</td>
<td>-0.65 (-0.76 to -0.53)</td>
<td>-0.65 (-0.77 to -0.54)</td>
<td>-0.71 (-0.83 to -0.58)</td>
<td>-0.70 (-0.82 to -0.57)</td>
</tr>
<tr>
<td>N/N&lt;sub&gt;unenriched&lt;/sub&gt;&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.55 (0.24-0.87)</td>
<td>0.54 (0.24-0.85)</td>
<td>0.46 (0.20-0.72)</td>
<td>0.48 (0.22-0.74)</td>
</tr>
<tr>
<td>NNS/NNS&lt;sub&gt;unenriched&lt;/sub&gt;&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1</td>
<td>0.81 (0.34-1.28)</td>
<td>0.81 (0.35-1.27)</td>
<td>0.75 (0.32-1.18)</td>
<td>0.78 (0.34-1.23)</td>
</tr>
<tr>
<td>Screen failure fraction&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0</td>
<td>0.32 (0.26-0.37)</td>
<td>0.33 (0.28-0.39)</td>
<td>0.39 (0.33-0.44)</td>
<td>0.39 (0.33-0.45)</td>
</tr>
<tr>
<td>Sample size/arm to rand.&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1090 (631-1549)</td>
<td>604 (376-832)</td>
<td>590 (369-811)</td>
<td>502 (318-686)</td>
<td>522 (335-709)</td>
</tr>
<tr>
<td>Trial duration (y)</td>
<td>5.9 (4.3-7.5)</td>
<td>5.2 (3.9-6.4)</td>
<td>5.1 (3.9-6.4)</td>
<td>4.9 (3.8-6)</td>
<td>5.1 (3.9-6.2)</td>
</tr>
<tr>
<td>Trial cost (M$)</td>
<td>100.9 (58.4-143.4)</td>
<td>61.1 (38-84.2)</td>
<td>60.0 (37.5-82.6)</td>
<td>52.4 (33-71.7)</td>
<td>54.5 (35-74.1)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Specified as mean ± standard deviation.
<sup>b</sup> Calculated as mean ÷ standard deviation.
<sup>c</sup> Numbers in parentheses are the 95% confidence intervals
<sup>d</sup> Sample size based on 25% difference, 80% power, alpha = 0.05, trial duration = 2 years.
N: number of subjects in trial
NNS: number of subjects needed to screen
Evidence of translation of enrichment cutpoints

Three different datasets were independently analysed for an ideal cut point using example algorithm (LEAP) hippocampal volume to predict aMCI that progress to AD-type dementia:

• ADNI-1, Descripa, VUMC

Published results to date are encouraging:

• Cut points in ADNI and the pooled cohort of Descripa and VUMC were reported as 5.39cm$^3$ and 5.34cm$^3$ for left and right hippocampus combined respectively [1,2]
• Cut points between Descripa and VUMC when analysed independently showed a variability of less than 2.5% [3]

Where do we need to be going in the future?

Progress on qualification of HCV could help define a pathway for other imaging biomarkers eg: tau PET

There is an increasing consensus that enrichment may involve more than one tool

• multiple biomarkers
• Biomarkers + genotype

Biomarkers as surrogate measures of efficacy may ultimately be required for AD treatments in the pre-clinical phase
How should the FDA and CAMD be working together to help?

There is heightened public awareness of the need to increase success in developing treatments for AD and other dementias.

The need to work together within CAMD to qualify biomarkers is greater than ever.
Sources of Variability in MRI biomarkers

Poor acquisition standardization (eg: different manufacturer, sequence, positioning)
Instrument instability over time (eg: scaling drift, change in calibration)
Poor data quality (eg: motion artefacts)
Algorithm / reader variability
  • Lack of standardization of algorithms (eg: centre dependent)
  • Algorithm interaction with instrument (eg: results are field-strength dependent)
Short term biological variability (over hours or days, eg: hydration)
treatment effect (expected or otherwise)
Phenotypical variation
  • Trial A and B might have had very different clinical inclusions criteria or different biomarkers used for eligibility

Likely all contribute: but which dominate?
Studying instrument variability with ADNI-1 Data

Data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) is used

Inclusion criteria:
- Subjects were scanned at 1.5T and 3T
- Subjects were scanned at baseline and month 12
- Two non-processed MP-RAGE scans are available for download

153 study subjects were included
33 different sites
10 different 1.5T scanners, 9 different 3T scanners
612 non-processed baseline scans are downloaded from ADNI

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Female %</th>
<th>Age</th>
<th>MMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>28</td>
<td>60.7%</td>
<td>74.6±8.7</td>
<td>22.8±2.1</td>
</tr>
<tr>
<td>MCI</td>
<td>74</td>
<td>35.1%</td>
<td>74.9±7.8</td>
<td>26.4±3.7</td>
</tr>
<tr>
<td>Control</td>
<td>51</td>
<td>64.7%</td>
<td>75.7±4.9</td>
<td>29.3±0.9</td>
</tr>
</tbody>
</table>
Variability measures on two methods

Automated hippocampal volumetry was independently performed with LEAP [1] and BMAS-HCV [2] on the 612 ADNI scans included

Compared are

• Volumes extracted from the two 1.5T scans
• Volumes extracted from the two 3T scans
• Volumes extracted from first 1.5T scan and first 3T scan

Two variability measures are computed for volumes $v_1$, $v_2$:

• Relative signed difference: $\Delta_{\text{signed}} = 2 \frac{(v_1 - v_2)}{(v_1 + v_2)}$
• Relative unsigned difference $\Delta_{\text{unsigned}} = 2 \left| \frac{(v_1 - v_2)}{(v_1 + v_2)} \right|$