Some Considerations about Targeted Clinical Trials Designs Based on Biomarkers

Lon S. Schneider, MD, MS

~~Data core leader –~~ NOT

Clinical core leader - Yes

University of Southern California ADRC, Los Angeles

ADC Data Core Leaders' Meeting

Boston

October 6, 2012

Keck School of Medicine of USC

Disclosures

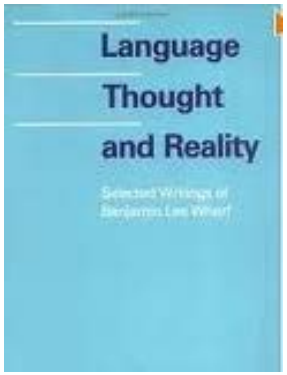
- Grant/Research Support
  - Industry:
    - Pfizer (ADCS), Baxter (ADCS), Genentech, Eli Lilly, Novartis
  - NIH:
    - USC ADRC, ADCS, ADNI, phytoSERMs, AD trials simulations, others
  - State of California:
    - AD Program, California Institute for Regenerative Medicine (CIRM)
- Consultant
  - Industry:
    - Abbott, AC Immune, Accera, Allergan, Allon, AstraZeneca, BiogenIdec, Elan, EnVivo, Exonhit, Genentech, GSK, Helicon, Ipsen, J & J, Kirin, Eli Lilly, FDA, Lundbeck, Merck, Myriad, Pain, Pfizer, Roche, Schwabe, Servier, SK Life Sciences, Takeda, Targacept, TauRx, Toyama, Zinfandel
  - Other:
    - Editorial boards for *The Lancet Neurology, Cochrane Collaboration*; other journals; the APA practice guidelines in AD, and guidelines committees for the American Association for Geriatric Psychiatry, American Geriatrics Society, World Federation of Societies of Biological Psychiatry; expert witness or consultant in civil and criminal federal and state cases for plaintiffs against Lilly, J & J, Pfizer; and for defendants AstraZeneca, Pfizer, California Dept of Justice, and US attorneys general

Disclaimers

- All observers are not led by the same physical evidence to the same picture of the universe

  -- Benjamin Lee Whorf

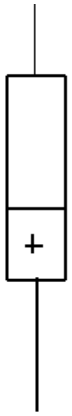
  Theory of linguistic relativity (1940)

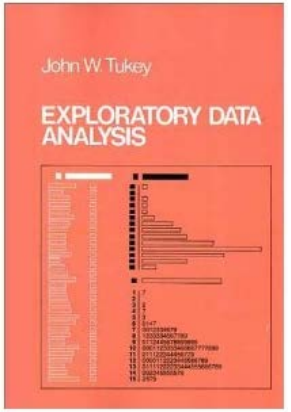


- It is often much worse to have good measurement of the wrong thing--especially when, as is so often the case, the wrong thing will ***in fact*** be used as an indicator of the right thing--than to have poor measurement of the right thing.

  -- John W. Tukey

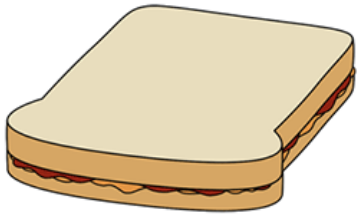
  Exploratory Data Analysis (1977)





- Ham sandwich theorem

  -- Stone and Tukey (1943)



PART 1

Issues with MCI and AD Trials

- Various expert-driven new diagnostic criteria that are biomarker-dependent and not yet determined to be helpful for trials
- Too many, (mainly) industry-driven drug targets
- Various expert- and historically-driven clinical outcomes that are selectively employed in trials
- Several ‘standard,’ but unvalidated-for-purpose, biomarkers are used variously for diagnosis, predicting outcomes, and as surrogate or supportive outcomes

Diagnostic Consensus?

Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria

Bruno Dubois*, Howard H Feldman*, Claudia Jacova, Steven T DeKosky, Pascale Barberger-Gateau, Jeffrey Cummings, André Delacourte, Douglas Galasko, Serge Gauthier, Gregory Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Yaakov Stern, Pieter J Visser, Philip Scheltens

The NINCDS–ADRDA and the DSM-IV-TR criteria for Alzheimer's disease (AD) are
standards in research; however, they have now fallen behind the unprecedented growth
Distinctive and reliable biomarkers of AD are now available through structural MRI, mole
PET, and cerebrospinal fluid analyses. This progress provides the impetus for our prop
criteria for AD. Our framework was developed to capture both the earliest stages, before ful
as the full spectrum of the illness. These new criteria are centred on a clinical core of earl
memory impairment. They stipulate that there must also be at least one or more abn
structural neuroimaging with MRI, molecular neuroimaging with PET, and cerebrospinal
or tau proteins. The timeliness of these criteria is highlighted by the many drugs in develo

Revising the definition of Alzheimer's disease: a new lexicon

Bruno Dubois, Howard H Feldman, Claudia Jacova, Jeffrey L Cummings, Steven T DeKosky, Pascale Barberger-Gateau, André Delacourte, Giovanni Frisoni, Nick C Fox, Douglas Galasko, Serge Gauthier, Harald Hampel, Gregory A Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Marie Sarazin, Leonardo C de Souza, Yaakov Stern, Pieter J Visser, Philip Scheltens

Lancet Neurol 2010; 9: 1118–27
Published Online October 11, 2010
DOI:10.1016/S1474-4422(10)70223-4
Reflection and Reaction page 1044

& Marie Curie University, s, Research Centre of the nstitute of the Brain and pinal Cord, Institute for Memory and Alzheimer's

Alzheimer's disease (AD) is classically defined as a dual clinicopathological entity. The recent advances in use of reliable biomarkers of AD that provide in-vivo evidence of the disease has stimulated the development of new research criteria that reconceptualise the diagnosis around both a specific pattern of cognitive changes and structural/biological evidence of Alzheimer's pathology. This new diagnostic framework has stimulated debate about the definition of AD and related conditions
urgency to this debate.
AD aims to advance th
and by proposing a com
of this lexicon is to con
dementia phases.

Journal of Internal Medicine 2004; 256: 240–246

KEY SYMPOSIUM

Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment

B. WINBLAD[1], K. PALMER[2], M. KIVIPELTO[2], V. JELIC[1], L. FRATIGLIONI[2], L.-O. WAHLUND[1], A. NORDBERG[3], L. BÄCKMAN[2], M. ALBERT[4], O. ALMKVIST[1], H. ARAI[5], H. BASUN[6], K. BLENNOW[7], M. DE LEON[8], C. DECARLI[9], T. ERKINJUNTTI[10], E. GIACOBINI[11], C. GRAFF[12], J. HARDY[13], C. JACK[14], A. JORM[15], K. RITCHIE[16], C. VAN DUIJN[17], P. VISSER[18] & R.C. PETERSEN[19]
[1]Division of Geriatric Medicine, Neurotec Department, Karolinska Institutet, Stockholm, Sweden; [2]Aging Research Center, Division of Geriatric

ELSEVIER
Alzheimer's & Dementia
Alzheimer's & Dementia 7 (2011) 263–269

The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease

Guy M. McKhann[a,b,*], David S. Knopman[c], Howard Chertkow[d,e], Bradley T. Hyman[f], Clifford R. Jack, Jr.[g], Claudia H. Kawas[h,i,j], William E. Klunk[k], Walter J. Koroshetz[l], Jennifer J. Manly[m,n,o], Richard Mayeux[m,n,o], Richard C. Mohs[p], John C. Morris[q], Martin N. Rossor[r], Philip Scheltens[s], Maria C. Carrillo[t], Bill Thies[t], Sandra Weintraub[u,v], Creighton H. Phelps[w]

ELSEVIER
Alzheimer's & Dementia 7 (2011) 270–279

The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease

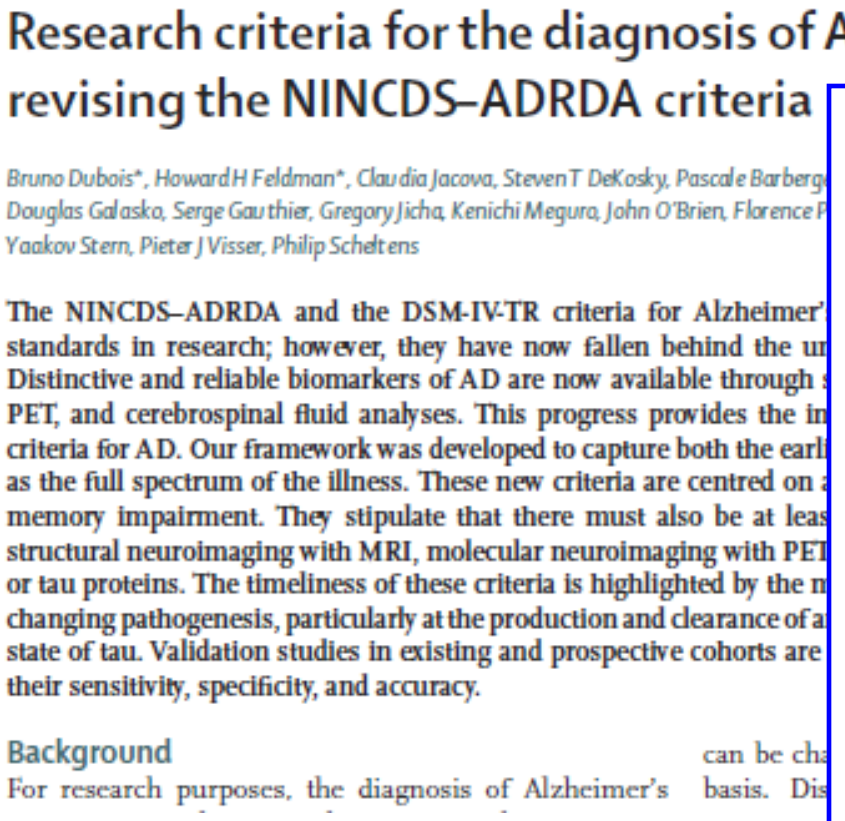
Marilyn S. Albert[a,*], Steven T. DeKosky[b,c], Dennis Dickson[d], Bruno Dubois[e], Howard H. Feldman[f], Nick C. Fox[g], Anthony Gamst[h], David M. Holtzman[i,j], William J. Jagust[k], Ronald C. Petersen[l], Peter J. Snyder[m,n], Maria C. Carrillo[o], Bill Thies[o], Creighton H. Phelps[p]
[a]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[b]Office of the Dean, University of Virginia, Charlottesville, VA, USA
[c]Department of Neurology, University of Virginia, Charlottesville, VA, USA
[d]Department of Pathology, Mayo Clinic, Jacksonville, FL, USA

ELSEVIER
Alzheimer's & Dementia
Alzheimer's & Dementia 7 (2011) 280–292

Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease

Reisa A. Sperling[a,*], Paul S. Aisen[b], Laurel A. Beckett[c], David A. Bennett[d], Suzanne Craft[e], Anne M. Fagan[f], Takeshi Iwatsubo[g], Clifford R. Jack, Jr.[h], Jeffrey Kaye[i], Thomas J. Montine[j], Denise C. Park[k], Eric M. Reiman[l], Christopher C. Rowe[m], Eric Siemers[n], Yaakov Stern[o], Kristine Yaffe[p], Maria C. Carrillo[q], Bill Thies[q], Marcelle Morrison-Bogorad[r], Molly V. Wagster[r], Creighton H. Phelps[r]

Revised Research Criteria for AD (2007)



Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria

Bruno Dubois*, Howard H Feldman*, Claudia Jacova, Steven T DeKosky, Pascale Barberg
Douglas Galasko, Serge Gauthier, Gregory Jicha, Kenichi Meguro, John O'Brien, Florence P
Yaakov Stern, Pieter J Visser, Philip Scheltens

The NINCDS–ADRDA and the DSM-IV-TR criteria for Alzheimer'
standards in research; however, they have now fallen behind the ur
Distinctive and reliable biomarkers of AD are now available through
PET, and cerebrospinal fluid analyses. This progress provides the in
criteria for AD. Our framework was developed to capture both the earli
as the full spectrum of the illness. These new criteria are centred on
memory impairment. They stipulate that there must also be at leas
structural neuroimaging with MRI, molecular neuroimaging with PET
or tau proteins. The timeliness of these criteria is highlighted by the n
changing pathogenesis, particularly at the production and clearance of a
state of tau. Validation studies in existing and prospective cohorts are
their sensitivity, specificity, and accuracy.

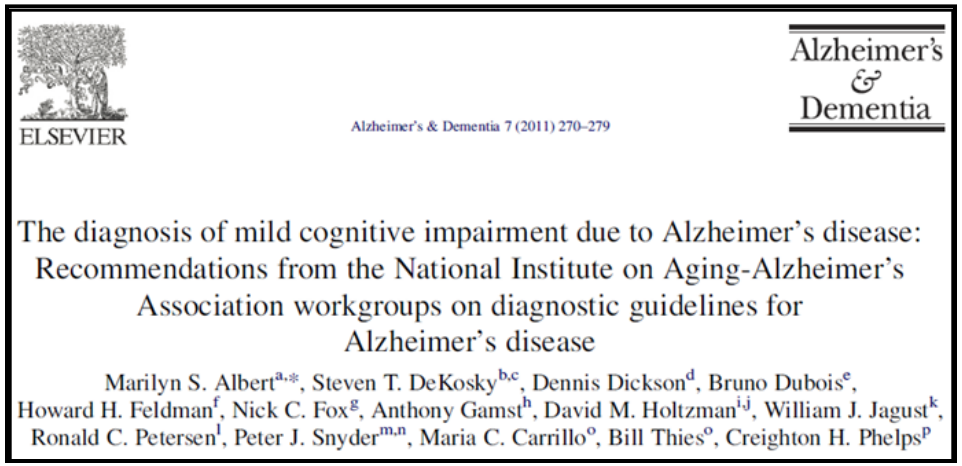
Background

For research purposes, the diagnosis of Alzheimer's can be cha basis. Dis

- The *International Working Group for New Research Criteria for the Diagnosis of AD* :
- Framework to capture the earliest stages…
- Must be at least one or more abnormal biomarkers
- Timeliness is highlighted by the many drugs in development that are directed at particularly at the production and clearance of Aβ
- Validation studies … are needed to advance these criteria

Consensus on Diagnosis (2011)?

- *The* Alzheimer's Association criteria
  - Preclinical
  - MCI associated with AD
  - AD (The new McKhann et al criteria)
- Common elements are specific putative biomarkers



ELSEVIER

Alzheimer's & Dementia

Alzheimer's & Dementia 7 (2011) 270–279

The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease

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]

Panel: Biomarkers of pathophysiology in the revised NIA-AA diagnostic criteria[4-6]

Biomarkers of brain β amyloidosis

- Increased uptake on amyloid imaging with PET*
- Decreased CSF amyloid β_{42}*

Biomarkers of neuronal injury (synaptic dysfunction and neuronal degeneration)

- Temporoparietal hypometabolism on ^{18}F-fluorodeoxyglucose PET*
- Medial temporal (hippocampal) atrophy*
- Increased CSF tau/phospho-tau*
- Temporoparietal hypoperfusion on single-photon emission CT

Other less validated biomarkers, biomarkers of collateral damage, or serial biomarkers

- Functional MRI activation studies, resting blood oxygen level-dependent functional connectivity, MRI perfusion, MR spectroscopy, diffusion tensor imaging
- Inflammatory (cytokines) and oxidative stress biomarkers (isoprostanes)
- Rates of brain atrophy

NIA-AA=National Institute on Aging and the Alzheimer's Association. *Markers included in an early proposal for revised criteria by Dubois and colleagues.[7]

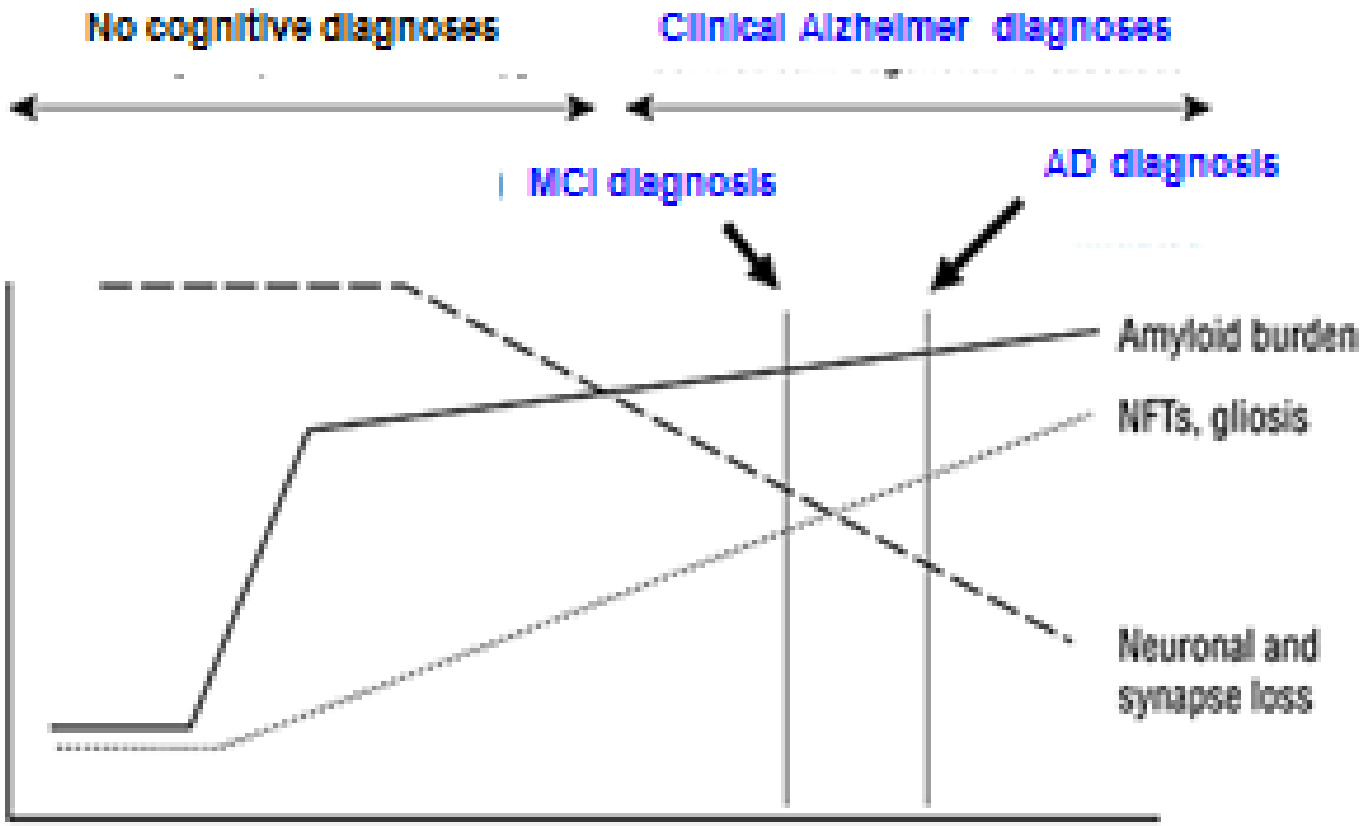
(Too Many?) Potential Aβ Targets for 'Disease-modifying' Therapies for AD

- Aβ production
  - α -secretase enhancement
  - β -secretase inhibition
  - γ -secretase inhibition
  - γ -secretase modulation
- Aβ degradation
  - Neprilysin activation
  - Insulin-degrading enzyme (IDE) activation
- Aβ removal
  - Vaccination
  - Passive immunization
  - Enhance receptor-mediated removal from CNS
  - Prevent entry from periphery
- Preventing Aβ toxicity
  - Prevent aggregation via Aβ binding
  - Prevent oligomerization (e.g., metal atteuation of proteins)
- Tau
  - Prevent tau aggregation, prevent tau hyperphosphorylation, facilitate tau phosphatases, stabilize microtubules
- Neuroprotection
  - Growth factor treatment or GF receptors activation, anti-apoptotic agents, metabolic/mitochondrial agents, block inflammation disease processes
- Neuroregeneration

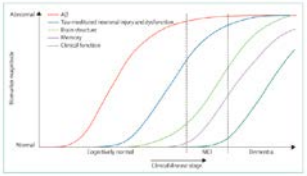
Issues to Address

- Causes of AD are not known
- No validated drug targets
- Where and when along the process could drugs work?
  - Early or late in the amyloid cascade?
  - Early or late in the clinical course?
  - An early intervention may not show a discernable effect for years or could show one immediately
- Do targets change over course of illness?
- Do outcomes differ for any given drug?

‘Standard’ biomarkers are informed by amyloid burden, tangles, and neuron loss



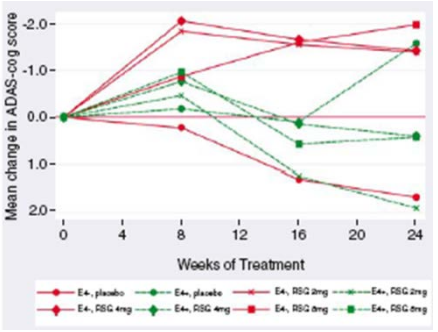
CSF Aβ
Aβ PET
CSF t-tau
CSF p-tau
MRI volumes

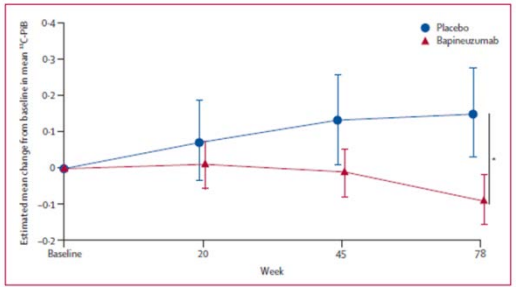


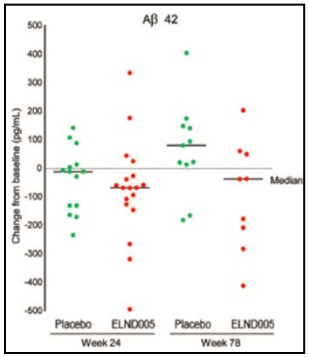
Modified from Hyman et al 2011

Phase 2 and 3 Trials Use Biomarkers to Predict or Assess Outcomes

- Rosiglitazone (Avandia), phase 2 and 3
  - ApoE carrier status
- Semagacestat, phase 2 and 3
  - CSF Aβ and tau
    - 40% decrease in Aβ in blood not in CSF
- Bapineuzumab, phase 2 and 3
  - Aβ PET, APOE
- Scyllo-inositol, phase 2
  - CSF Aβ and tau
- Solanezumab, phase 2 and 3
  - CSF Aβ and tau







Ongoing phase 2 targeted designs uses biomarker for entry

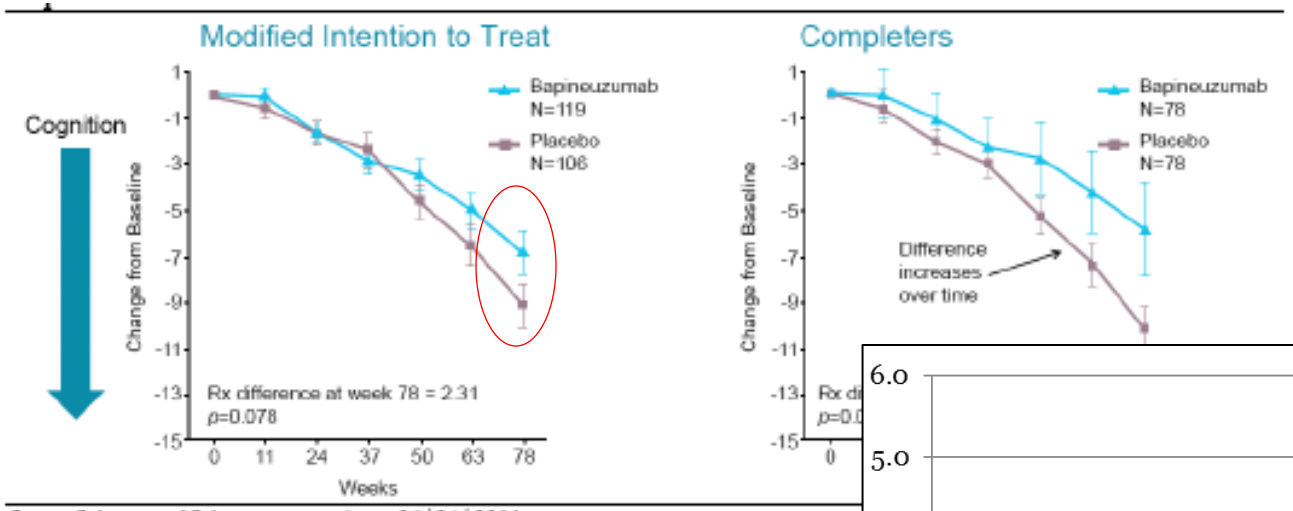
- GSI: BMS 708163
  - Prodromal AD, MMSE 24-30, plus Aβ +
  - 75 sites, N = 270, 1 dose and placebo, 2-year follow-up (but 2 highest doses dropped)
  - Primary: safety and CSF markers
  - "ADNI knock-off"
- mAb: Gantenerumab (Roche)
  - Prodromal AD, MMSE > 23, Aβ-PET positive
  - 63 ex-US sites, N = 360, 2 doses and placebo, 2-year follow-up
  - Primary: CDR-sb and Aβ change

- Is this the new normal for trials?

Trials Outcomes Analyzed by ApoE Status

- Rosiglitazone
- Tarenflurbil
- Bapineuzumab phase 2
- Bapinuzumab phase 3
- Future trials
  - Solanezumab
  - BMS
  - Roche
  - Pioglitazone (Takeda/Zinfandel)
- Apparent rationale:
  - To suggest differential outcomes with biomarkers
  - The idea that there must be subgroups of drug-responsive patients

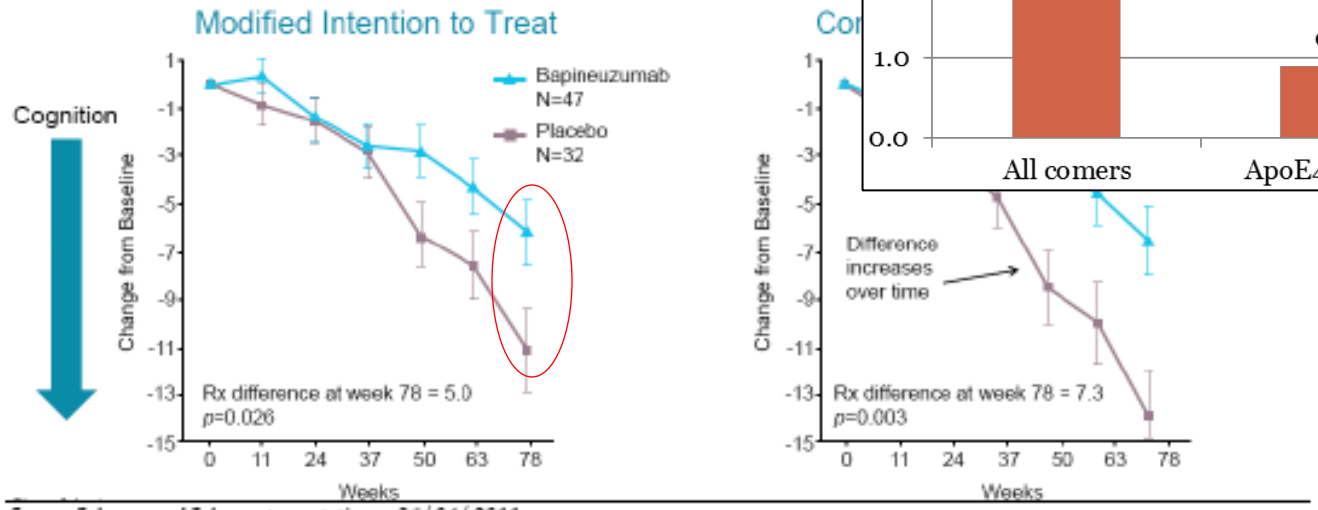
Bapineuzumab 201 Trial



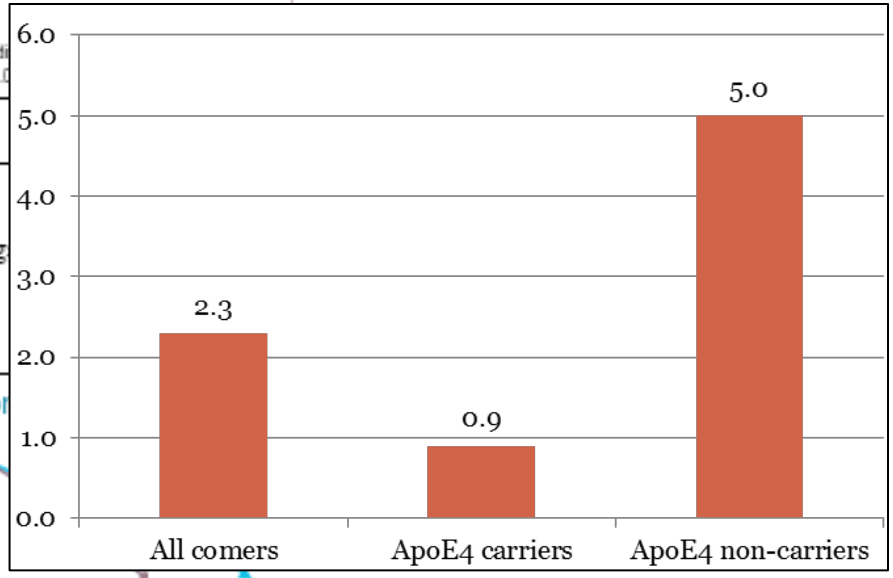
Source: Johnson and Johnson presentation – 26/04/2011

Bapineuzumab also showed a statistically significant impact on cog
ApoE4 non-carriers.

Bapineuzumab Phase II data in ApoE4 non-carriers



Source: Johnson and Johnson presentation – 26/04/2011

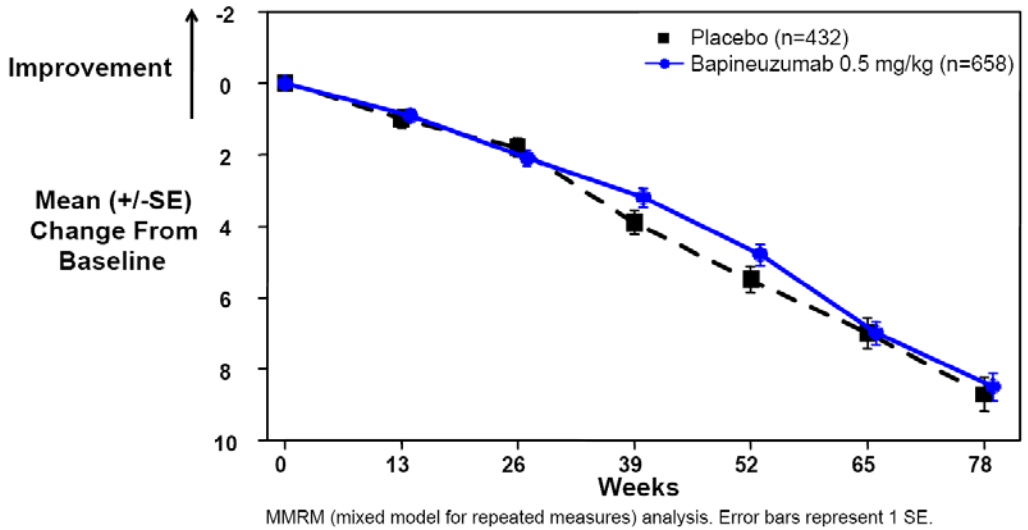


Bapineuzumab 301 and 302 Trials

Change in ADAS-Cog 11 by Treatment Group Over 78 Weeks (APOE ε4 Carriers)

(mITT population)

Treatment Difference at Week 78		
Bapineuzumab	Mean (95% CI)	**p-value**
0.5 mg/kg	-0.2 (-1.4, 1.0)	**0.798**



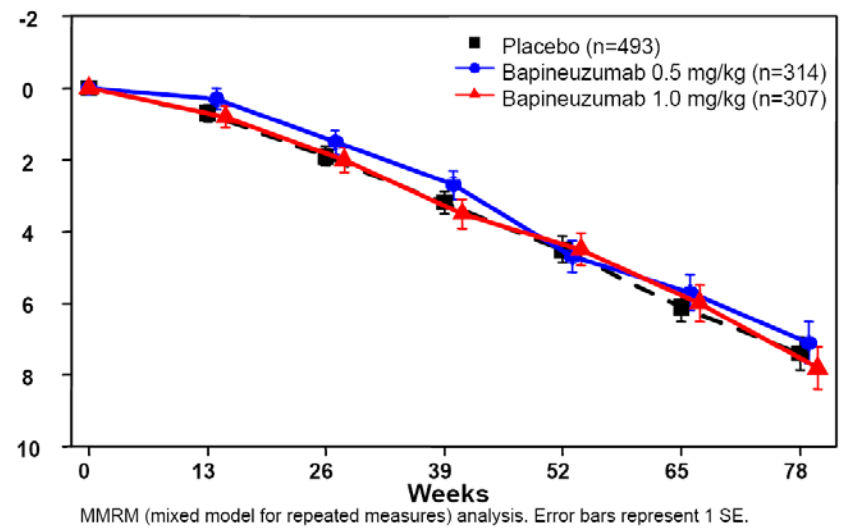
MMRM (mixed model for repeated measures) analysis. Error bars represent 1 SE.

APOE ε4 carriers

ıange in ADAS-Cog 11 by Treatment Group Over 78 Weeks (APOE ε4 Non-Carriers)

(mITT population)

Treatment Difference at Week 78		
Bapineuzumab	Mean (95% CI)	**p-value**
0.5 mg/kg	-0.3 (-1.8, 1.1)	**0.642**
1.0 mg/kg	0.4 (-1.1, 1.8)	**0.620**



MMRM (mixed model for repeated measures) analysis. Error bars represent 1 SE.

APOE ε4 non-carriers

Salloway et al 2012, Sperling et al 2012

PART 2

Simulating Stratified Medicine, Targeted Design Trials in MCI and AD with Aβ and ApoE Markers

Lon S. Schneider, MD, Richard E. Kennedy, MD, PhD, Gary R. Cutter, PhD

University of Southern California Keck School of Medicine, Los Angeles, CA,
University of Alabama at Birmingham, Birmingham, AL, USA.

ISCTM Autumn Meeting
October 3, 2012
Marina del Rey, California

Keck School of Medicine of USC

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Background for Aβ Targeted Design

- Proposed criteria for prodromal or early AD requires a positive biomarker
- *ad hoc* groups recommend that clinical trials for prodromal AD would be more efficient if a CSF $A\beta_{42}$ biomarker were required
  - "to show a 40% reduction in progression on ratings, with 80% power, an alpha error P ≤ 0.05, and a 2-year drop-out rate < 40% would require about 100 or 150 patients for one or another primary outcome per group when patients are selected [using a CSF $A\beta_{42}$ biomarker] compared to twice as many without the biomarker criteria"
- Therefore: One should test the potential efficiency of these recommendations by simulating a range of clinical trials scenarios

Methods: patient selection and methods

- Use clinical trials datasets to select subjects fulfilling certain clinical trials criteria
- Amnestic MCI criteria or MCI due to AD selected as though they were applying for clinical trials:
    - (1) aMCI diagnosis as above
    - (2) aMCI with CSF $A\beta_{42} \leq 192$ pG/mL
    - (3) aMCI with t-tau/$A\beta_{42} > 0.39$
  - Latter two criteria are expert-recommended = “prodromal AD”
- **Outcomes**: ADAS-cog and CDR-sb performed at 6-month intervals
- **Clinical trials scenarios:**
  - Sample sizes of 50, 100, 200, and 400 per group
  - 12 and 24 month long trials
  - Dropout rates of 20% and 40% in both groups incorporated into scenarios
- **Placebo group outcome:**
  - the score for patient at the specified time point in the ADNI database
- **Treatment group outcome:**
  - effect sizes from 0.15 to 0.75 (i.e., very small to moderately large)
- **For each patient:**
  - Treatment effect randomly generated from a X^2 distribution with mean equal to expected effect
  - Each effect was shifted by subtracting 2 times the expected effect, then adding the result to the patient’s score at the specified time point in the database
  - Even when a patient was reused in the analysis the actual value used would be modified by this randomly selected amount

Methods: statistical analysis

- Primary analyses: Mixed effects linear model (covariance pattern model) which adjusts for missing data to test for differences
  - A full model used with group effect, visit effect, and group by visit interactions, with age and gender as covariates, and a reduced model with visit, age, and gender effects. A compound symmetric covariance structure was used to model the correlation between visits for each participant. Parameters estimated using maximum likelihood
  - P-values for the group (treatment) effect were found using -2 times the difference in the log likelihood of the models which follows a X^2-square distribution with the appropriate degrees of freedom
- Secondary analyses: LOCF and complete cases (not further discussed)
- The missing data pattern present in ADNI was used to simulate dropouts
- 1000 simulations for each scenario to estimate power to 3 digits
- Power = proportion of 1000 simulated trials per scenario with α error $p \leq 0.05$
- Analyses R 2.10.1 and R nlme package 3.1-89
- Data downloaded Dec 7, 2009: http://www.loni.ucla.edu/twiki/bin/view/ADNI/ADNIClinicalFAQ

Schneider et al Alzh & Dem 2010

Results: Characteristics and ratings by selection criteria

(199 of the 400 aMCI patients had CSF examinations)

	No marker required	Low $A\beta_{42}$	High t-tau/$A\beta_{42}$	P Value
N	400	148	137	
Age, years, mean, SD	74.92 (7.41)	74.66 (7.09)	74.66 (7.45)	0.84
Gender, male %	64.5%	64.9%	62.8%	0.92
Education, college %	64.3%	62.8%	62.0%	0.92
APOE e4 genotype %	**54.0%**	**64.2%**	**66.0%**	**0.17**
MMSE, baseline, mean (SD)	27.01 (1.78)	26.79 (1.79)	26.83 (1.82)	0.32
CDR-sb, baseline, mean (SD)	1.61 (0.88)	1.65 (0.91)	1.63 (0.89)	0.95
CDR-sb, 12mo., mean (SD)	2.27 (1.52)	2.51 (1.39)	2.51 (1.42)	0.03
CDR-sb, 24mo., mean (SD)	3.06 (2.23)	3.44 (2.14)	3.49 (2.15)	0.03
ADAS-cog, baseline, (SD)	**11.56 (4.42)**	**12.25 (4.54)**	**12.41 (4.55)**	**0.07**
ADAS-cog, 12mo., mean (SD)	12.55 (6.19)	13.34 (5.93)	13.59 (5.92)	0.06
ADAS-cog, 24mo., mean (SD)	**14.12 (7.43)**	**15.76 (7.08)**	**15.85 (7.12)**	**0.01**
Dementia, 24 mo., mean (SD)	28.5%	35.8%	38.0%	0.23

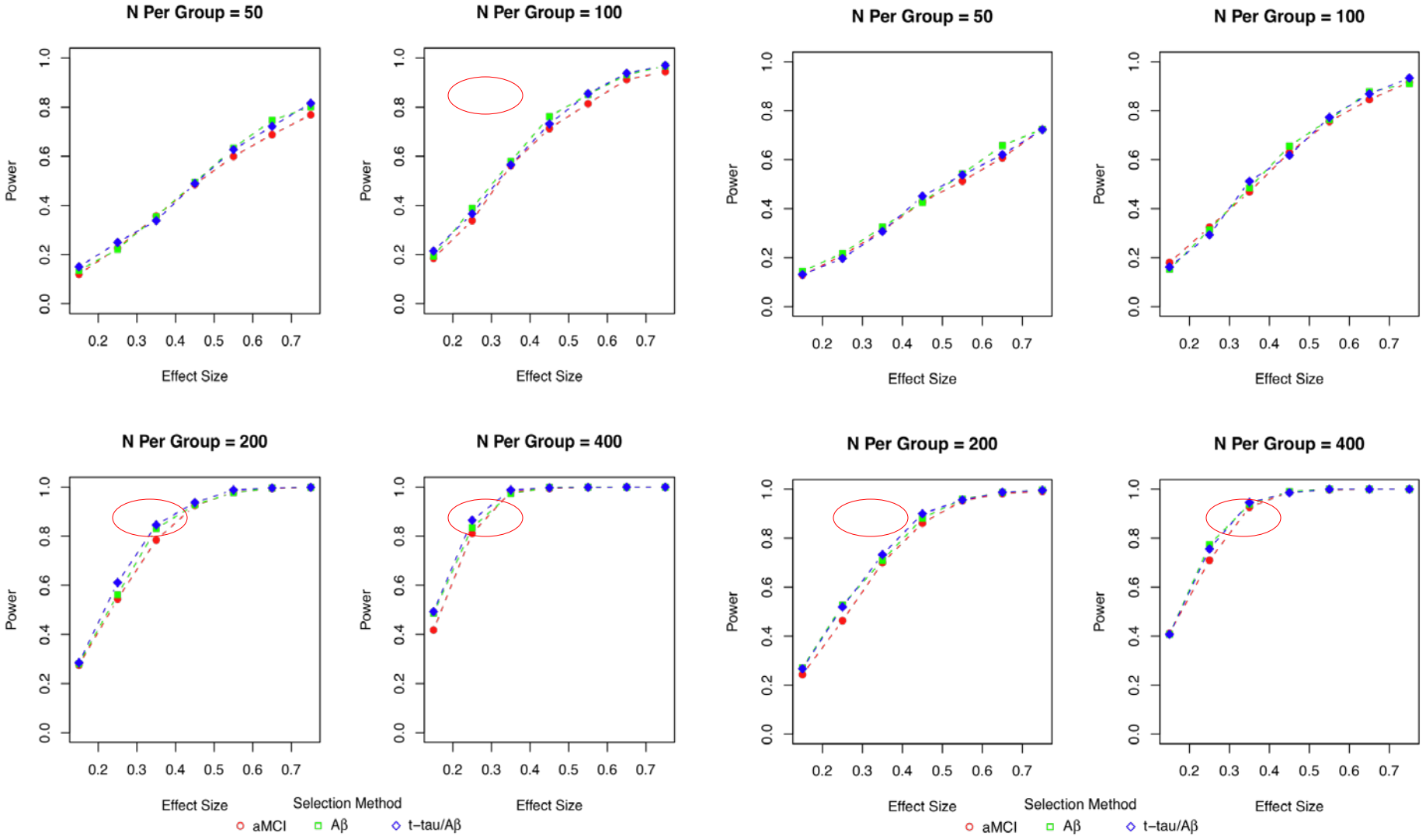
- > 95% classified as 'MCI due to AD,' 58% with FH of dementia
- 44.0% used ChEIs, 9% ChEIs+memantine; 53.5% neither
- 96%, 90%, 81%, and 72% had outcomes at 6-, 12-, 18- and 24- mo

Schneider et al Alzh & Dem 2010

Power for ADAS-cog outcomes in 24-month trials

20% dropouts

40% dropouts



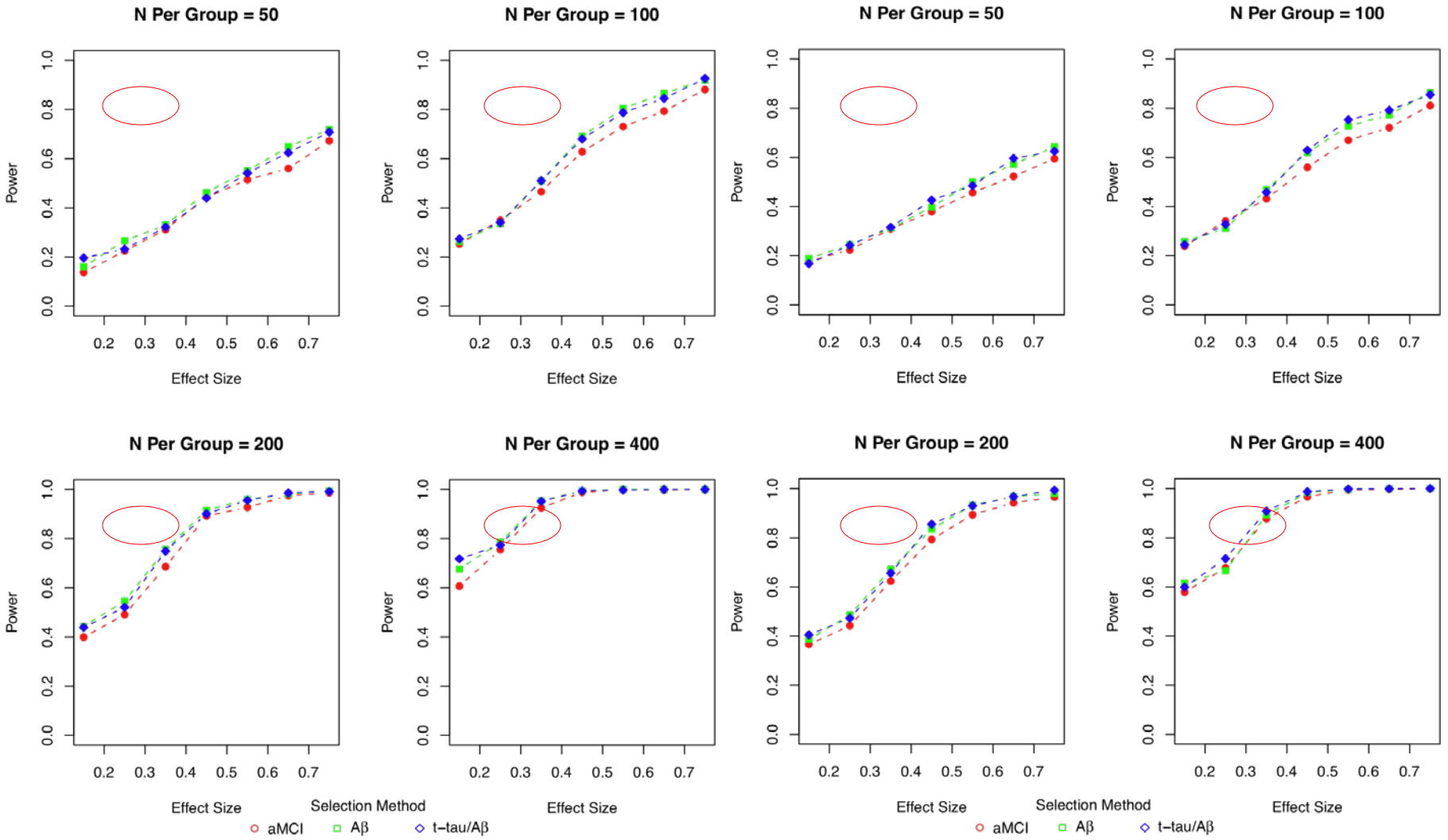
●= aMCI ■ = aMCI + low $A\beta_{42}$ ◆ = aMCI + high t-tau/ $A\beta_{42}$

Schneider et al Alzh & Dem 2010

Power for CDR-sb outcomes in 24-month trials

20% dropouts

40% dropouts



●= aMCI ■ = aMCI + low $A\beta_{42}$ ◆ = aMCI + high t-tau/ $A\beta_{42}$

Schneider et al Alzh & Dem 2010

Power for ADAS-cog in 24-month trials

N per Group	Dropout %	Effect Size	Selection Method	Treatment Group Mean	Placebo Group Mean	Treatment Group SD	Placebo Group SD	Power Mixed Model
100	20	0.45	aMCI	0.33	2.85	6.03	5.61	0.71
100	20	0.45	Aβ	1.04	3.73	6.25	5.88	0.76
100	20	0.45	t-tau/Aβ	0.99	3.65	6.41	5.94	0.73
200	20	0.35	aMCI	0.32	2.85	6.08	5.65	0.78
200	20	0.35	Aβ	1.05	3.71	6.28	5.86	0.83
200	20	0.35	t-tau/Aβ	0.96	3.64	6.40	5.95	0.85
200	40	0.35	aMCI	0.89	2.85	5.97	5.65	0.70
200	40	0.35	Aβ	1.65	3.68	6.18	5.86	0.71
200	40	0.35	t-tau/Aβ	1.57	3.65	6.30	5.95	0.73
200	40	0.45	aMCI	0.32	2.87	6.10	5.65	0.86
200	40	0.45	Aβ	1.06	3.70	6.34	5.87	0.88
200	40	0.45	t-tau/Aβ	0.93	3.68	6.36	5.99	0.90
400	20	0.25	aMCI	1.45	2.86	5.92	5.63	0.81
400	20	0.25	Aβ	2.23	3.70	6.15	5.88	0.84
400	20	0.25	t-tau/Aβ	2.17	3.68	6.23	5.98	0.87
400	40	0.25	aMCI	0.86	2.85	6.00	5.66	0.71
400	40	0.25	Aβ	1.67	3.70	6.27	5.89	0.77
400	40	0.25	t-tau/Aβ	1.54	3.68	6.32	6.00	0.76
400	40	0.35	aMCI	1.46	2.86	5.92	5.63	0.93
400	40	0.35	Aβ	2.25	3.73	6.14	5.88	0.94
400	40	0.35	t-tau/Aβ	2.16	3.67	6.23	6.00	0.95

To ensure an approximate power of 80% to 90% for the mixed model analysis, simulations show that for small effects of 0.25, typical to that of cholinesterase inhibitors, somewhat fewer than 400 patients per group are needed with a dropout rate of 20%, and for medium size effects of 0.45, somewhat greater than 100 per group are needed with a dropout rate of 20%. Requiring low $A\beta_{1\text{-}42}$ biomarker ("Aβ") or high total tau to $A\beta_{1\text{-}42}$ ratio ("t-tau/Aβ") in the selection criteria resulted in very small increases in statistical power; these participants showed greater placebo decline but also increased variability, i.e., standard deviation of change. Simulation parameters included α=0.05, effect sizes of 0.15 to 0.75 with Chi-squared random errors, and 20% and 40% dropouts with mixed model analysis for participants with missing data.

Power for CDR-sb outcomes in 24-month

N per Group	Dropout %	Effect Size	Selection Method	Treatment Group Mean	Placebo Group Mean	Treatment Group SD	Placebo Group SD	Power Mixed Model
200	20	0.35	aMCI	0.91	1.48	2.22	1.97	0.69
200	20	0.35	Aβ	1.22	1.83	2.23	1.94	0.76
200	20	0.35	t-tau/Aβ	1.30	1.90	2.23	1.93	0.75
200	20	0.45	aMCI	0.73	1.48	2.22	1.98	0.89
200	20	0.45	Aβ	1.04	1.83	2.26	1.94	0.92
200	20	0.45	t-tau/Aβ	1.11	1.91	2.25	1.93	0.90
200	40	0.45	aMCI	0.73	1.48	2.20	1.96	0.79
200	40	0.45	Aβ	1.03	1.83	2.24	1.94	0.84
200	40	0.45	t-tau/Aβ	1.11	1.92	2.25	1.93	0.86
400	20	0.25	aMCI	1.05	1.48	2.16	1.98	0.76
400	20	0.25	Aβ	1.39	1.83	2.15	1.94	0.79
400	20	0.25	t-tau/Aβ	1.46	1.90	2.15	1.94	0.77
400	20	0.35	aMCI	0.91	1.48	2.24	1.98	0.93
400	20	0.35	Aβ	1.22	1.83	2.23	1.95	0.95
400	20	0.35	t-tau/Aβ	1.30	1.90	2.23	1.93	0.95
400	40	0.25	aMCI	1.05	1.48	2.16	1.98	0.68
400	40	0.25	Aβ	1.39	1.83	2.16	1.93	0.67
400	40	0.25	t-tau/Aβ	1.46	1.91	2.15	1.94	0.72
400	40	0.35	aMCI	0.91	1.48	2.23	1.99	0.88
400	40	0.35	Aβ	1.22	1.83	2.23	1.94	0.89
400	40	0.35	t-tau/Aβ	1.29	1.91	2.23	1.93	0.91

To ensure an approximate power of 80% to 90% for the mixed model analysis, simulations show that for small effects of 0.25, somewhat more than 400 patients per group are needed with a dropout rate of 20%, and for medium size effects of 0.45, somewhat less than 200 per group are needed with a dropout rate of 20%. Requiring low amyloid-$\beta_{1\text{-}42}$ biomarker ("Aβ") or high t-tau/$A\beta_{1\text{-}42}$ ("t-tau/Aβ") as selection criteria resulted in very small increases in statistical power. Gain in power was less prominent as total power increased. Simulation parameters included α=0.05, effect sizes of 0.15 to 0.75 with Chi-squared random errors, and 20% to 40% dropouts analyzed with mixed model analysis for participants with missing data.

Targeted Trials Based on ApoE Genotype

ADCS Studies Used and ADNI

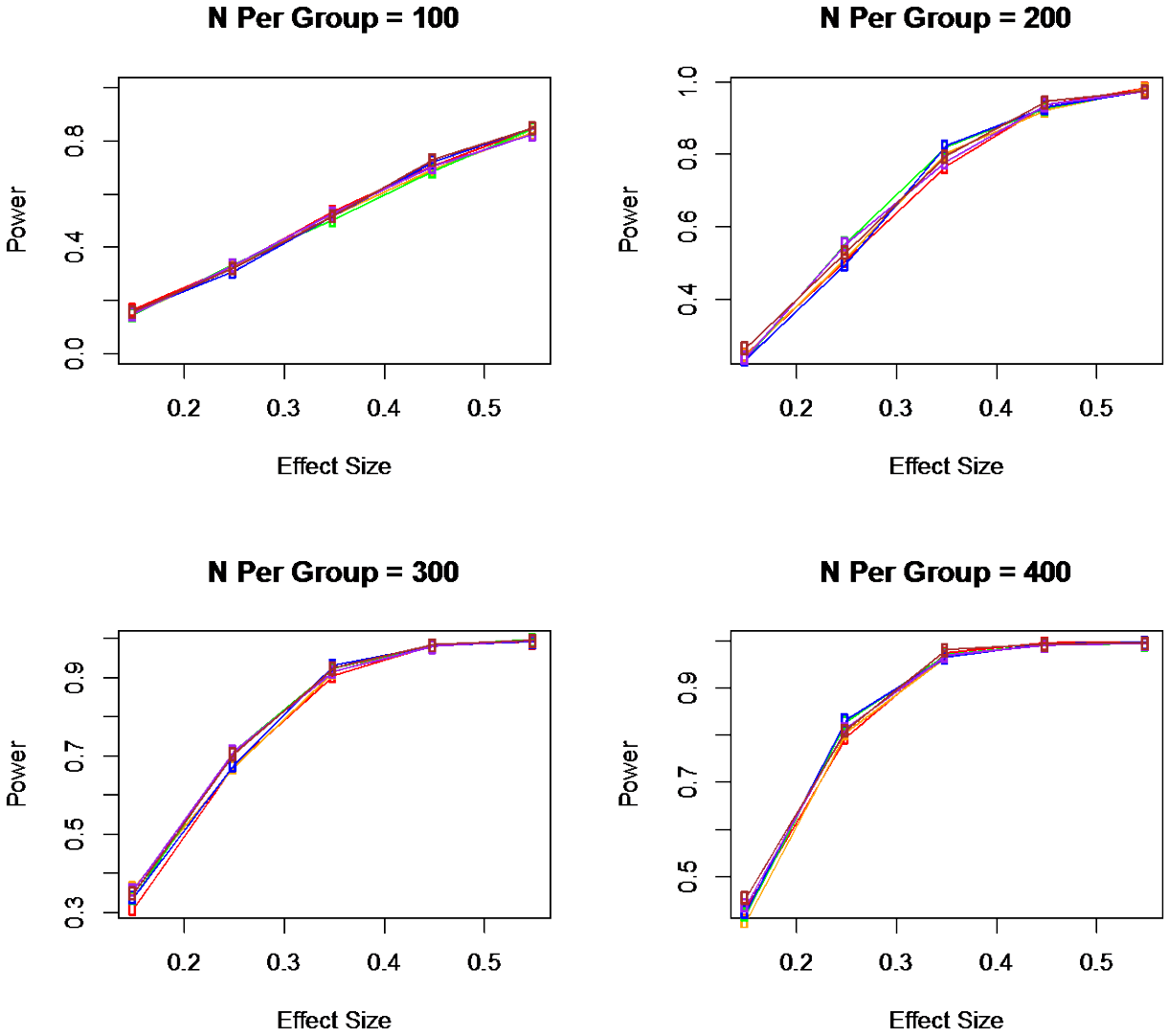
Study, dates	Design	Intervention	N	Duration (mos)
Selegiline, vit E, 1993-1996	severe AD	Vitamin E, selegiline	341	24
Prednisone 1995-1998	mild to mod AD	Prednisone	138	16
CE 1995-1999	mild to mod AD	Conjugated estrogens	120	15
MIS 1999-2004	MCI	Donepezil, vit E	769	36
Simvastatin (LL) 2003-2008	mild to mod AD	Simvastatin	406	18
Vitamins B (HC) 2003-2007	mild to mod AD	B vitamins	409	18
DHA 2006-2009	mild to mod AD	DHA	402	18
ADNI 2005-2010	Observational, mild AD, MCI	None	800	36 (AD) 48 (MCI)

Clinical characteristics among AD and MCI participants by ApoE4 carrier status

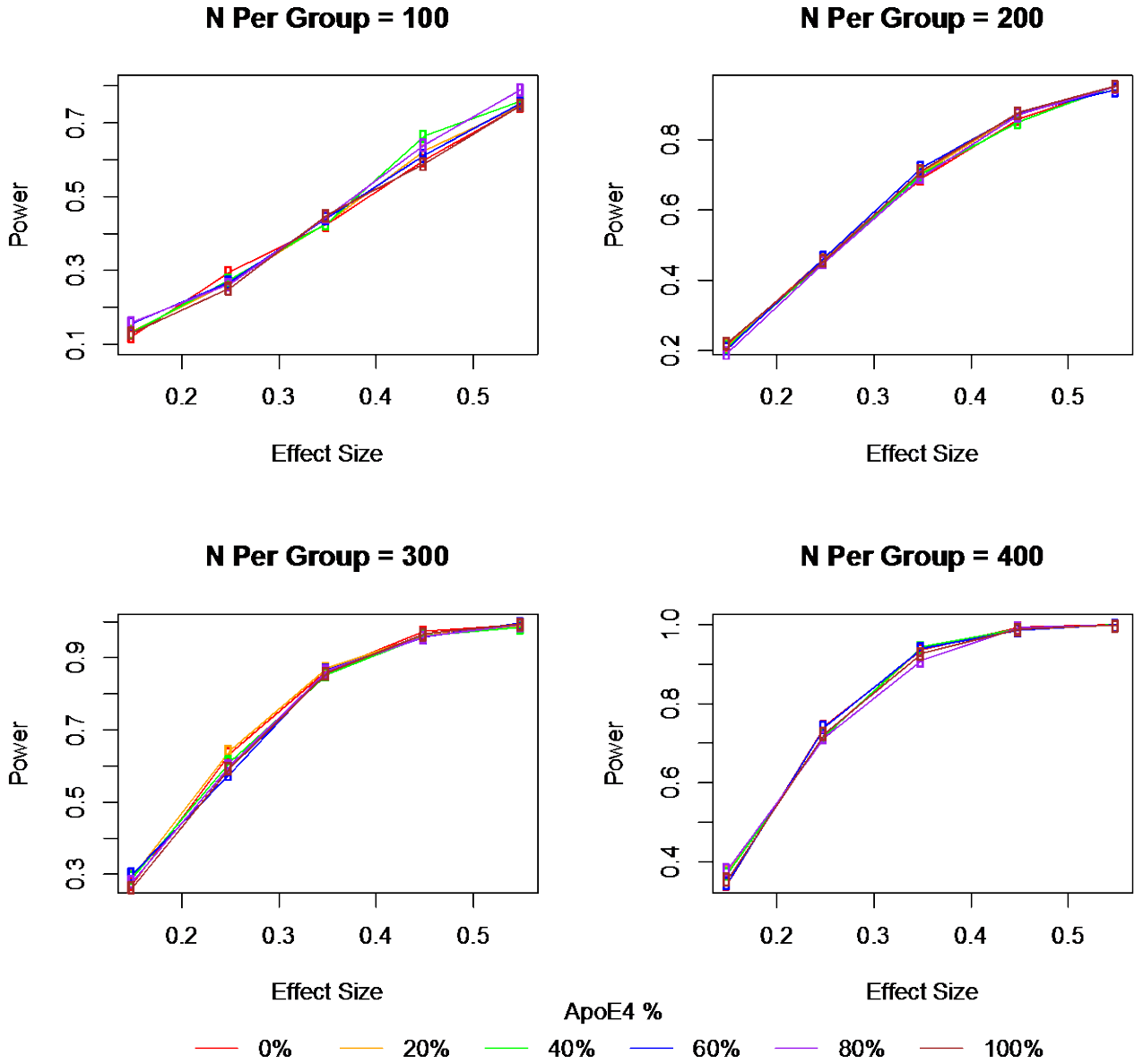
Mild to Moderate AD Overall				
		E4-	E4+	P-value
	N	(N=545)	(N=873)	
Age, years	1368	75.8 (9.5)	74.7 (7.7)	<0.001
Educ, yrs	1374	14.2 (3.3)	14.2 (2.9)	0.9
Hispanic	1374	31 (6%)	32 (4%)	0.077
Married	1411	367 (67%)	654 (75%)	0.001
White	1374	482 (91%)	769 (91%)	0.85
Female	1374	303 (57%)	451 (53%)	0.19
ADAS-Cog				
Baseline	1392	22.3 (9.2)	22.2 (8.7)	0.82
6 mo	1252	23.7 (10.1)	24.3 (9.9)	0.18
12 mo	1129	25 (11)	27 (11)	0.19
18 mo	793	27 (12)	29 (12)	0.042
24 mo	133	26.4 (9.9)	28.8(12.6)	0.57

MCI Overall				
		E4-	E4+	P-value
	N	(N=544)	(N=648)	
Age, yrs	1134	73.4 (8.2)	72.9 (6.6)	0.054
Educ., yrs	1134	15.0 (3.2)	15.0 (3.1)	0.73
Hispanic	1134	27 (5%)	15 (2%)	0.013
Married	1182	394 (73%)	525 (82%)	<0.001
White	1134	468 (91%)	580 (94%)	0.046
Female	1134	206 (40%)	271 (44%)	0.18
ADAS-Cog				
Baseline	402	10.4 (4.2)	12.1 (4.4)	<0.001
6 mo	1038	10.2 (5.2)	12.2 (5.2)	<0.001
12 mo	972	10.6 (5.6)	13.0 (5.9)	<0.001
18 mo	872	10.8 (5.8)	14.0 (7.0)	<0.001
24 mo	814	10.7 (6.5)	14.7 (7.3)	<0.001

Power for ADAS-cog Outcomes in 18-month AD Trials Based on ApoE Genotype



Power for ADAS-cog Outcomes in 24-month MCI Trials Based on ApoE Genotype



A word on effect sizes (and presumed power)

Semantics of Effect Sizes

Table 1. The first two rows simulate clinical trials of prodromal AD (aMCI) and biomarker-positive prodromal AD (low Aβ) with sample sizes of 100 per group, 20% dropouts, effect sizes set at 0.35 SD units, and show resulting percent slope reduction and Power. In the second two rows slope reduction is set at 50% and resulting effect sizes and Power are shown.

	Baseline ADAS-cog	Treatment Change	Placebo Change	Drug-placebo diff.	Effect size	% slope reduction	Power
Setting effect size at 0.35							
aMCI	11.6 (4.4)	0.88 (5.92)	2.86 (5.62)	1.98	0.35	**69.2**	0.56
Low $A\beta_{1-42}$	12.2 (4.5)	1.66 (6.18)	3.71 (5.85)	2.05	0.35	**55.2**	0.58
Setting slope reduction at 50%							
aMCI			2.86 (5.62)	1.43	**0.25**	50	0.32*
Low $A\beta_{1-42}$			3.71 (5.85)	1.86	**0.32**	50	0.49*

ES = effect size, calculated as (treatment-placebo change)/standard deviation of placebo change.
*Calculated by Donohue et al using two sample t-test power calculations
Data from Schneider et al [1]

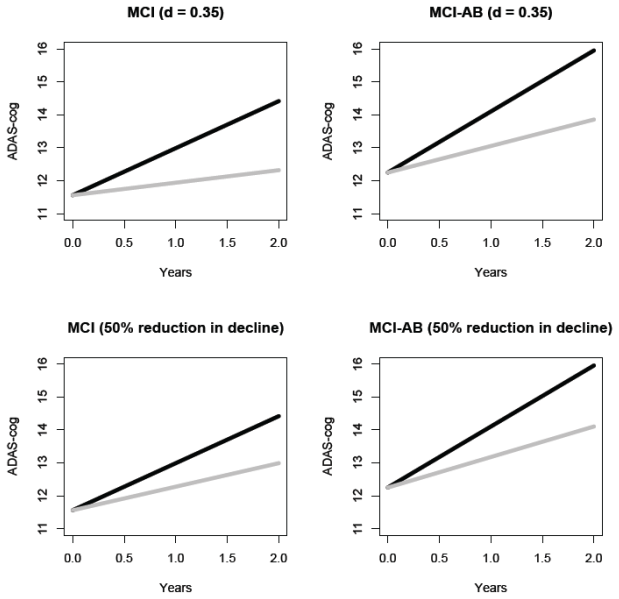


Table 2. Comparison of ADAS-cog and CDR-sb change in placebo groups over 18-months in randomized controlled clinical trials and calculations of effect sizes, based on 50% change in slope

Trial	ADAS-cog change, (SD)	50% slope reduction	Slope effect size	Power*	Sample size, 80% Power	CDRsb change, (SD)
ADCS trial 1	8.14 (8.68)	4.07	0.47	0.84	184	Not done
ADCS trial 2	6.54 (8.17)	3.27	0.40	0.71	250	2.51 (2.57)
Company A trial 1	5.49 (9.39)	2.24	0.24	0.32	686	2.55 (3.03)
Company A trial 2	4.34 (8.56)	2.17	0.25	0.35	634	2.05 (2.82)
Company B	7.36 (9.28)	3.68	0.40	0.71	250	2.50 (3.04)
Company C trial 1	6.44 (8.69)	3.22	0.37	0.64	290	2.43 (3.12)
Company C trial 2	5.85 (8.86)	2.92	0.33	0.55	364	2.74 (3.17)
Company D	9.10 (8.33)	4.55	0.55	0.93	134	2.99 (2.92)

Adapted from reference [2]
ES= within group effect size, change/SD
* Power is calculated using two sample t-test power calculations with a 20% dropout estimate

Conclusions and Discussion

‘MCI due to AD’ Results Summary

- 70-74% of aMCI patients were $A\beta_{42}$ biomarker positive; 54% were ApoE ε4 carriers
- Patients show little mean change, considerable heterogeneity in course
- Little to no difference in power across the 3 MCI inclusion criteria or ApoE carrier status, and AD
- Requiring $A\beta_{42}$ biomarker criteria or requiring ApoE ε4 carriers (or excluding them) didn’t have much of an effect on power
- Greater mean differences between placebo and treatment with biomarker criteria (for ADAS-cog), BUT there are greater increases in SDs that reduced the standardized effect sizes

Discussion

- Requiring positive biomarkers, whether Aβ or APOE, may select from the extremes of the distribution
- It is unknown if low CSF $A\beta_{42}$ patients or APOE ε4 non-carriers would be more likely to respond to an experimental drug
- The opposite could be true:
  - Targeted design trials that select only low $A\beta_{42}$ patients or ApoE ε4 non-carriers may inadvertently select those who are less likely to benefit
- Targeted clinical trials designs
  - The efficiency of a targeted design depends on the effectiveness of the drug in both the biomarker positive and negative groups, the proportion of biomarker positive patients in the sample, and the accuracy of the assay
  - The proposed treatment must be substantially more effective in the biomarker positive patients than in the excluded biomarker negative group

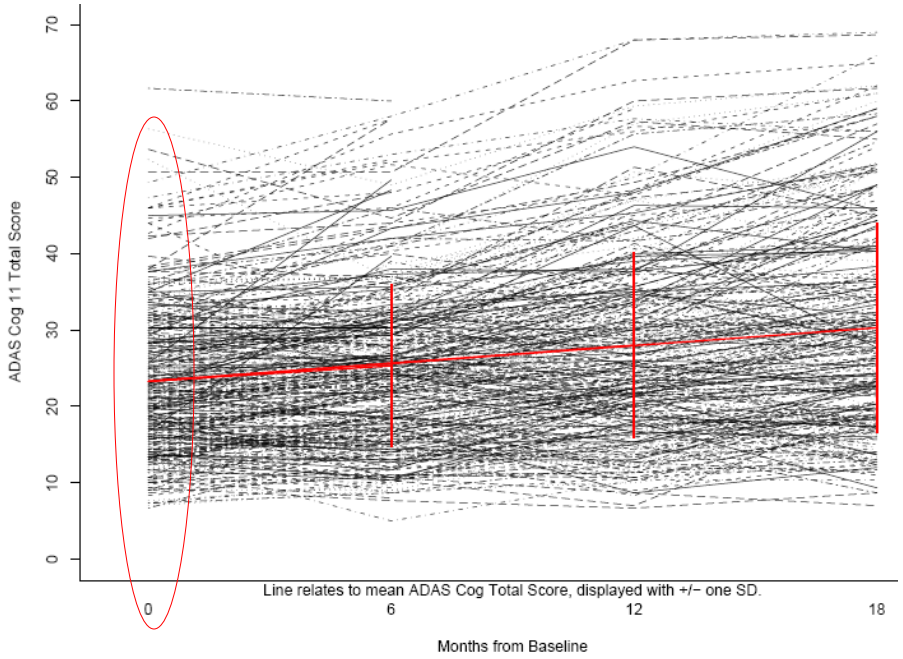
Conclusions

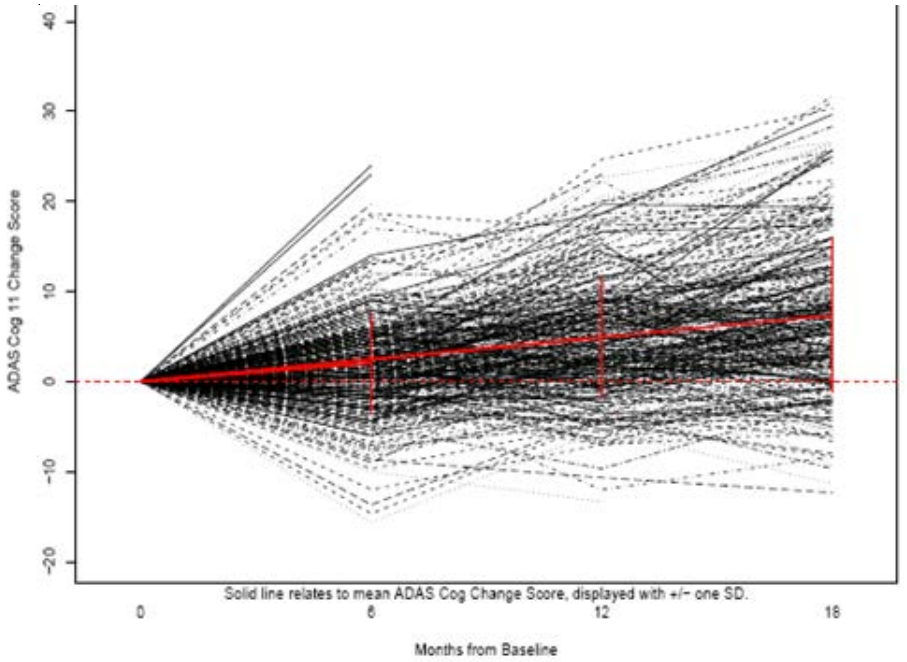
- Selecting prodromal AD patients for a clinical trial based on CSF $A\beta_{42}$ or APOE ε4 biomarker criteria will likely identify more severe patients but not enhance trials statistical power
- Absent a strong rationale to do otherwise it may be more relevant to not require current biomarkers for trials entry *in this setting* and to restrict their use as explanatory or stratification variables *when there are reasons to do so*
- Modeling, analysis, and simulations might provide a reasonable way to manage design considerations in clinical trials, better than expert opinion, conventional wisdom

THE END

The Placebo Groups of 18-month Trials

What mild to moderate AD looks like at the patient level

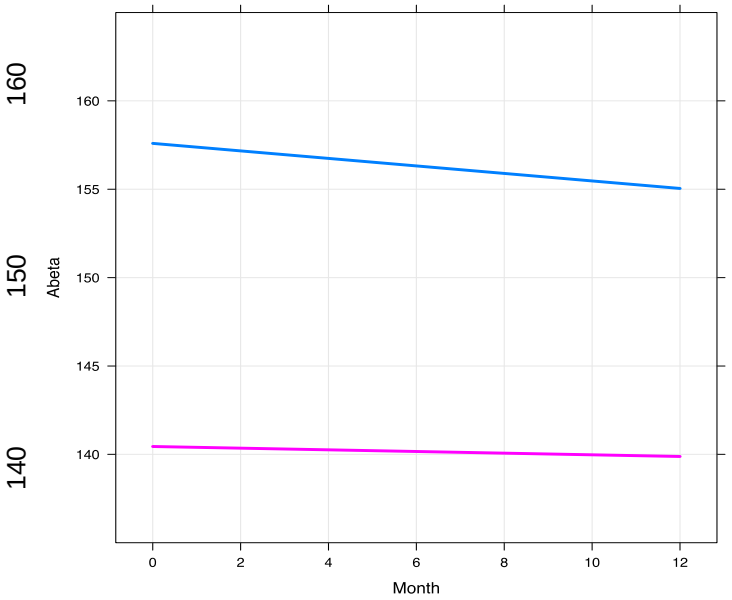


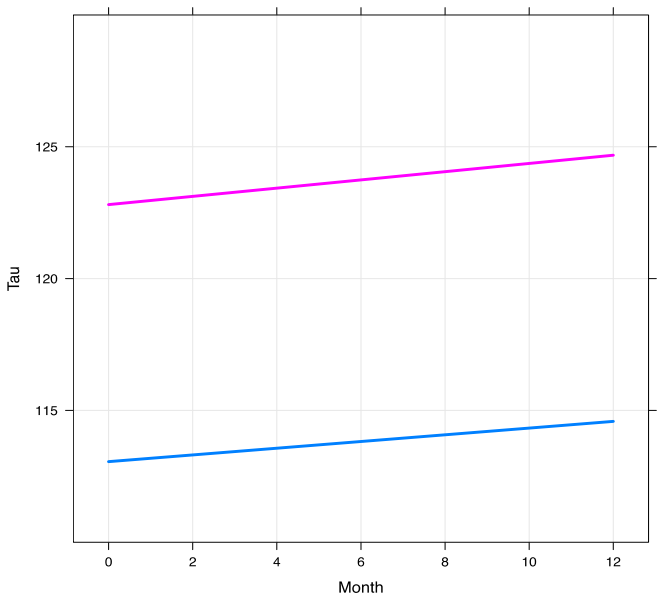


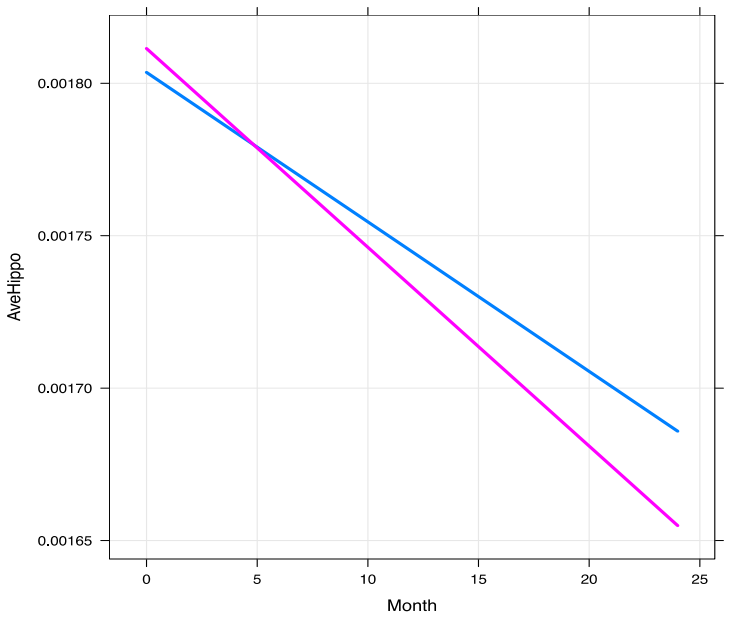
The 6-month test-retest reliability is 0.86 (the NTB is reported as 0.92)

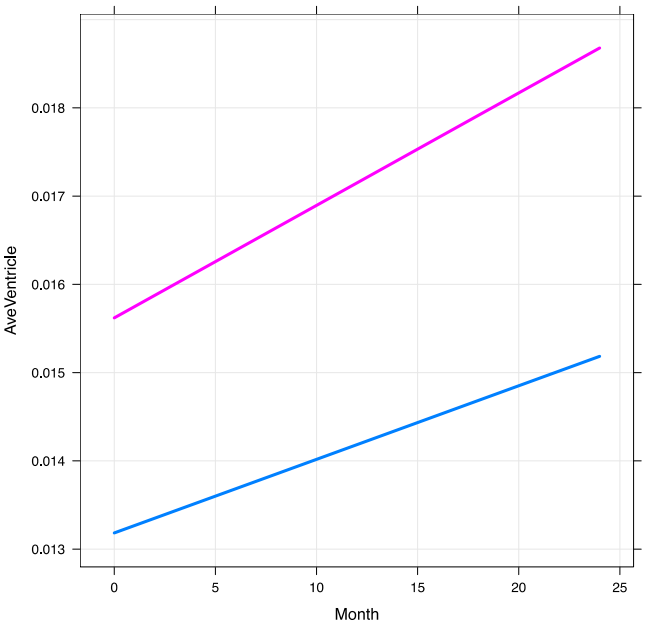
AD is relentlessly progressive, but not uniformly so.
Between 15–22% of patients show only slow or no decline
(Johnsen et al., 2003; Perrault et al., 2002; Holmes and Lovestone, 2003)

Biomarker Change (Aβ,Tau, HC, Ventricles)









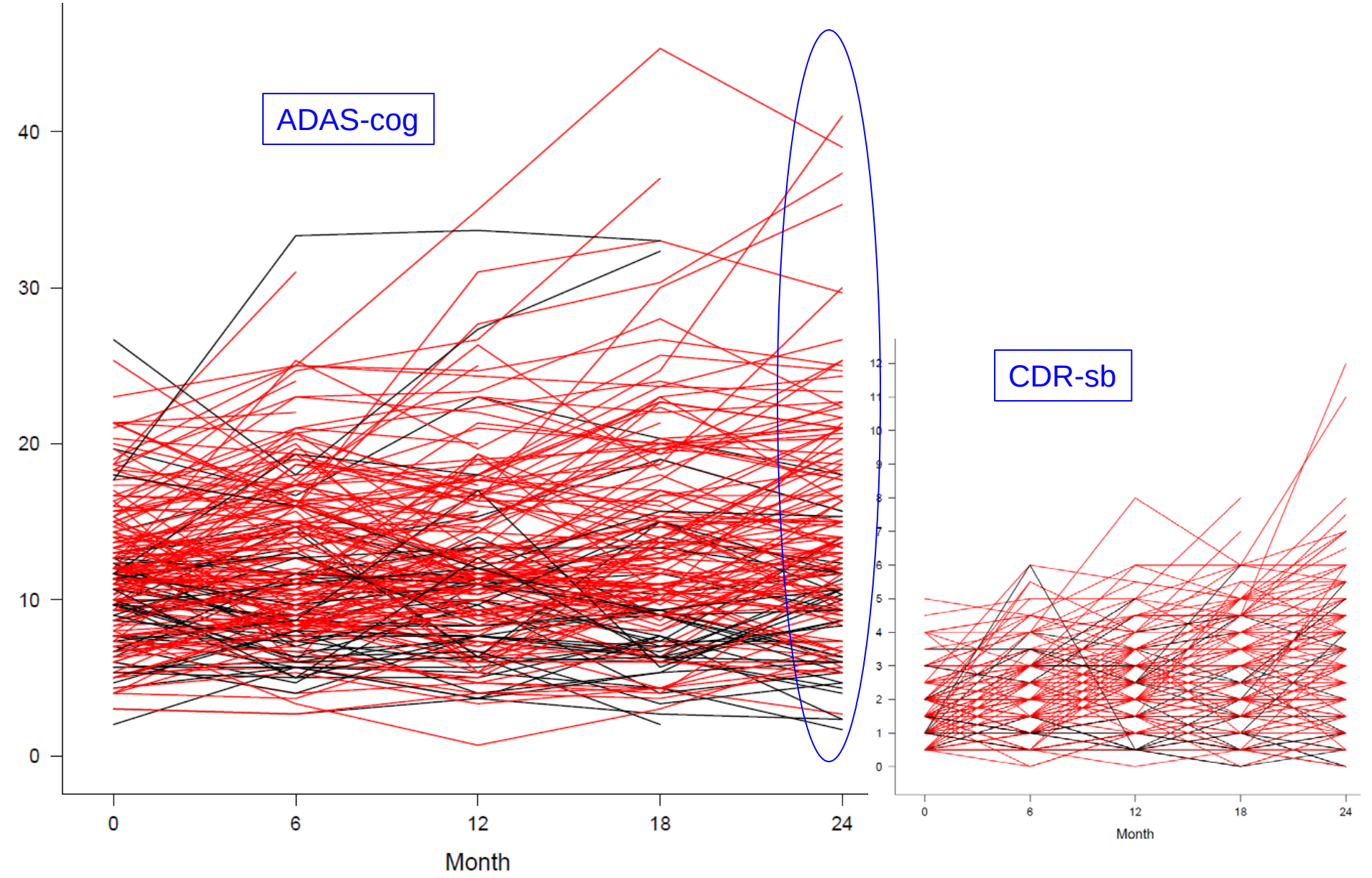
Overarching Context

- Considerable obstacles to translating pre-clinical research to clinical
- Urgency to do more trials with fewer (or more?) patients; to “get a signal” earlier….there are too many drugs and no validated targets
- Clinical trials often don’t turn out as planned, often underpowered to test the hypothesis
  - We blame the statistics, models, sites, placebos, cholinesterase inhibitors, outcomes measures
- We then try to improve the next trial by tweaking, e.g., inclusion criteria, outcomes, follow-ups, and biomarkers
- We believe that this will “reduce heterogeneity”
- These are complex problems and we stand to be disappointed if we rely on simple solutions

Outline

- Background: post hoc analyses of AD trials based on APOE 4 carriage have provided interesting and contradictory results.
  - some results might be due to play-of-chance in underpowered analyses,
  - other results may be due to actual interaction of the drug with the subgroup.
  - APOE 4 is the strongest risk factor for AD and associated with age of onset of AD it has received particular attention for stratified medicine approaches.
- Review trials that published outcomes based on APOE carriage
- Present trials simulations derived from ADNI on Abeta carriage
- Present trials simulations derived from ADCS trials and ADNI that empirically test the efficiency of developing drugs based on trials scenarios of APOE carriage
  - specifically, what might be gained by certain stratified medicine assumptions.
- Conclusions:
  - Previous trials using targeted designs in AD were either misleading or didn't achieve intended purpose.
  - Using an ApoE or Abeta biomarker doesn't affect trials much at all
  - Except hypothetically if you just used E2 carriers (< 10% of AD) then there is little change
- Discuss: the conditions under which targeted designs could work and suggestions on making focused trials better

MCI CSF $A\beta_{42}$ positives (--) and negatives (--)



ELSEVIER

Alzheimer's & Dementia 6 (2010) 367–377

Alzheimer's & Dementia

Featured Articles

Requiring an amyloid-β_{1-42} biomarker for prodromal Alzheimer's disease or mild cognitive impairment does not lead to more efficient clinical trials

Lon S. Schneider[a,*], Richard E. Kennedy[b], Gary R. Cutter[b]; and the Alzheimer's Disease Neuroimaging Initiative

[a]*Departments of Psychiatry and Neurology, Keck School of Medicine, and Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA*

[b]*School of Public Health, University of Alabama, Birmingham, AL, USA*

ADAS-cog

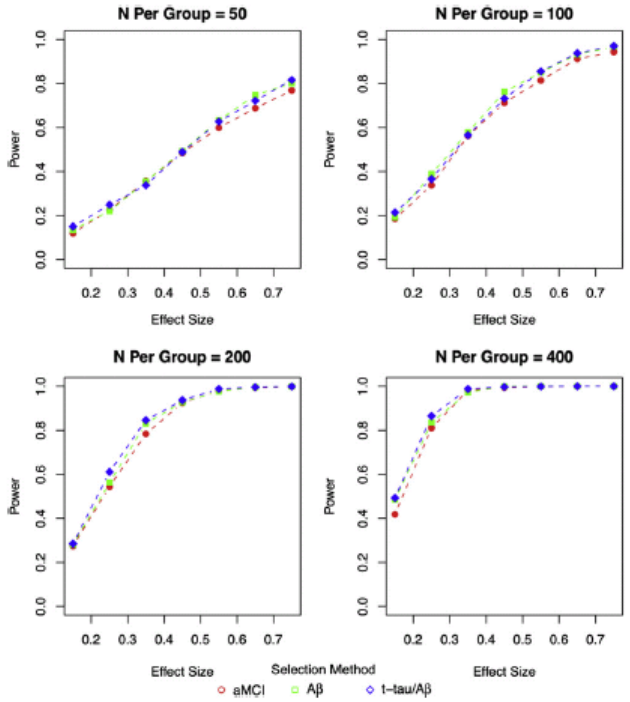


Fig. 1. Power for Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog) outcomes in 24-month-long trials with 20% dropouts. Power calculations for the ADAS-cog in patients with amnestic mild cognitive impairment (aMCI), with aMCI plus low cerebrospinal fluid (CSF) $A\beta_{1-42}$, and aMCI plus high CSF t-tau/$A\beta_{1-42}$ by effect size and sample size. Addition of a biomarker selection criterion did not appreciably increase power under any of the scenarios. Simulation parameters included $\alpha = 0.05$, χ^2 random errors, and 20% dropouts with mixed model analysis for participants with missing data.

CDR-sb

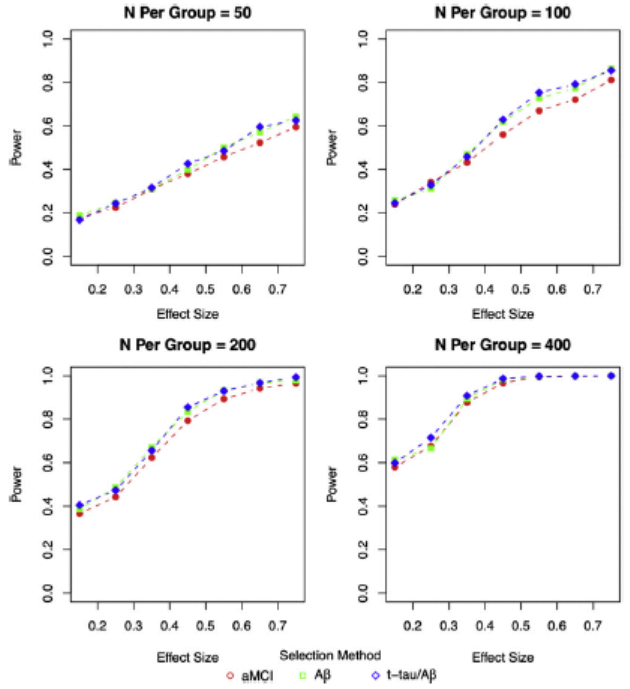
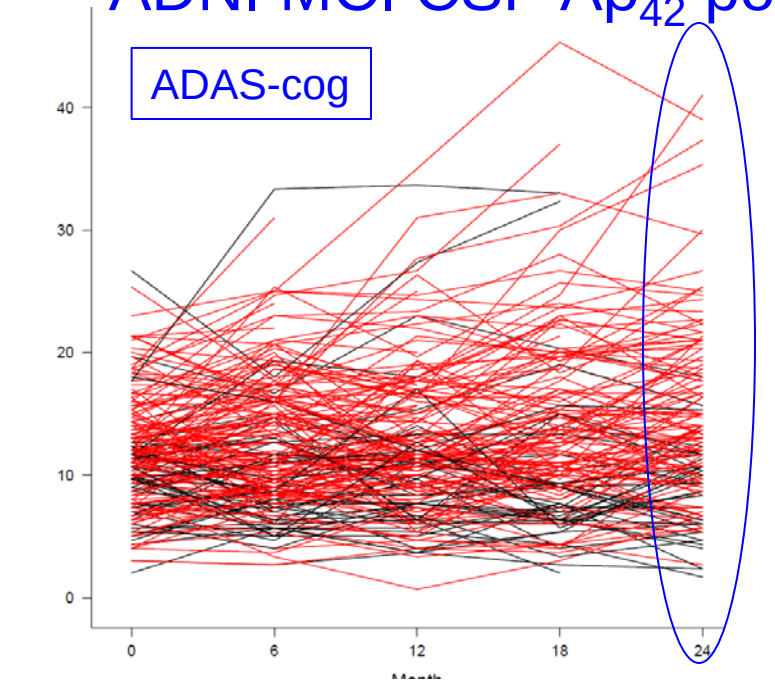
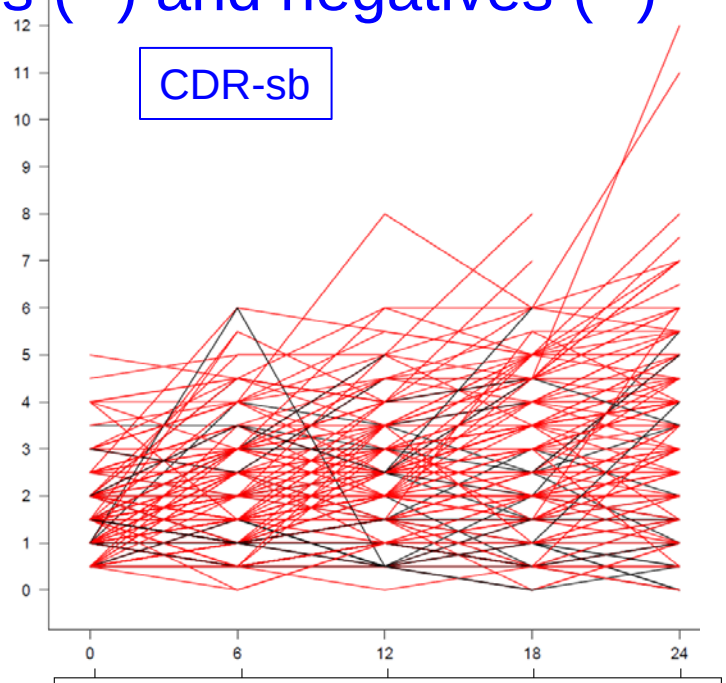
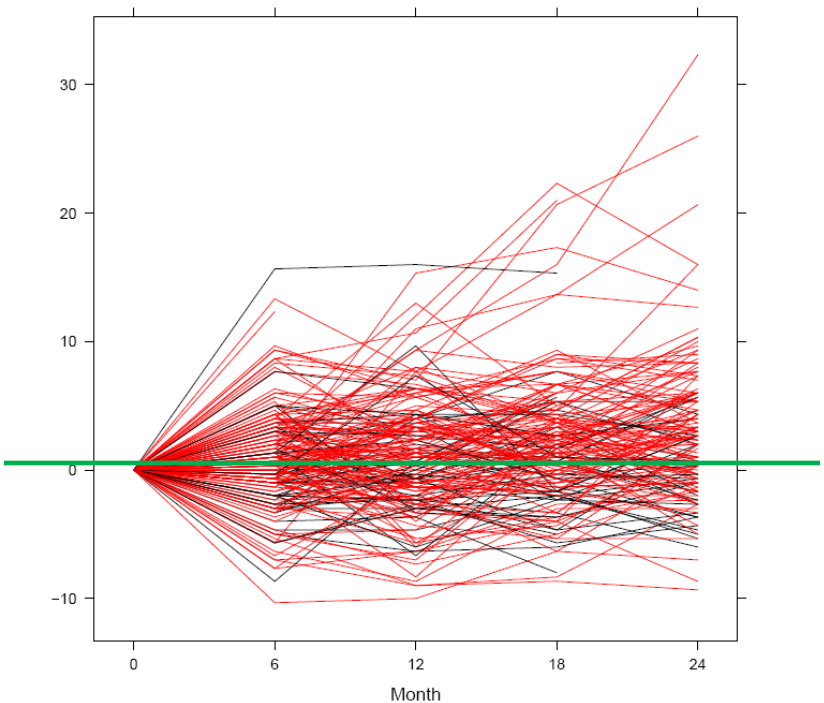


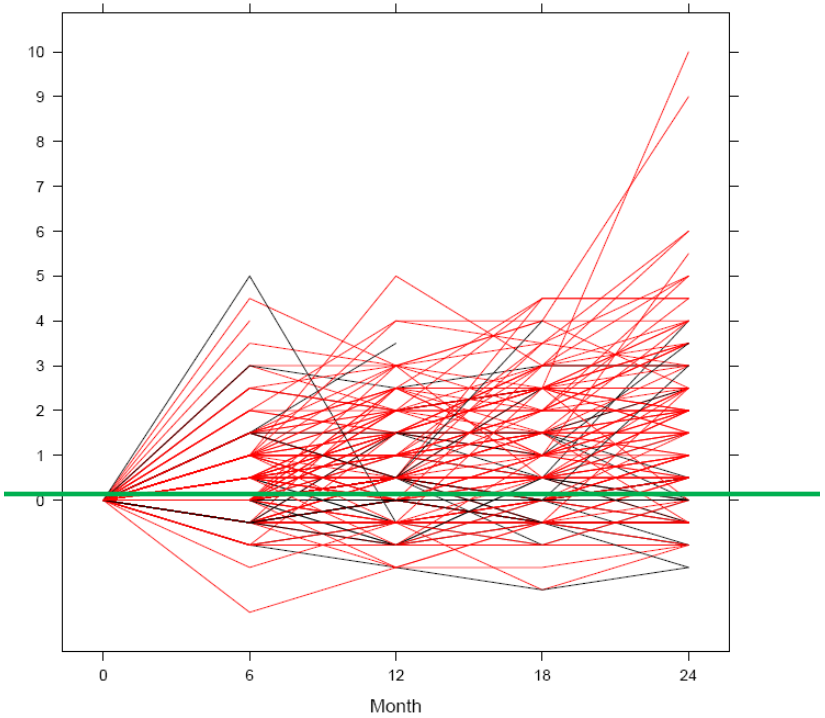
Fig. 4. Power for CDR-sb outcomes in 24-month-long trials with 40% dropouts. Power calculations for the CDR-sb in patients with aMCI, aMCI plus low CSF $A\beta_{1-42}$, and aMCI plus high CSF t-tau/ $A\beta_{1-42}$ by effect size and sample size. Addition of biomarkers resulted in only small increases in statistical power, mainly with smaller sample sizes and large effect sizes. Simulation parameters included $\alpha = 0.05$ with 24-month-long trials, χ^2 random errors, and 40% dropouts with mixed model analysis for participants with missing data.

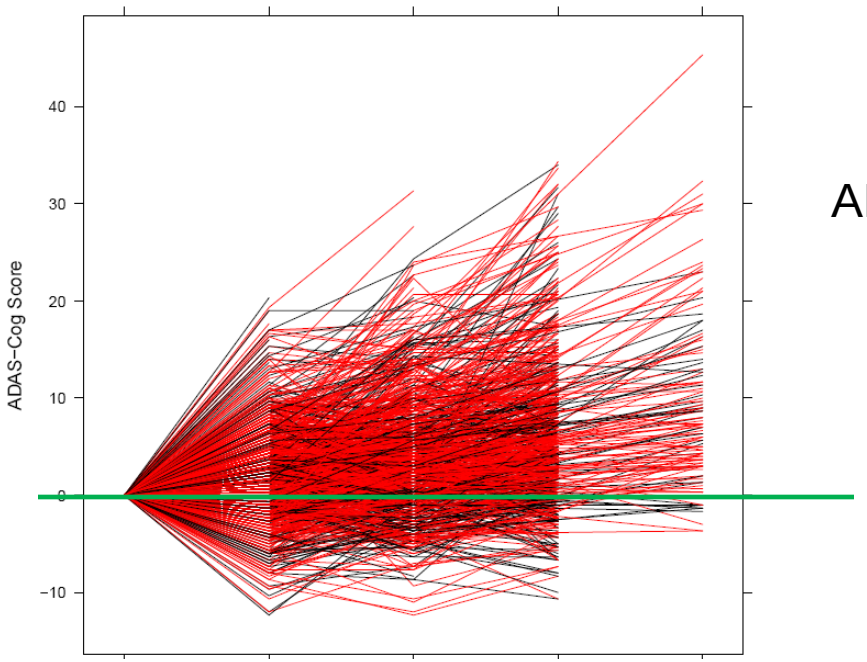
ADNI MCI CSF $A\beta_{42}$ positives (--) and negatives (--)



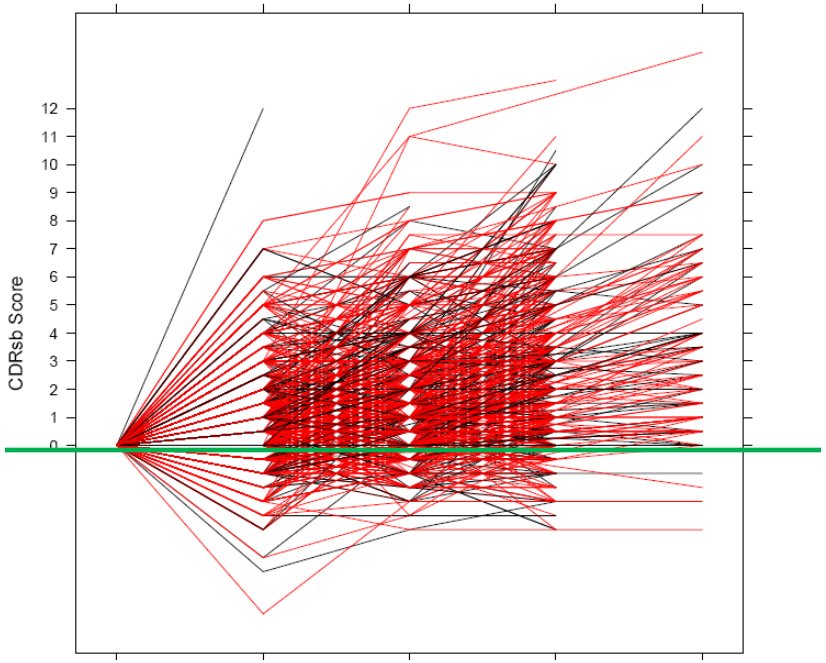


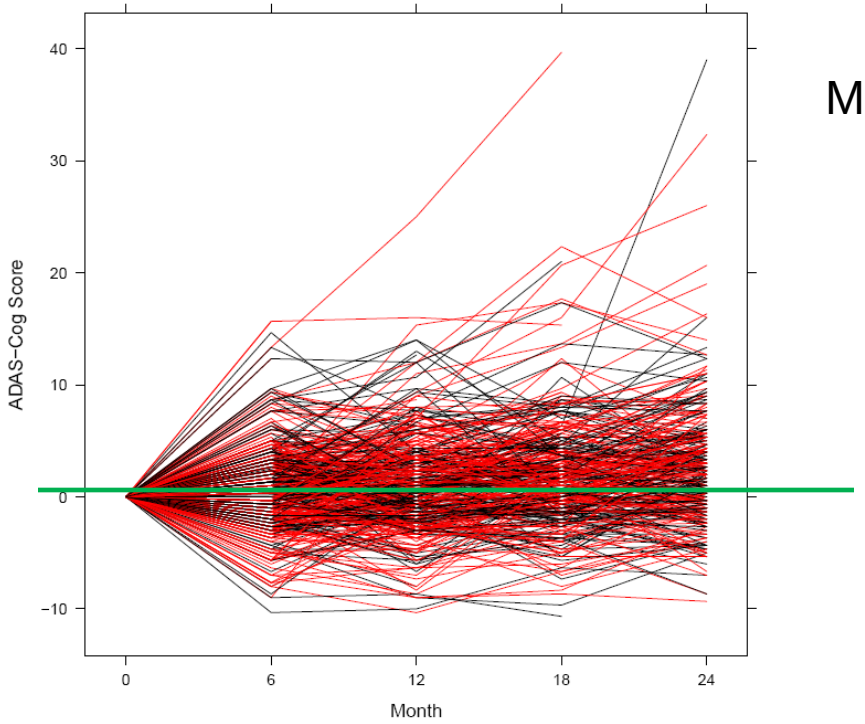




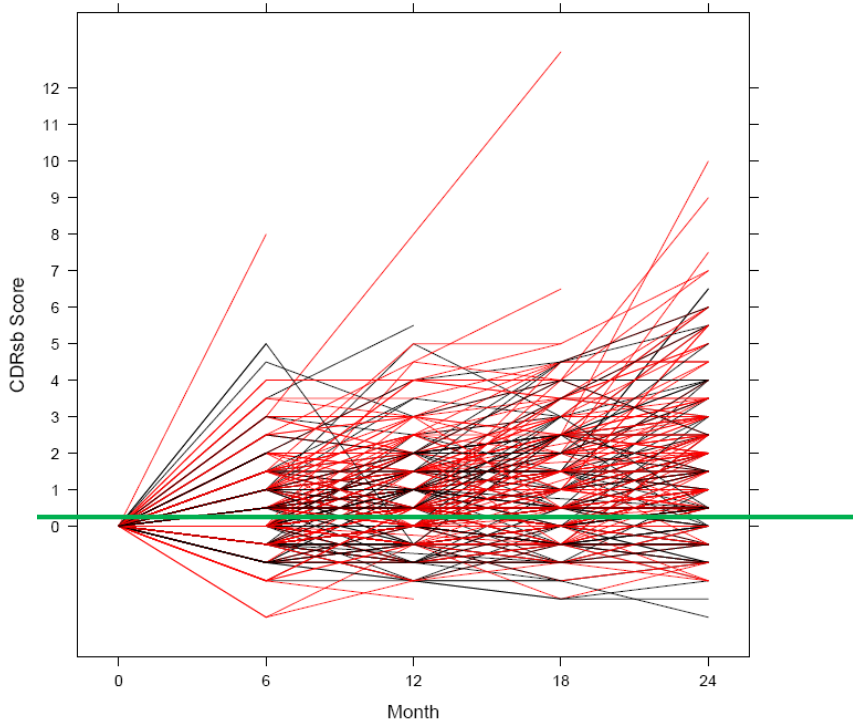

ADAS−Cog Score
40
30
20
10
0
−10

AD

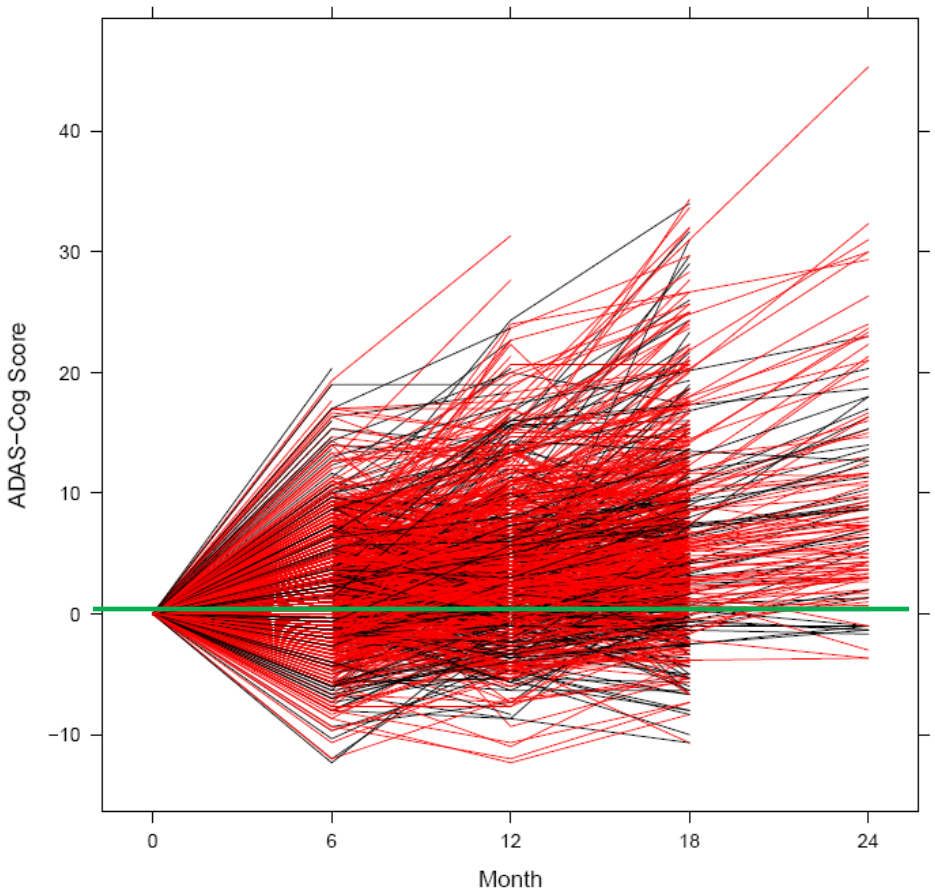

CDRsb Score
12
11
10
9
8
7
6
5
4
3
2
1
0

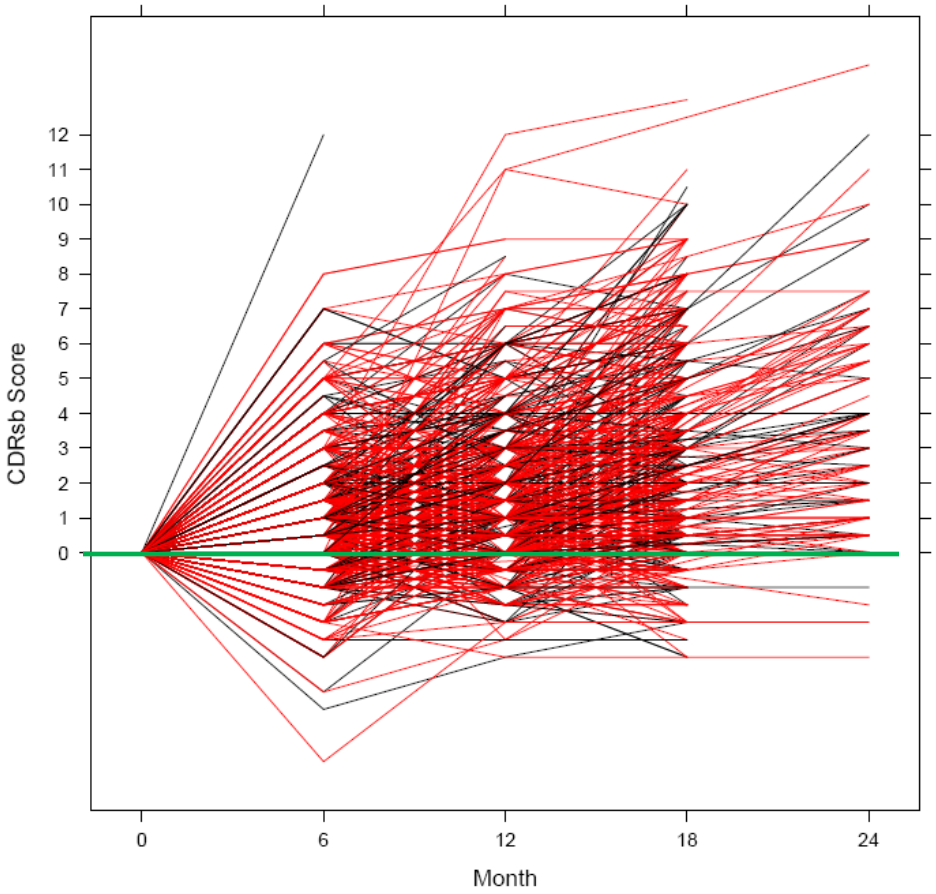

ADAS−Cog Score
40
30
20
10
0
−10
0
6
12
18
24
Month

MCI


CDRsb Score
12
11
10
9
8
7
6
5
4
3
2
1
0
0
6
12
18
24
Month

ADAS-cog and CDR-sb Change in AD





AD studies: ADNI, DHA, ES, HC, and PR, for ADAS-cog N =1042 at baseline, 906 at month 6, 816 at month 12, 688 at month 18, and 133 at month 24; for CDRsb, there are 1057 at baseline, 970 at month 6, 899 at month 12, 749 at month 18, and 133 at month 24

Limitations

- Results depend on the extent that ADNI represents clinical trials sample
- Substantial majority of MCI and AD patients already had low $A\beta_{42}$ and high t-tau/ $A\beta_{42}$ and are APOE ε4 carriers
- Precision, timing and standardization of the assay?
- Using other cutoffs for biomarkers, other selection criteria may give different results and provoke different considerations