MRI Hippocampal Volume for Enrichment

Derek Hill[1,2] & Jerry Novak[3], Pat Cole[4], Diane Stephenson[5]

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

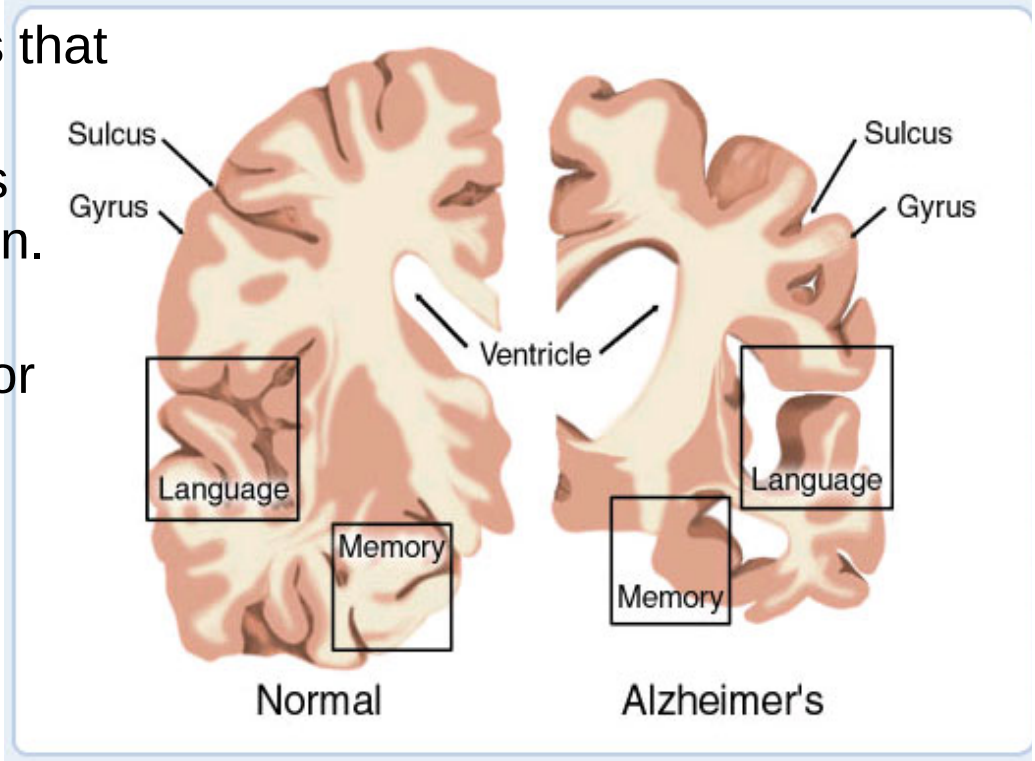
IXICO has technologies that are referred to in this presentation, and some results presented are obtained from IXICO technologies, but these technologies are not promoted in this presentation.

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

Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease

Stephen Salloway, M.D., Reisa Sperling, M.D., Nick C. Fox, M.D., Kaj Blennow, M.D., William Klunk, M.D., Murray Raskind, M.D., Marwan Sabbagh, M.D., Lawrence S. Honig, M.D., Ph.D., Anton P. Porsteinsson, M.D., Steven Ferris, Ph.D., Marcel Reichert, M.D., Nzeera Ketter, M.D., Bijan Nejadnik, M.D., Volkmar Guenzler, M.D., Maja Miloslavsky, Ph.D., Daniel Wang, Ph.D., Yuan Lu, M.S., Julia Lull, M.A., Iulia Cristina Tudor, Ph.D., Enchi Liu, Ph.D., Michael Grundman, M.D., M.P.H., Eric Yuen, M.D., Ronald Black, M.D., and H. Robert Brashear, M.D., for the Bapineuzumab 301 and 302 Clinical Trial Investigators*

Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease

Rachelle S. Doody, M.D., Ph.D., Ronald G. Thomas, Ph.D., Martin Farlow, M.D., Takeshi Iwatsubo, M.D., Ph.D., Bruno Vellas, M.D., Steven Joffe, M.D., M.P.H., Karl Kieburtz, M.D., M.P.H., Rema Raman, Ph.D., Xiaoying Sun, M.S., and Paul S. Aisen, M.D., for the Alzheimer's Disease Cooperative Study Steering Committee; and Eric Siemers, M.D., Hong Liu-Seifert, Ph.D., and Richard Mohs, Ph.D., for the Solanezumab Study Group

Value of regulatory decisions on biomarkers

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 qualification of HCV for enrichment

EMA qualified low hippocampal volume in 2011 with a “user beware” caveat. FDA is requesting much more data

Alzheimer's & Dementia ■ (2013) 1–9

Research Article

Coalition Against Major Diseases/European Medicines Agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease

Derek L. G. Hill[a], Adam J. Schwarz[b], Maria Isaac[c], Luca Pani[c], Spiros Vamvakas[c], Robert Hemmings[c], Maria C. Carrillo[d], Peng Yu[b], Jia Sun[b,e], Laurel Beckett[f], Marina Boccardi[g], James Brewer[h], Martha Brumfield[i], Marc Cantillon[j], Patricia E. Cole[b], Nick Fox[k], Giovanni B. Frisoni[g], Clifford Jack[l], Thomas Kelleher[m], Feng Luo[m], Gerald Novak[n], Paul Maguire[o], Richard Meibach[p], Patricia Patterson[q], Lisa Bain[r], Cristina Sampaio[s], David Raunig[t], Holly Soares[m], Joyce Suhy[u], Huanli Wang[f], Robin Wolz[a], Diane Stephenson[i,*]

Standardization: for structural MRI: a recipe exists

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

J Magn Reson Imaging. 2008 April ; 27(4): 685–691. doi:10.1002/jmri.21049.

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

Clifford R. Jack Jr., MD[1,*], Matt A. Bernstein, PhD[1], Nick C. Fox, MD[2], Paul Thompson, PhD[3], Gene Alexander, PhD[4], Danielle Harvey, PhD[5], Bret Borowski, RTR[1], Paula J. Britson, BS[1], Jennifer L. Whitwell, PhD[1], Chadwick Ward, BA[1], Anders M. Dale, PhD[6], Joel P. Felmlee, PhD[1], Jeffrey L. Gunter, PhD[1], Derek L.G. Hill, PhD[7], Ron Killiany, PhD[8], Norbert Schuff, PhD[9], Sabrina Fox-Bosetti, PhD[9], Chen Lin, PhD[1], Colin Studholme, PhD[9], Charles S. DeCarli, MD[10], Gunnar Krueger, PhD[11], Heidi A. Ward, PhD[1,12], Gregory J. Metzger, PhD[13], Katherine T. Scott, PhD[11], Richard Mallozzi, PhD[12], Daniel Blezek, PhD[12], Joshua Levy, PhD[14], Josef P. Debbins, PhD[4,12], Adam S. Fleisher, MD[6], Marilyn Albert, PhD[15,16], Robert Green, MD[17], George Bartzokis, MD[3], Gary Glover, PhD[18], John Mugler, PhD[19], Michael W. Weiner, MD[9], and for the ADNI Study

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

Alzheimer's & Dementia

Perspectives

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,*], Frederik Barkhof[b], Matt A. Bernstein[a], Marc Cantillon[c], Patricia E. Cole[d], Charles DeCarli[e], Bruno Dubois[f], Simon Duchesne[g,h], Nick C. Fox[i], Giovanni B. Frisoni[j], Harald Hampel[k], Derek L. G. Hill[l,m], Keith Johnson[n,o], Jean-François Mangin[p], Philip Scheltens[q], Adam J. Schwarz[r], Reisa Sperling[s], Joyce Suhy[t], Paul M. Thompson[u], Michael Weiner[v,w], Norman L. Foster[x]

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

Frisoni GB, Jack CR, et al for the EADC - European Alzheimer's Disease Consortium and the ADNI - Alzheimer's Disease Neuroimaging Initiative.
The EADC-ADNI Harmonized Protocol for Hippocampal Segmentation on Magnetic Resonance: Evidence of Validity.
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**

1.5T vs 1.5T	LEAP^	BMAS-HCV*
r	0.994	0.995
Δ_{signed} [%]	-0.19±1.93	0.04±1.94
$\Delta_{unsigned}$ [%}	1.51±1.22	1.47±1.26

3T vs 3T	LEAP^	BMAS-HCV*
r	0.992	0.993
Δ_{signed} [%]	-0.21±2.50	0.29±2.17
$\Delta_{unsigned}$ [%]	1.52±1.37	1.65±1.43

1.5T vs 3T	LEAP^	BMAS-HCV*
r	0.979	0.978
Δ_{signed} [%}	-1.17±3.07	-1.01±3.72
$\Delta_{unsigned}$ [%]	2.68±1.89	3.03±2.36

^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

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

[a] Specified as mean ± standard deviation.
[b] Calculated as mean ÷ standard deviation.
[c] Numbers in parentheses are the 95% confidence intervals

[d] 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

Yu et al, Neurobiology of Aging, 35 (4), 808-818, 2

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]

[1] IA van Rossum et al. Injury markers predict time to dementia in subjects with MCI and amyloid pathology. Neurology. 2012; 79 (17), 1809-1816
[2] S Vos, I. van Rossum, L. Burns, D Knol, P Scheltens, H Soininen et al. Test sequence of CSF and MRI biomarkers for prediction of AD in subjects with MCI. Neurobiology of Aging. 2012; 33: 2272–2281
[3] L Clerx et al. Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. Neurobiology of aging. 2013; 34 (8), 2003-2013

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

	Number	Female %	Age	MMSE
AD	28	60.7%	74.6±8.7	22.8±2.1
MCI	74	35.1%	74.9±7.8	26.4±3.7
Control	51	64.7%	75.7±4.9	29.3±0.9

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: $\Delta_{signed} = 2\,(v_1-v_2)/(v_1+v_2)$
- Relative unsigned difference $\Delta_{unsigned} = 2\,|(v_1-v_2)/(v_1+v_2)|$

[1] Wolz et al, NeuroImage 2010
[2] Belaroussi et al, *Poster session, AAIC 2012*