Neuroimaging enrichment biomarkers for CNS diseases

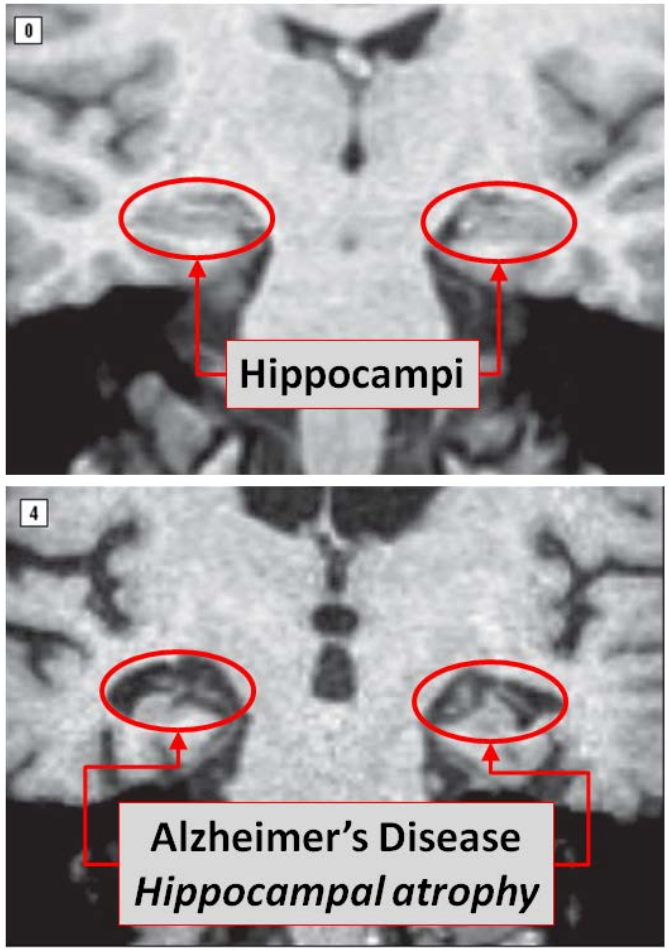
Adam Schwarz
(Eli Lilly and Company)
On behalf of CAMD imaging qualification team
Special thanks to Peng Yu and Derek Hill

Lilly

Outline

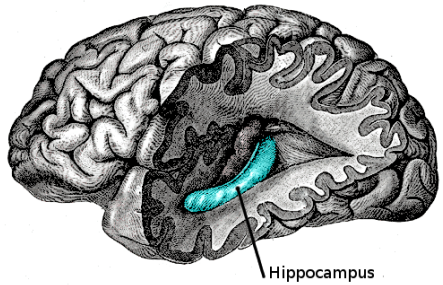
- Hippocampal volume (HV) in AD (case study of an enrichment biomarker)
- Overview of evidentiary considerations for biomarkers
  - General considerations
  - Mapping to HV and context of use for trial enrichment
- NIA-AA recommendations for clinical research in MCI due to AD
- Performance characteristics of HV in MCI
  - Heterogeneity of clinically-defined MCI population (differential clinical progression)
  - Supporting data from the literature
  - Test-retest
  - Sensitivity to different HV algorithms
  - Operational considerations

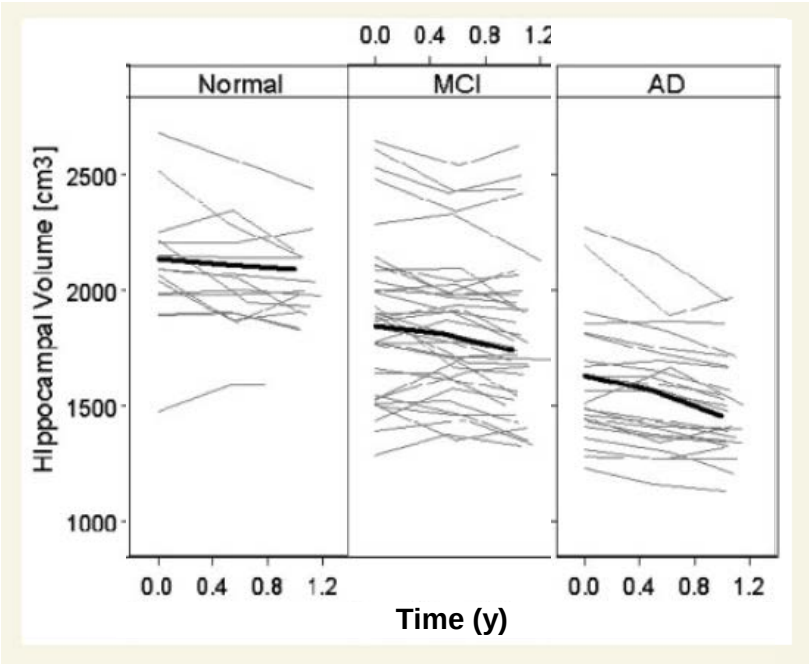
Hippocampal atrophy in Alzheimer's Disease



deCarli *et al.* (2007) *Arch Neurol* **64** 108

L. Seress / Wikipedia Commons



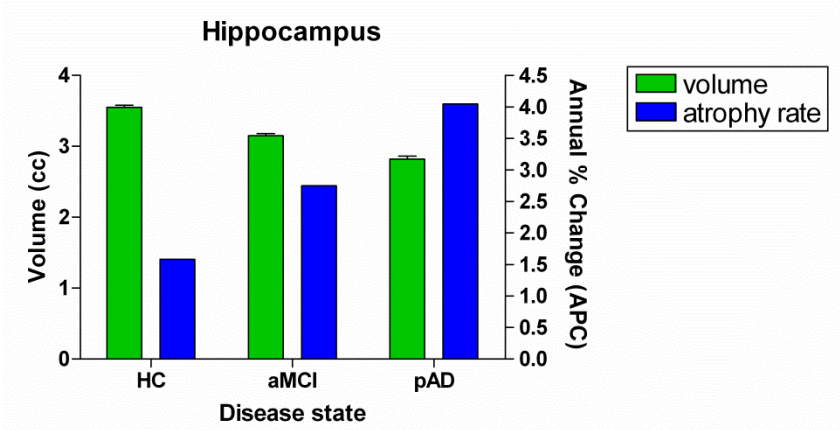


Schuff *et al.* (2009) *Brain* **132** 1067

AD = Alzheimer's Disease. MCI = Mild Cognitive Impairment.

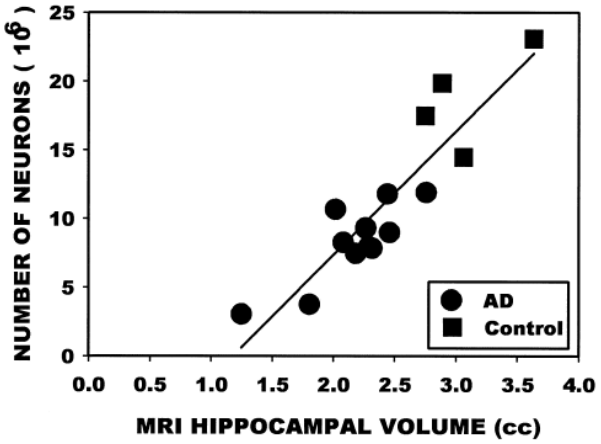
Brain atrophy as measured by structural MRI reflects neuropathology of AD

Disease stage



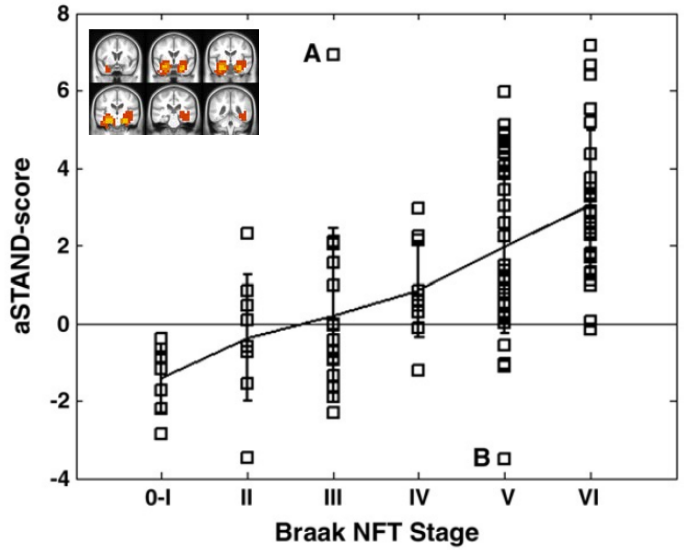
ADNI

Neurodegeneration



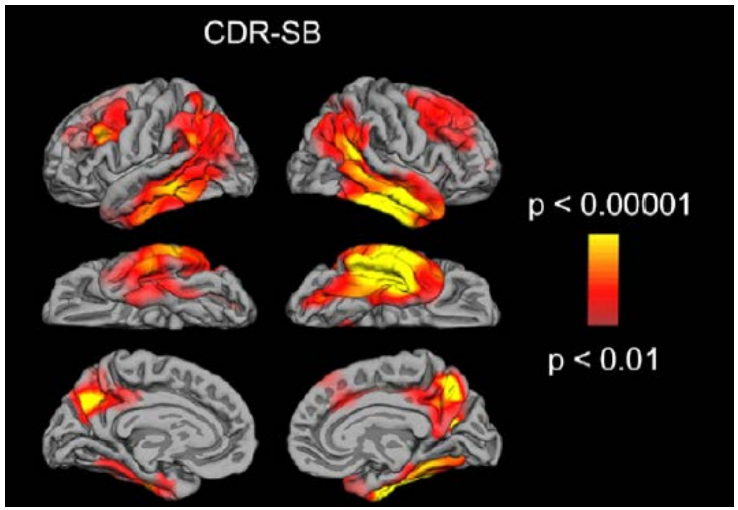
Bobinksi M *et al.* (2000) *Neuroscience* 95(3) 721

Post-mortem Braak stage

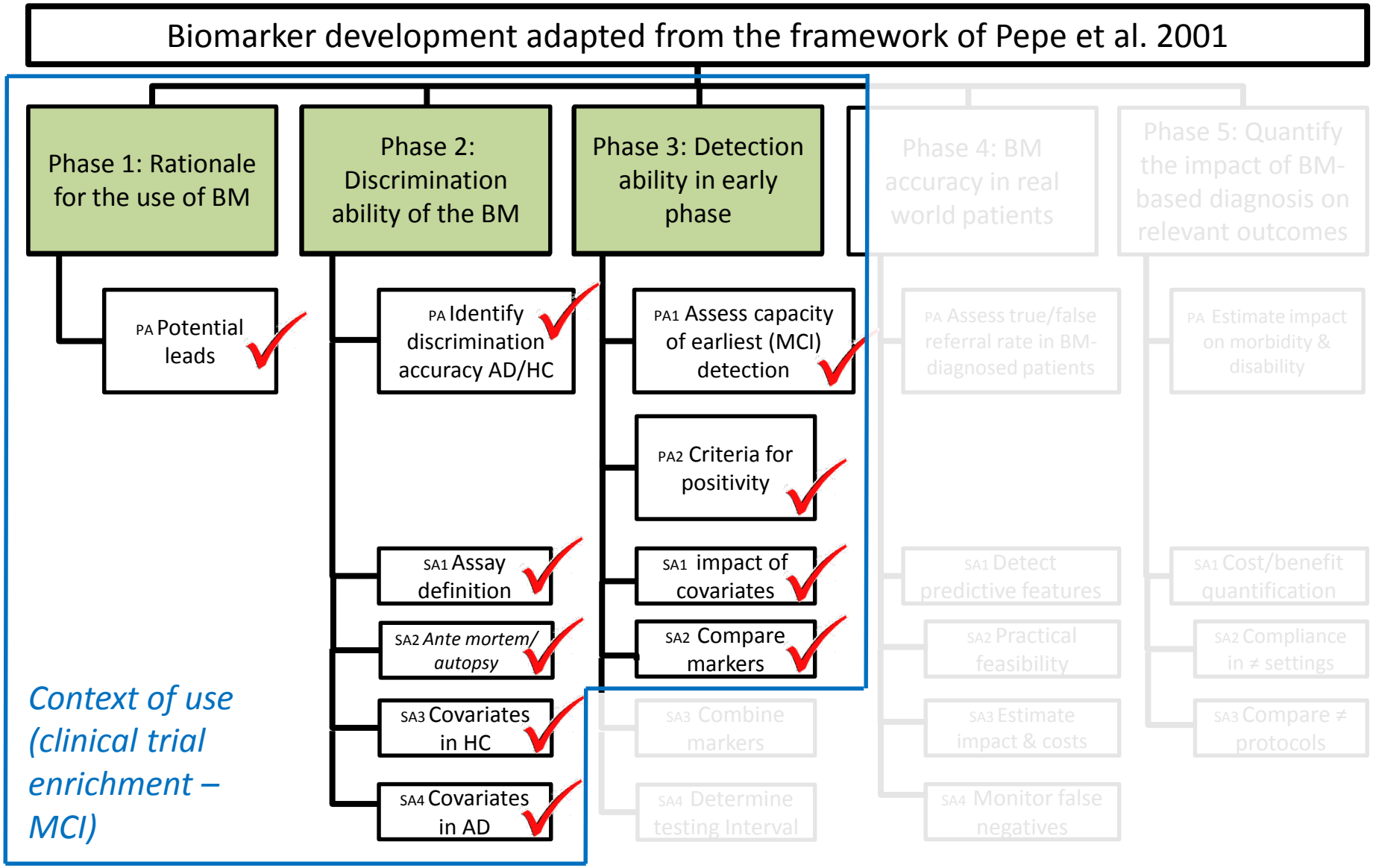


Vemuriet *al.* (2009) *NeuroImage* **42**(2) 559

Cognitive function



McDonald *et al.* (2010) *Neurobiol Aging, in press*



Slide courtesy of Marina Boccardi & Giovanni Frisoni

A Prototypical Process for Creating Evidentiary Standards for Biomarkers and Diagnostics

CA Altar[1], D Amakye[2], D Bounos[3], J Bloom[4], G Clack[5], R Dean[4], V Devanarayan[6], D Fu[7], S Furlong[5], L Hinman[8], C Girman[9], C Lathia[10], L Lesko[11], S Madani[12], J Mayne[13], J Meyer[8], D Raunig[12], P Sager[5], SA Williams[14], P Wong[8] and K Zerba[15]

A framework for developing evidentiary standards for qualification of biomarkers is a key need identified in the Food and Drug Administration's Critical Path Initiative.[1] This article describes a systematic framework that was developed by Pharmaceutical Research and Manufacturers of America (PhRMA) committees and tested at a workshop in collaboration with the Food and Drug Administration and academia. With some necessary refinements, this could be applied to create an appropriately individualized evidentiary standard for any biomarker purpose.

Altar CA et al. (2008) Clin Pharm Ther 83(2) 368

Canonical feature of AD. Causally related to core amnestic phenotype.

N/A (non-chemical marker)

N/A (outside Context of Use)

Evidence from many studies (meta-analysis). *Explicit replication part of proposed HCV analysis plan*.

Standardized methods of acquisition and analysis commonly applied. 510(k)/CE-marked analysis software available. *Hippocampal harmonization.*

No real benchmark. Performs similarly to alternatives.

Table 1 Prototype "evidence map"—categorical description of different types of scientific evidence potentially relevant to biomarker qualification; subcategorical graded weight of evidence from least to most

Evidence type	Grade D	Grade D+/C–	Grade C	Grade C+/B–	Grade B	Grade B+/A–	Grade A
Theory on biological plausibility	Observed association only	Theory, indirect evidence of relevance of the biomarker from animals	As for lower grade but evidence is direct	Theory, indirect evidence of relevance in humans	Theory, direct evidence in humans, non-causal pathway possible	As for lower grade, but biomarker on causal path	Human evidence based mathematical model of biology showing biomarker is on causal pathway
Interaction with pharmacologic target	Biomarker identifies target in *in vitro* binding			Biomarker identifies target in *in vivo* binding in animals	Biomarker identifies target in *in vivo* studies or from human tissue, no truth standard		Biomarker identifies target in *in vivo* studies or from tissues in humans, with accepted truth standard
Pharmacologic mechanistic response	*In vitro* evidence that the drug affects the biomarker	*In vitro* evidence that multiple members of this drug class affects the biomarker	*In vivo* evidence that this drug affects biomarker in animals	As for lower grade but effect shown across drug class	Human evidence that this drug affects the biomarker OR animal evidence of specificity	Human evidence across this mechanistic drug class	Human evidence that multiple members of this drug class affect the biomarker and the effect is specific to this class/mechanism
Linkage to clinical outcome of a disease or toxicity		Biomarker epidemiologically associated with outcome without any intervention	Biomarker associated with change in outcome from intervention in another drug class	As for lower grade but in this drug class	As for lower grade but multiple drug classes albeit inconsistent or a minority of disease effect		As for lower grade but consistent linkage and explains majority of disease effect
Mathematics replication, confirmation		An algorithm is required to interpret the biomarker and was developed from this dataset		Algorithm was developed from a different dataset and applied here prospectively			Algorithm developed from different dataset, replicated prospectively in other sets and applied prospectively here
Accuracy and precision (analytic validation)				Sources of technical variation are unknown but steps are taken to ensure consistent test application	Major sources or variation known and controlled to be less than biological signal; standardization methods applied		All major sources of technical imprecision are known, and controlled test/assay accuracy is defined against standards
Relative performance		Does not meet performance of benchmark		Similar performance to benchmark			Exceeds performance of benchmark or best alternative biomarker

Not all types of evidence required all seven grades to be completed.

Altar CA et al. (2008) Clin Pharm Ther 83(2) 368

Biomarkers of neurodegeneration are embedded in the 2011 NIA-AA research criteria for MCI due to AD

The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging and Alzheimer's Association workgroup

Marilyn S. Albert[a,*], Steven T. DeKosky[b,c], Dennis Dickson[d], Bruno Dubois[e], Howard H. Feldman[f], Nick C. Fox[g], Anthony Gamst[h], David M. Holtzman[i,j], William J. Jagust[k], Ronald C. Petersen[l], Peter J. Snyder[m,n], Maria C. Carrillo[o], Bill Thies[o], Creighton H. Phelps[p]

Table 3
MCI criteria incorporating biomarkers

Diagnostic category	Biomarker probability of AD etiology	Aβ (PET or CSF)	Neuronal injury (tau, FDG, sMRI)
MCI–core clinical criteria	Uninformative	Conflicting/indeterminant/untested	Conflicting/indeterminant/untested
MCI due to AD—intermediate likelihood	Intermediate	Positive	Untested
		Untested	Positive
MCI due to AD—high likelihood	Highest	Positive	Positive
MCI—unlikely due to AD	Lowest	Negative	Negative

Abbreviations: AD, Alzheimer's disease; Aβ, amyloid beta peptide; PET, positron emission tomography; CSF, cerebrospinal fluid; FDG, fluorodeoxyglucose; sMRI, structural magnetic resonance imaging.

Albert M et al. (2011) Alzheimers & Dementia

A systematic survey of the published literature indicated strong evidence for low hippocampal volume as an enrichment biomarker in MCI

ELSEVIER

CrossMark

Alzheimer's & Dementia 10 (2014) 421–429

Alzheimer's & Dementia

Featured Articles

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,v], Diane Stephenson[i,*]

De novo calculations confirmed literature findings and robustness to HCV measurement algorithm

Table 1
Results of Coalition Against Major Diseases' *de novo* analysis. The AUC for four different hippocampal volume quantification algorithms applied to ADNI-1 data indicate the prediction by MRI hippocampal volume of clinical conversion to Alzheimer's dementia within two years.

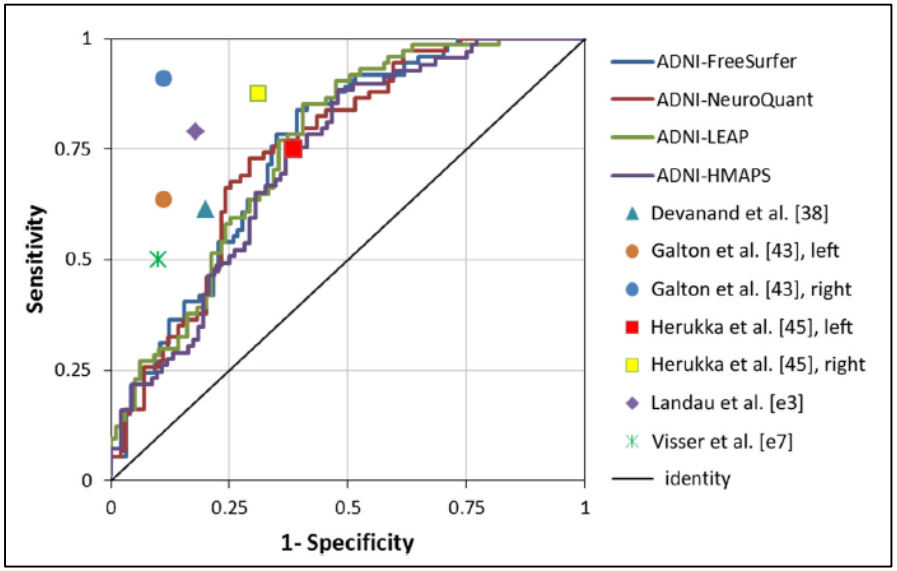
Algorithm	Training, n	Testing, n	AUC based on clinical conversion
LEAP	149	173	0.7565
NeuroQuant	149	173	0.7516
FreeSurfer	148	171	0.7536
HMAPS	128	161	0.7290

Abbreviations: AUC, area under the receiver–operating characteristic curves; LEAP, Learning Embeddings for Atlas Propagation; HMAPS, Hippocampus Multi-Atlas Propagation and Segmentation.

Table 2
AUC values reported in the Coalition Against Major Diseases literature review

Study	n	AUC based on clinical conversion
Bakkour et al. [e9]	49	0.65
Devanand et al. [38]	139	0.77
Fleisher et al. [e10]	129	0.60
Galluzzi et al. [42]	90	0.73

Abbreviation: AUC, area under the receiver–operating characteristic curves.



Hill DLG *et al*. (2014) *Alzheimers & Dementia* **10** 421

Analytic validation: test-retest reliability

ELSEVIER

CrossMark

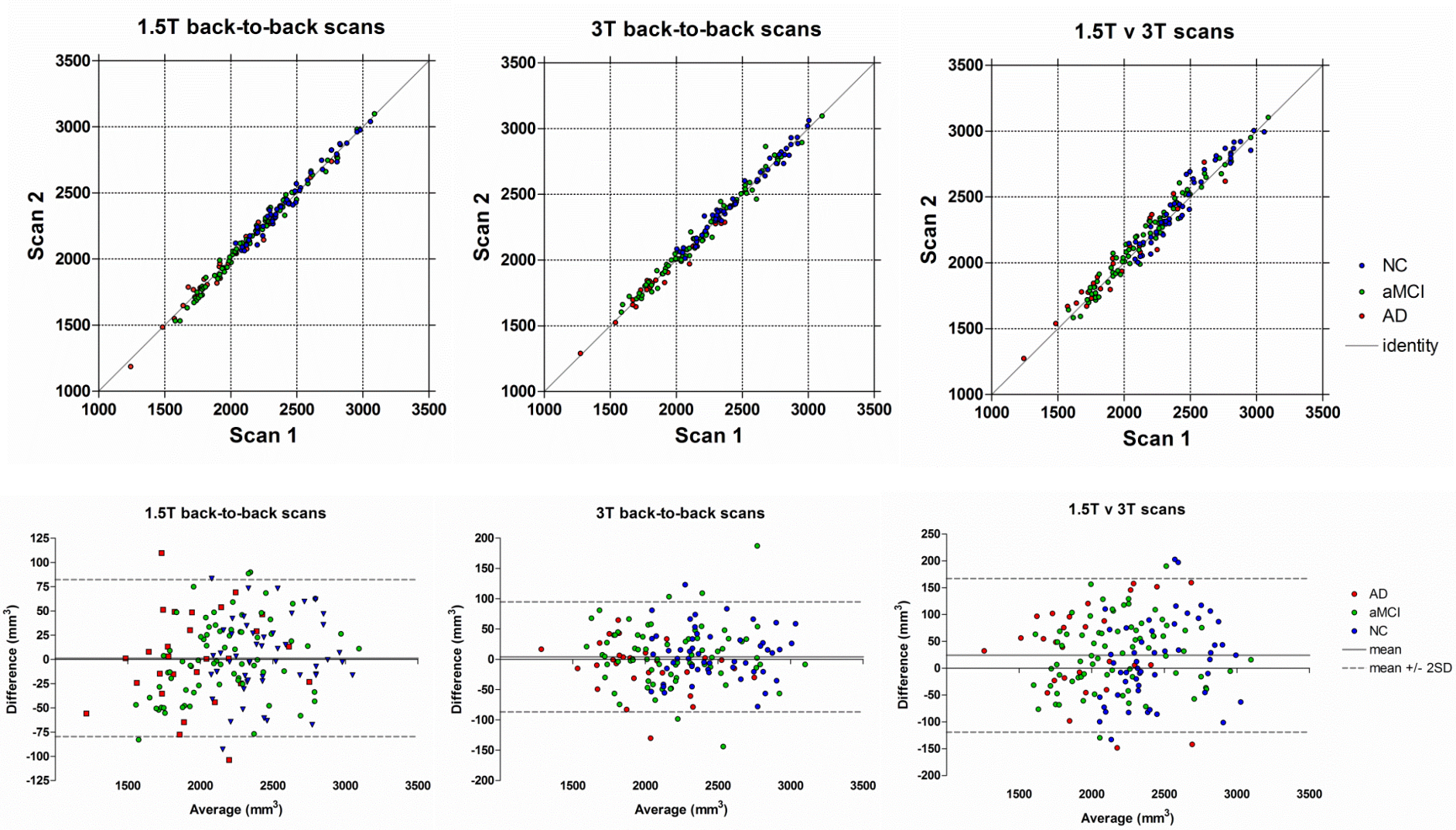
Alzheimer's & Dementia 10 (2014) 430–438

Alzheimer's & Dementia

Robustness of automated hippocampal volumetry across magnetic resonance field strengths and repeat images

Robin Wolz[a,b], Adam J. Schwarz[c], Peng Yu[c], Patricia E. Cole[c], Daniel Rueckert[b], Clifford R. Jack, Jr.,[d], David Raunig[e], Derek Hill[a,*], for The Alzheimer's Disease Neuroimaging Initiative

Hippocampal volume measurements are highly reliable (test-retest)



Wolz R *et al*. (2014) *Alzheimers & Dementia* **10** 430

Operational considerations and practical implications for trials

Neurobiology of Aging 35 (2014) 808–818

Contents lists available at ScienceDirect

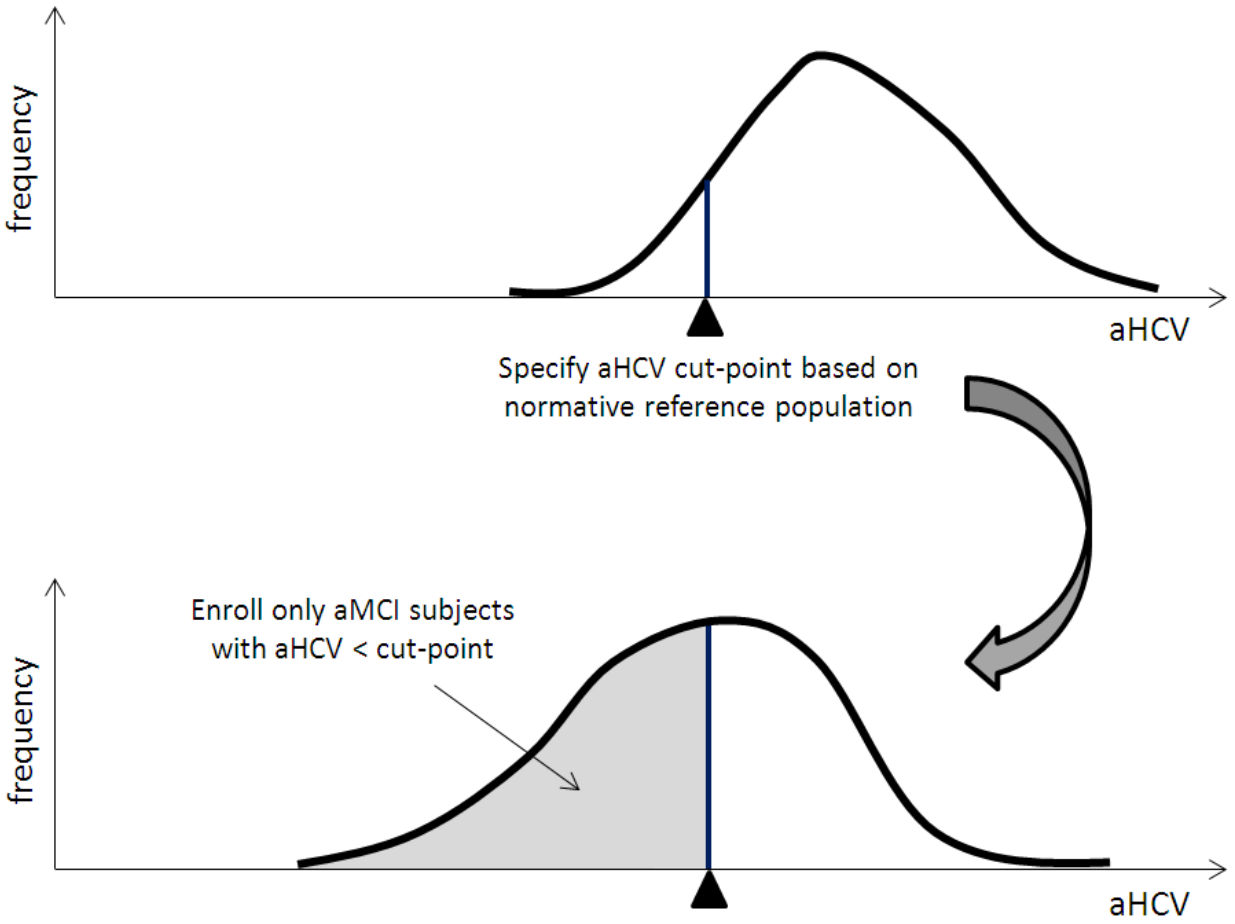
Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Operationalizing hippocampal volume as an enrichment biomarker for amnestic mild cognitive impairment trials: effect of algorithm, test-retest variability, and cut point on trial cost, duration, and sample size

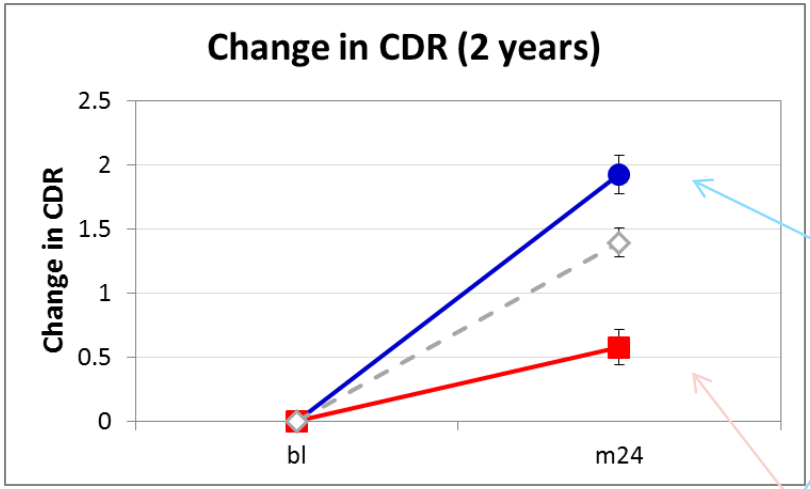
Peng Yu[a], Jia Sun[a,b], Robin Wolz[c,d], Diane Stephenson[e], James Brewer[f], Nick C. Fox[g], Patricia E. Cole[h], Clifford R. Jack Jr[i], Derek L.G. Hill[c,g], Adam J. Schwarz[h,*], for the Coalition Against Major Diseases and the Alzheimer's Disease Neuroimaging Initiative

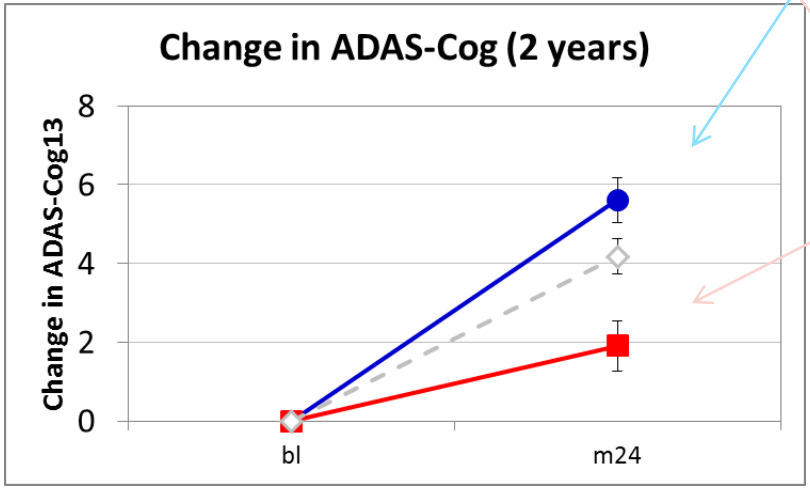
Cut-point defined with respect to normative reference range



MCI subjects with smaller hippocampi progress more rapidly

- Enriched population (HV < 25% of normal)
- All MCI subjects
- Subjects excluded (HV >= 25% of normal)

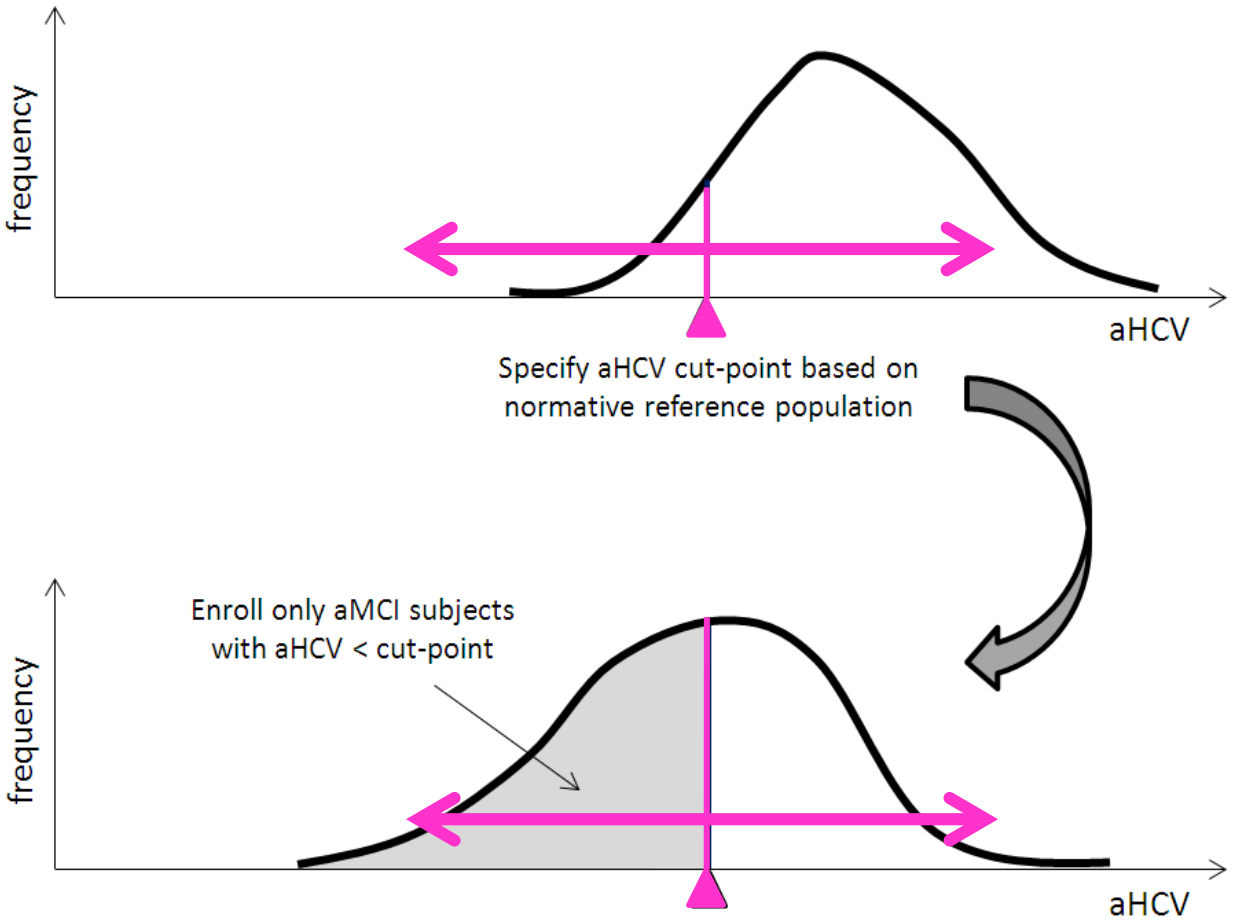




Subjects with smaller HV at baseline progress more rapidly

Slower progressing subjects are excluded

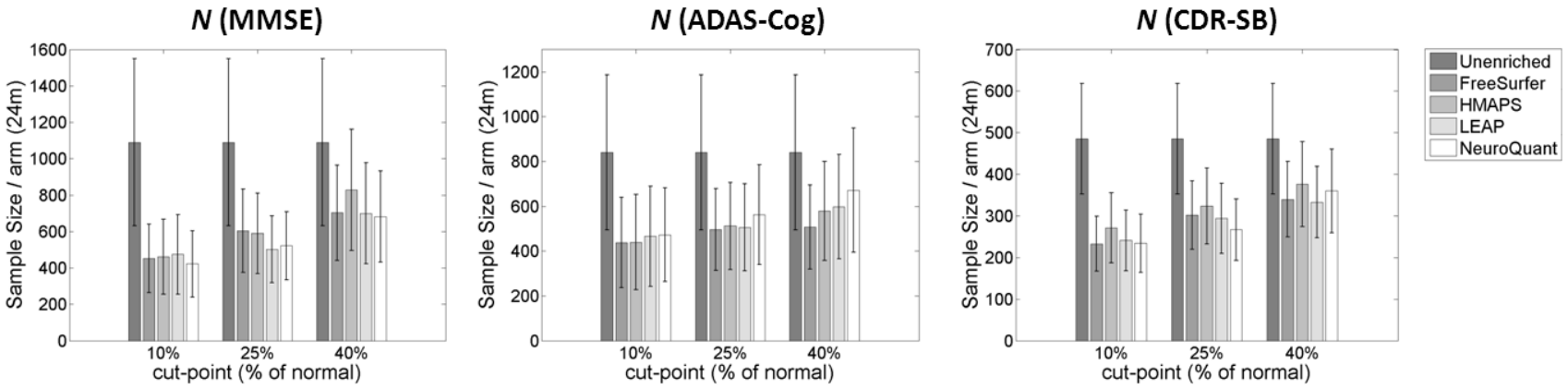
Cut-point defined with respect to normative reference range



How do the enriched trial characteristics depend on the choice of cut-point?

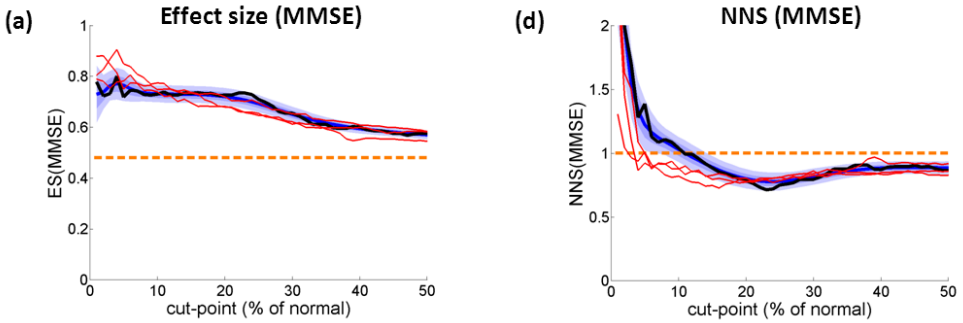
MCI subject selection based on low hippocampal volume results in smaller sample sizes

This improvement is not sensitive to algorithm and is maintained across a range of cut-points.

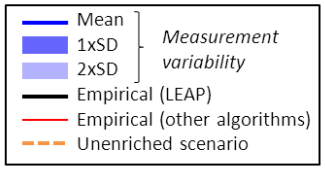


Yu P *et al*. (2014) *Neurobiol. Aging* **35** 808

Enriched population yields smaller sample size but increased screen fail rate → implications for clinical trial operations

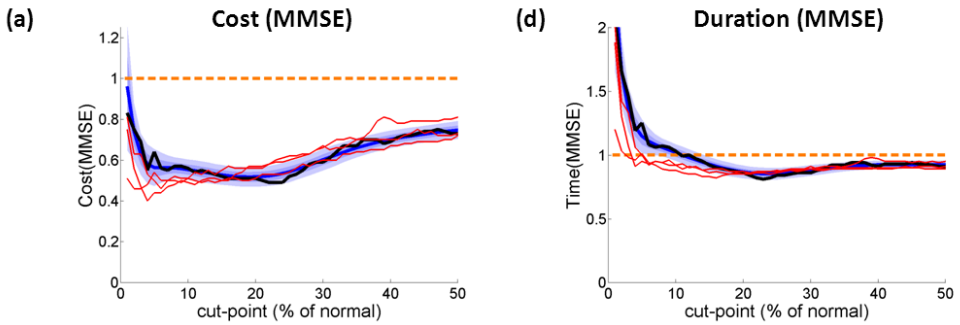


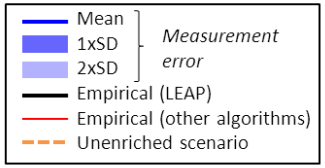
← NNS = Number needed to screen (to enroll projected sample size)



Yu P *et al*. (2014) *Neurobiol. Aging* **35** 808

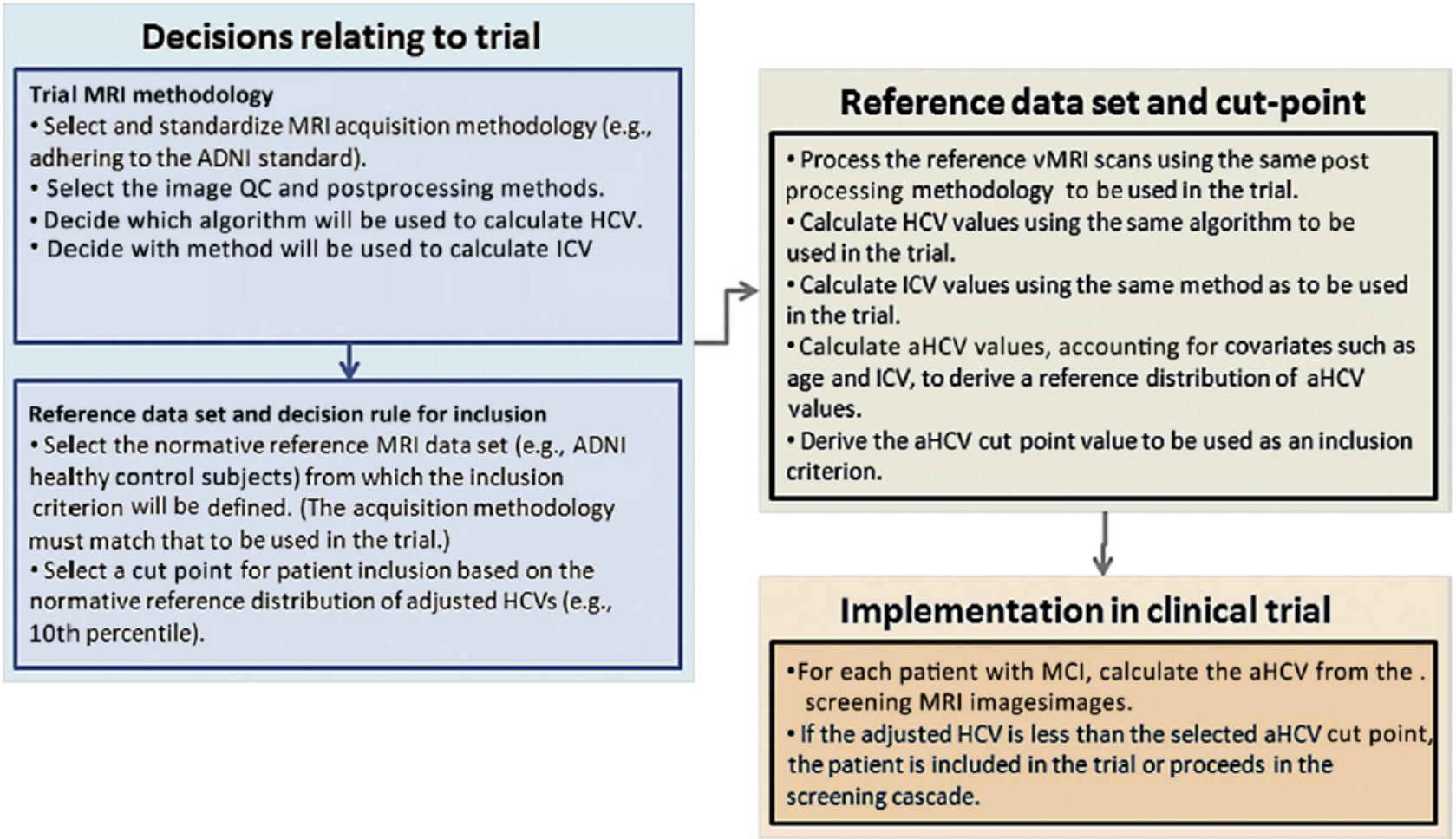
Enriched population yields smaller sample size but increased screen fail rate → implications for clinical trial operations





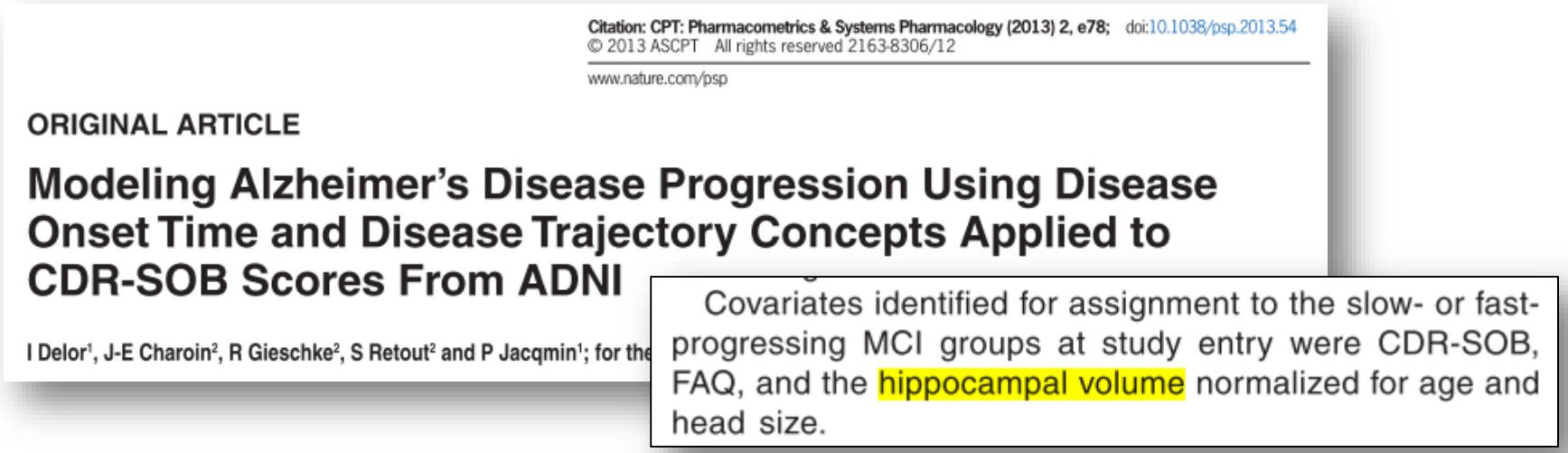
Yu P *et al*. (2014) *Neurobiol. Aging* **35** 808

An operational recipe for the use of HCV to enrich clinical trials



Hill DLG *et al*. (2014) *Alzheimers & Dementia* **10** 421

Gantanerumab MCI *post hoc* analysis (SCarlet RoAD)



Citation: CPT: Pharmacometrics & Systems Pharmacology (2013) 2, e78; doi:10.1038/psp.2013.54

www.nature.com/psp

ORIGINAL ARTICLE

Modeling Alzheimer's Disease Progression Using Disease Onset Time and Disease Trajectory Concepts Applied to CDR-SOB Scores From ADNI

I Delor[1], J-E Charoin[2], R Gieschke[2], S Retout[2] and P Jacqmin[1]; for the

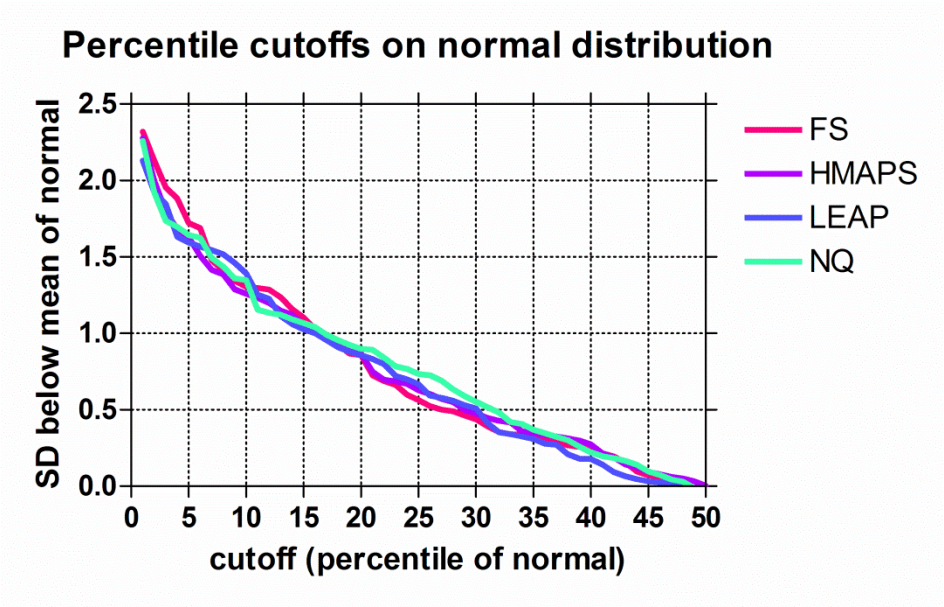
Covariates identified for assignment to the slow- or fast-progressing MCI groups at study entry were CDR-SOB, FAQ, and the hippocampal volume normalized for age and head size.

"Different progression rates from person to person, and the field's inability to predict with any precision how quickly a given person will progress, are longstanding problems in Alzheimer's disease trials. In this instance, the fast progressors—i.e., those whose hippocampal volume and CDR-SB performance declined the most over the duration of the trial—appeared to benefit […]"
http://www.alzforum.org/news/conference-coverage/aducanumab-solanezumab-gantenerumab-data-lift-crenezumab-well
AAIC-2015

Summary

- Evidentiary considerations and research guidelines relevant to the context of use were reviewed
- Key evidentiary questions to be addressed by a putative biomarker include:
  - Heterogeneity of the clinically-defined target population
  - Strength of supporting data and robustness of findings across different studies, cohorts, geographies
  - Test-retest of the method *per se*
  - Sensitivity to technical variations
  - Operational considerations (including time and cost)
- Hippocampal volume (HV) provides a case study of a neuroimaging enrichment biomarker for MCI due to AD, for which the above points have been addressed
- Biomarker qualification could improve chance of success, reduce number of subjects exposed to an experimental treatment that may have side effects, and reduce time/cost of trials.

- Backup



- 10th percentile ~ 1.3 SD below normal mean
- 25th percentile ~ 0.6 SD below normal mean
- 40th percentile ~ 0.2 SD below normal mean