Randomized controlled trials in mild cognitive impairment

Sources of variability

OPEN

Ronald C. Petersen, PhD, MD
Ronald G. Thomas, PhD
Paul S. Aisen, MD
Richard C. Mohs, PhD
Maria C. Carrillo, PhD
Marilyn S. Albert, PhD
For the Alzheimer's Disease Neuroimaging Initiative (ADNI) and Foundation for NIH (FNIH) Biomarkers Consortium AD MCI Placebo Data Analysis Project Team

Correspondence to Dr. Petersen: peter8@mayo.edu

ABSTRACT

Objective: To examine the variability in performance among placebo groups in randomized controlled trials for mild cognitive impairment (MCI).

Methods: Placebo group data were obtained from 2 National Institute on Aging (NIA) MCI randomized controlled trials, the Alzheimer's Disease Cooperative Study (ADCS) MCI trial and the Alzheimer's Disease Neuroimaging Initiative (ADNI), which is a simulated clinical trial, in addition to industry-sponsored clinical trials involving rivastigmine, galantamine, rofecoxib, and donepezil. The data were collated for common measurement instruments. The performance of the placebo participants from these studies was tracked on the Alzheimer's Disease Assessment Scale–cognitive subscale, Mini-Mental State Examination, and Clinical Dementia Rating–sum of boxes, and for progression on these measures to prespecified clinical study endpoints. *APOE* status, where available, was also analyzed for its effects.

Results: The progression to clinical endpoints varied a great deal among the trials. The expected performances were seen for the participants in the 2 NIA trials, ADCS and ADNI, with generally worsening of performance over time; however, the industry-sponsored trials largely showed stable or improved performance in their placebo participants. *APOE*4 carrier status influenced results in an expected fashion on the study outcomes, including rates of progression and cognitive subscales.

Conclusions: In spite of apparently similar criteria for MCI being adopted by the 7 studies, the implementation of the criteria varied a great deal. Several explanations including instruments used to characterize participants and variability among study populations contributed to the findings.

Neurology® 2017;88:1751–1758

GLOSSARY

AD = Alzheimer disease; **ADAS-Cog** = Alzheimer's Disease Assessment Scale–cognitive subscale; **ADCS** = Alzheimer's Disease Cooperative Study; **ADNI** = Alzheimer's Disease Neuroimaging Initiative; **aMCI** = amnestic mild cognitive impairment; **CDR** = Clinical Dementia Rating; **FDA** = Food and Drug Administration; **InDDEx** = Investigation Into Delay to Diagnosis of Alzheimer's Disease With Exelon; **MCI** = mild cognitive impairment; **MMSE** = Mini-Mental State Examination; **MRK** = Merck rofecoxib trial; **SB** = sum of boxes.

Randomized controlled trials are necessary for the development of therapeutics in medicine. However, they are very expensive and, if not designed and conducted with great attention to detail, can yield uninterpretable results due to methodologic issues. This has been of special concern in the field of Alzheimer disease (AD), since over the last decade a substantial number of compounds have been tested for efficacy with no new therapeutics being approved by the US Food and Drug Administration (FDA) in over 10 years.[1] A relevant factor contributing to these

Supplemental data at Neurology.org

From the Department of Neurology (R.C.P.), Mayo Clinic College of Medicine, Rochester, MN; Department of Neurosciences (R.G.T.), University of California, San Diego, La Jolla; University of Southern California (P.S.A.), Los Angeles; Eli Lilly and Company (R.C.M.), Indianapolis, IN; Division of Medical & Scientific Relations (M.C.C.), Alzheimer's Association, Chicago, IL; and Johns Hopkins School of Medicine (M.S.A.), Baltimore, MD. R.C.M. is currently with Global Alzheimer's Platform Foundation, Washington, DC.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found in the coinvestigators list at Neurology.org.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article. The Article Processing Charge was funded by the authors.

failures may be trial designs that do not test a well-specified hypothesis or that are not implemented sufficiently precisely to test the hypothesis.

Recently, the National Institute on Aging and the Alzheimer's Association proposed criteria for the entire AD spectrum from preclinical phases through dementia.[2] Intermediary in that process is the construct of mild cognitive impairment (MCI), which is typically characterized as a clinical state between normal cognition of aging and very early dementia.[3–5]

Over the last decade, there have been several clinical trials involving participants with MCI, primarily amnestic MCI (aMCI), but all have failed to show efficacy.[6–10] A key issue in these studies pertains to participant selection, since the variability in the performance of the placebo participants was considerable.[11] MCI is a clinical syndrome like dementia and likely represents a continuum of symptoms, and with biomarkers, specificity for underlying etiology of MCI due to AD is improving. With this motivation, the Foundation for the National Institutes of Health Biomarkers Consortium in conjunction with the Alzheimer's Association launched a project to explore 7 MCI clinical trials that had been conducted in an effort to examine the extent to which placebo-treated patients from these studies behaved in a fashion commensurate with previous longitudinal studies of MCI. Specifically, the study addressed the following questions:

1. Were the participants equivalent at baseline in all studies?
2. Was the measured change in cognitive function over the follow-up period significantly different across the different studies?
3. Were the participants significantly different at the time of exit, i.e., progression to the endpoint in the study, considering that some studies used clinical progression to dementia while others used the Clinical Dementia Rating (CDR)?
4. How did the participants in the various clinical trials perform in the initial 24 months of the studies since that time frame allows us to put 6 of the studies on the same scale?
5. Were there demographic (age, sex, *APOE* status) or baseline severity measures (Mini-Mental State Examination [MMSE]) that can account for any differences in progression or clinical status across studies?

Table Design features of randomized controlled trials in mild cognitive impairment

Study	Agent	No. of Partic	Age, y	Sex, % F	*APOE4*, %	Entry criteria	Exit criteria	ADAS-Cog, mean (SD), range 0–66	MMSE, mean (SD), range 1–30	CDR-SB, mean (SD), range 0–18	Duration, mo	No. of sites	Outcome
ADCS	Donepezil; vitamin E; NCT 00000173	259	72.9	47	53	CDR, LM, MMSE	Clinical dementia	11.03 (4.2)	27.3 (1.8)	1.9 (0.8)	36	59	Neg[a]
ADNI	None	398	74.7	35	53	CDR	Clinical dementia	11.5	27.0	1.6	12	58	NA
InDDEx	NCT 00106899; rivastigmine	510	70.6	51	46[b]	LM, MMSE, CDR	Clinical dementia due to AD	10.2 (4.4)	26.9 (1.8)	1.4 (0.9)	48	65	Neg
Gal-11	NCT 00000174; galantamine; NCT 00236431	496	70.1	52	43[c]	NYUPDR, CDR, NYUPDR	CDR ≥ 1	Median = 16.0 (4.9); range 2–61		1.6 (0.04)	24	177	Neg
Gal-18	Galantamine NCT 00236574	526	70.9	59	43[c]	CDR, NYUPDR	CDR ≥ 1	Median = 18.0; range 2–63		1.7 (0.04)	24	177	Neg
MRK	Rofecoxib	732	74.8	31	36[d]	MMSE, CDR, BLESSED, AVLT	CDR ≥ 1, CONFIRM, 2 mo	9.8 (3.8)	27.3 (1.7)	1.4 (0.8)	48	46	Neg
Eisai	Donepezil; NCT 00293176	387	69.8	43	42[e]	CDR, LM, MMSE	Clinical dementia	18.2 (7.0)		1.5 (0.9)	12	74	Neg

Abbreviations: AD = Alzheimer disease; ADAS-Cog = Alzheimer's Disease Assessment Scale–cognitive subscale; ADNI = Alzheimer's Disease Neuroimaging Initiative; AVLT = Auditory Verbal Learning Test; BLESSED = Blessed Memory Information Test; CDR = Clinical Dementia Rating; InDDEx = Investigation Into Delay to Diagnosis of Alzheimer's Disease With Exelon; LM = Wechsler Memory Scale-Revised Logical Memory Subtest delayed recall; MMSE = Mini-Mental State Examination; NYUPDR = New York University Paragraph Delayed Recall; Partic = participants; SB = sum of boxes.

[a] Positive donepezil effect for 12 and 24 months in *APOE4* carriers.

[b] On 248/510 participants who had DNA drawn.

[c] Combined Gal-11 and Gal-18 of these consented to DNA.

[d] On 669/732 participants who had DNA drawn.

[e] On 541/821 participants who had DNA drawn.

METHODS Datasets from the 7 large studies with sample sizes greater than 250 participants (2 by the same sponsor using essentially identical methodologies) were collected and analyzed. The total number of placebo participants analyzed was 3,308. The duration of the trials ranged from 2 to 4 years.

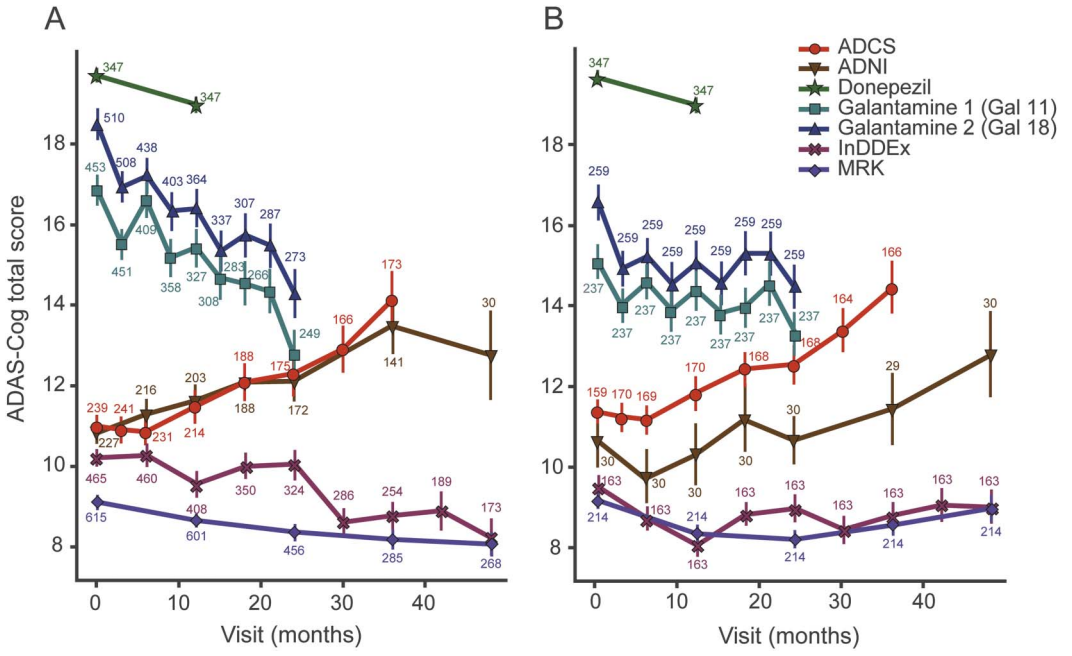
The essential features of the 7 trials are shown in the table. Using the Alzheimer's Disease Assessment Scale–cognitive subscale (ADAS-Cog) as an example of how such common metrics were developed, the first step was to determine whether the individual ADAS-Cog items were collected, or only the total score was captured.[12] If total scores were available, the next question was to determine if the total score was calculated using the ADAS-Cog with 11, 12, or 14 items. To allow the assessment of population demographic differences, key descriptive variables were located and extracted from each trial at the participant level: age, sex, race, education, and baseline MMSE (if available), investigator, and country.[13] Joining of datasets necessitated locating unique participant identifiers, and also indicators identifying study populations, e.g., intention to treat, per protocol, or safety. Longitudinal assessment of response measures required determination of the visit schedule and translation of visit codes to common time units. Other challenges included *APOE* status data (there was no standard method of archiving and determining the set of codes used to indicate missing data items).

Study designs. The characteristics of the 7 studies included in these analyses are outlined in the table. In general, all used the same criteria for aMCI, but varied in their implementation procedures.[3] Most used a CDR of 0.5 and a MMSE score in the range of ≥24 at entry.[12,13] The specific criteria for the memory impairment varied among the studies and included an education-adjusted delayed recall score on Logical Memory from the Wechsler Memory Scale–Revised, delayed recall on the New York University paragraph recall test, or delayed recall on the Auditory Verbal Learning Test.[14–16] The exit criteria included the clinical diagnosis of dementia or progression to a CDR of ≥1. The duration of the studies ranged from 12 (donepezil) to 48 months (rivastigmine). The 2 NIH trials, the Alzheimer's Disease Cooperative Study (ADCS) and Alzheimer's Disease Neuroimaging Initiative (ADNI), were identical with respect to criteria, while the Eisai (donepezil) trial, the 2 Janssen galantamine trials (Gal-11 and Gal-18), the Novartis rivastigmine trial (Investigation Into Delay to Diagnosis of Alzheimer's Disease With Exelon [InDDEx]), and the Merck rofecoxib trial (MRK) had different entry and exit criteria.

Statistical analyses. The mean ADAS-Cog values were compared using a 1-way analysis of variance at baseline and again at 24 months. Following a significant omnibus test, p values for pairwise comparisons were calculated and adjusted using the Holm procedure. The R statistical computing environment, version 3.1, was used for both statistical analysis and data management. Only data from the participants in the ADNI trial who were not using cholinesterase inhibitors or memantine were included so their data corresponded to the placebo participants from the other trials.

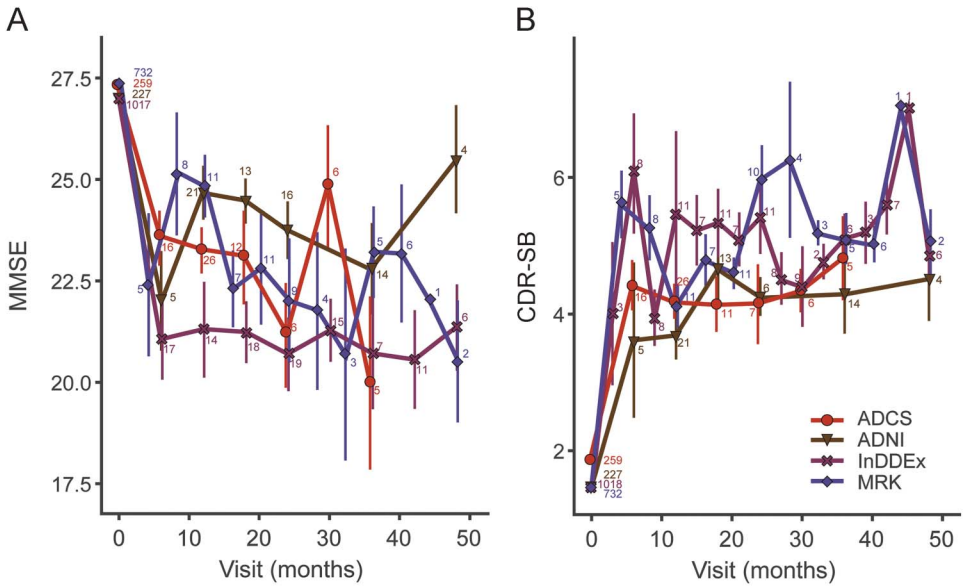
RESULTS **Overall strategy.** The data from the 7 clinical trials were combined to evaluate the overall performance using instruments in common among the various studies. To allow us to compare the trials, several common metrics were evaluated, including the ADAS-Cog, the MMSE, and the CDR–sum of boxes (SB). In addition, word recall scores from the ADAS-Cog were used to assess memory function among the trials.

Figure 1 Alzheimer's Disease Assessment Scale–cognitive subscale (ADAS-Cog) performance



Mean performance with 95% confidence intervals on the ADAS-Cog for the mild cognitive impairment clinical trials for all participants (A) and for participants who completed the trial (B). ADCS = Alzheimer's Disease Cooperative Study; ADNI = Alzheimer's Disease Neuroimaging Initiative; InDDEx = Investigation Into Delay to Diagnosis of Alzheimer's Disease With Exelon; MRK = Merck rofecoxib trial.

Figure 2 Mini-Mental State Examination (MMSE) and Clinical Dementia Rating–sum of boxes (CDR-SB) performance

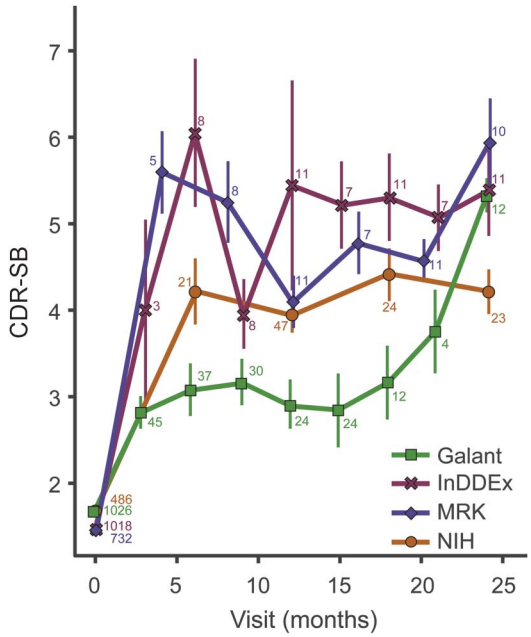


Mean performance with 95% confidence intervals for 4 trials on the MMSE (A) and on the CDR-SB (B). ADCS = Alzheimer's Disease Cooperative Study; ADNI = Alzheimer's Disease Neuroimaging Initiative; InDDEx = Investigation Into Delay to Diagnosis of Alzheimer's Disease With Exelon; MRK = Merck rofecoxib trial.

Demographics. The table demonstrates the demographic and clinical features of participants in all 7 trials at baseline. As is apparent, the general characteristics were reasonably similar, although there were slight differences in the cognitive status of the participants in the various trials at baseline.

Overall comparisons. Figure 1 shows the performance features of the ADAS-Cog among the 7 clinical trials. Those participants in the ADNI trial who were on cholinesterase inhibitors were removed from the analyses. Panel A demonstrates performance of all participants, and panel B the completers' analysis. As is evident, the ADAS-Cog appeared to show cognitive worsening as expected in the ADCS and ADNI trials (NIH trials); however, the participants in the Gal-11 and Gal-18 trials appeared to improve. Similarly, in the InDDEx, donepezil, and MRK trials, the participants seemed to improve or remain stable. These conclusions were attenuated in the completers' analysis, but the same trends were seen. When the NIH trials were combined and compared to the 5 combined industry trials, there was strong statistical difference in slopes, with the ADAS-Cog performance on the NIH trials worsening with the industry trials remaining stable or improving for all participants ($p < 0.01$) as well as for the completers ($p < 0.01$). Similar analyses for the MMSE total participants and completers and CDR-SB total participants and completers (data not shown) corroborate the ADAS-Cog findings. That is, participants appeared to worsen in a predictable fashion for the ADCS-MCI and ADNI trials but remained stable or improved slightly in the other 5 trials.

Figure 3 Clinical Dementia Rating–sum of boxes (CDR-SB) over 24 months



Mean CDR-SB with 95% confidence intervals for 4 trials truncated at 24 months. InDDEx = Investigation Into Delay to Diagnosis of Alzheimer's Disease With Exelon; MRK = Merck rofecoxib trial.

Figure 2A demonstrates the MMSE at entry and at point of progression to endpoint in the individual trials. Due to attrition over the trials, some data points represent a small number of participants. While the values at entry were similar, it appears that the participants in the InDDEx trial were a bit less impaired than participants in the ADCS and ADNI trials ($p < 0.01$). The MRK participants were similar to those in the NIH trials (ADCS and ADNI).

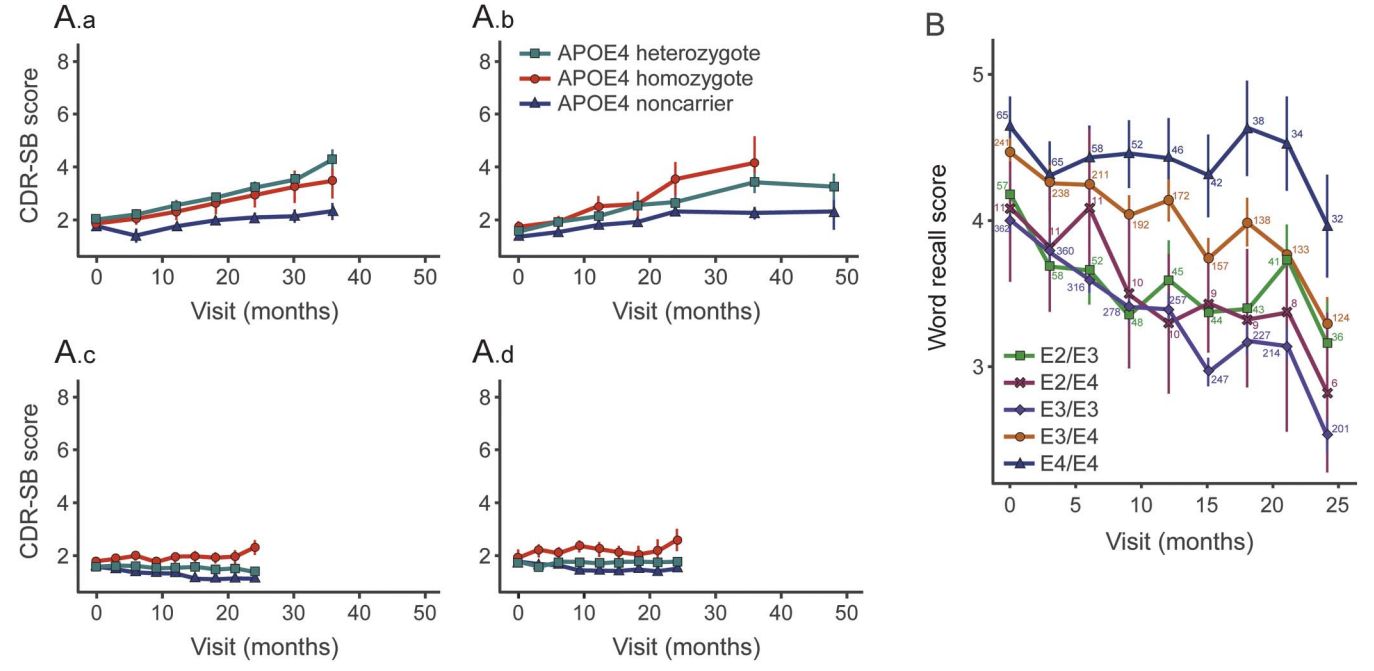
In a similar fashion, figure 2B indicates that, with respect to the CDR-SB, the InDDEx and MRK participants were less impaired at entry than the participants in the ADCS and ADNI trials. However, at the time of reaching the endpoint in the respective studies, the InDDEx and MRK participants appeared to be more impaired than the ADCS and ADNI participants ($p < 0.01$).

In an attempt to compare studies in a more standardized fashion, the 2 NIH studies, ADCS and ADNI, were combined, and the Gal-11 and Gal-18 studies were combined. These data were then plotted relative to the InDDEx and MRK trials with regard to CDR-SB in figure 3. The comparisons were also truncated at 24 months since all studies continued to at least that point. As is apparent from figure 3, the InDDEx and MRK trial participants were slightly less impaired than the NIH participants at entry, but the combined Gal, InDDEx, and MRK participants were more impaired than the NIH participants at 24 months ($p < 0.001$). While the Gal-combined participants were somewhat less impaired throughout the study until the 24-month point, the InDDEx and MRK studies' participants were more impaired throughout most of the study, up to 24 months ($p < 0.01$).

As is shown in figure 4A, the *APOE4* carrier status clearly had an effect on performance, with greater impairment on the CDR-SB among *APOE4* homozygotes across all trials. Similarly, in figure 4B, the delayed recall scores on the ADAS-Cog in the Gal-11 and Gal-18 trials also showed a consistent relationship, with poorest performance being observed in the participants who were *APOE4* homozygotes, followed by those who were heterozygotes, and with *APOE3/3* participants performing the best. As is shown in the table, the percentage of participants who were *APOE4* carriers varied a great deal by trial.

DISCUSSION In spite of the fact that all 7 trials used ostensibly the same MCI criteria for amnestic MCI participants, the performance of the participants in the studies varied considerably. Most of the trials included participants with multidomain aMCI and the degree of impairment in the nonmemory domains varied. In addition, there may have been differences between academic and commercial sites, but this factor was not addressed. The data suggested that the studies were not equivalent with regard to the nature

Figure 4 **Influence of *APOE***



Mean performance with 95% confidence intervals on Clinical Dementia Rating– sum of boxes (CDR-SB) by *APOE* carrier status (A.a, Alzheimer's Disease Cooperative Study mild cognitive impairment (MCI); A.b, Alzheimer's Disease Neuroimaging Initiative MCI; A.c, Gal-11; A.d, Gal-18) and for mean word recall with 95% confidence intervals from the Alzheimer's Disease Assessment Scale by *APOE* carrier status (B).

of the participants enrolled. In general, it appears that the ADCS and ADNI participants were very similar and were more impaired at entry than participants in Gal-11, Gal-18, InDDEx, or MRK trials. In particular, the InDDEx and MRK participants performed better at entry compared to the Gal-11 or Gal-18 participants, and performed more poorly at the time of reaching their endpoints. These conclusions can be derived from evaluations of the ADAS-Cog, MMSE, and CDR-SB data and suggest that in these trials, the participants may have been less impaired at entry but the exit criteria may have been more strict, resulting in the fact that the participants were categorized as MCI for a longer period of time yielding lower progression rates. There may have been an element of a placebo effect present as well. Therefore, while the studies appeared to be designed in a similar fashion, subtle entry, exit, and procedural differences led to vastly different outcomes for the participants.

In addition, there may have been subtle differences in the conduct of the trials that led to increased variability in performance. For example, in the InDDEx trial, the investigators made the judgment of progression to the endpoint, clinical dementia, without access to the neuropsychological test data. The clinicians were shown the ADAS-Cog and behavioral measures but not the specific components of the neuropsychology battery. In addition, the site clinicians submitted all of the data to a central review committee that required significant cognitive changes on several metrics prior to declaring that the participant had reached the point of dementia, again without access to neuropsychology test scores. It is possible that the participants in the InDDEx trial, for example, were therefore more impaired at exit than in the other trials, thus having a different threshold for reaching the endpoint of the study. Ultimately, this resulted in a lower progression rate and led to the extension of the study to 3 and 4 years in duration.

The role of *APOE4* carrier status cannot be overemphasized. As is shown in the table, among the trials, there appeared to be more *APOE4* carriers in the ADCS and ADNI trials than in the others. Since *APOE4* carrier status is associated with amyloid deposition and progression to dementia due to AD, this may have been an important factor in the progression rates observed across the trials.[17] The ADCS and ADNI trials behaved in a more predictable fashion, with participants progressing in a regular fashion through the MCI state with worsening metrics on the ADAS-Cog, MMSE, and CDR-SB, and an approximate annual progression rate of 16% per year was obtained. The progression rates in the other studies were variable and less than what they had anticipated, and in part, this could be due to fewer participants being *APOE4* carriers.

A complicating factor may have arisen concerning the state-dependent endpoints. That is, each study required the MCI participants to achieve the diagnosis of dementia or a surrogate, e.g., CDR 1. If a continuous cognitive measure had been used as the outcome, as suggested by the recent FDA guidelines, more consistent findings may have arisen.[18] However, this assertion must be tempered by the variable performance on one of the common measures, e.g., ADAS-Cog, as is shown in figure 2.

The non-NIH trials were also conducted in many countries using multiple languages. The MCI construct was relatively new, especially during the time that these trials were conducted, and the cultural differences in characterizing participants with mild impairments as well as those reaching the dementia threshold could be variable. It is not clear what effect the translation of the multiple instruments into several languages would have on the outcome, but that likely added to the variability of the performance of the trials. Several of the trials, e.g., InDDEx trial, commented on the inclusion of participants with multiple comorbidities and exclusions of participants with other relevant conditions like depression. As such, the participant populations at the MCI stage may have been sufficiently variable with regard to comorbidities and cultural differences to add an element of complexity to interpreting the data.

There are several lessons to be learned from the comparison of these trials. No single trial was better or worse than any of the others, but there were multiple sources of variability. From the ADCS and ADNI trials, it appears that, if the studies involve a single language and culture, and include multiple academic centers, the implementation of the criteria for aMCI may be more standardized. However, given the subtle nature of the clinical abnormalities in some of these participants, there can be differences in performances of the groups. One factor affecting the outcome was the severity of the participants at the point of entry. Subtle differences in the memory impairment criteria can have a significant effect on the rate of progression of the participants. As such, those trials that had less impaired participants had lower progression rates than those with participants who were more severely impaired at entry. This may have resulted in some participants being further from the endpoint of the study at entry, and the rate of progression in participants with milder disease is known to be slower.[19]

The role of *APOE4* carrier status is also apparent. It is likely that those participants who were *APOE4* carriers had more underlying pathology at entry than noncarriers, and that resulted in a greater degree of impairment at entry, which could then have influenced the rate of progression and time to reach the endpoint. The amount and type of data available to the clinicians to make the endpoint assessment is also important. Some

Comment: MCI trials—Categorical "square pegs" in dimensional "round holes"?

Given repeated disappointments from trials of treatments for Alzheimer disease (AD), one hoped for better results from upstream interventions at the mild cognitive impairment (MCI) stage. However, results there have been almost as bleak. A new report[1] considers explanations for the latter: varying participant selection, observation methods, and outcomes assessment. Cleverly, the authors avoided differences in efficacy of various treatments by considering only the trials' placebo-assigned control groups. This design permitted inclusion of the Alzheimer's Disease Neuroimaging Initiative (ADNI) MCI cohort, assembled similarly to that of the Alzheimer's Disease Cooperative Study (ADCS).

The findings are sobering. They highlight difficulties in trial design within a clinical diagnosis, MCI, that is increasingly understood as a slice in time, without natural boundaries, in the evolution of AD. The article's figure 2 shows marked variability in the various trial cohorts' Alzheimer's Disease Assessment Scale–cognitive subscale (ADAS-Cog) scores at baseline. Importantly also, ADAS scores for all trial cohorts except the ADCS and ADNI trials declined over time, suggesting that familiarity or practice effects may have obscured disease-related cognitive decline. As for outcomes, randomization and masking may protect against bias, but imprecision in diagnosis (owing again to lack of natural boundaries) may still have diluted treatment effects.

Figure 3 suggests a rapid exaggeration in symptoms at the first follow-up examination. Such a phenomenon could reflect selective enrollment of persons with near-normal Mini-Mental State Examination scores who were, in fact, sicker than their entry scores suggested. One might wonder about laxity of some clinics that were reimbursed on a per-person-enrolled basis. After this first follow-up, symptom progress appears haphazard. This suggests there was not much disease-related decline thereafter for the treatments to mitigate.

What to do now? Creation of a category for the slice in time that is MCI, while useful clinically, can be problematic for trial design. Even so, the longitudinal performance declines among the ADCS and ADNI cohorts suggest that strict criteria and meticulous methods can offer a target for disease-modifying interventions. Addition of more objective endpoints (various biomarkers, perhaps even multiple) may also help.[2] By contrast, functional difficulty, while of fundamental clinical importance, may be problematic as a trial outcome, since its variable origins may obfuscate real treatment effects. If only we understood better the true nature of AD—not just its features or manifestations or tautological animal models—we might have better luck assessing treatments that modify its progression.

1. Petersen RC, Thomas RG, Aisen PS, et al. Randomized clinical trials in mild cognitive impairment: sources of variability. Neurology 2017;88:1751–1758.
2. Leoutsakos JM, Gross AL, Jones RN, Albert MS, Breitner JCS. Alzheimer progression score: development of a biomarker summary outcome for AD prevention trials. J Prev Alz Dis 2016;3:229–235.

John C.S. Breitner, MD, MPH

From the Centre for Studies on Prevention of AD, Douglas Mental Health University Institute Research Centre, and the Department of Psychiatry and Behavioral Sciences, McGill University Faculty of Medicine, Montreal, Canada.

Disclosure: The author reports no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

Study funding: Supported by McGill University, the Government of Canada, Fonds de Rechérche du Quebec–Santé, and an unrestricted research gift fund from Pfizer Canada.

trials required the clinical diagnosis of dementia with review by a central committee, and other trials used progression to CDR 1 as the endpoint. It is clear that participants can progress from MCI to dementia while still at the CDR 0.5 stage, and consequently, variability can be introduced by the nature of the endpoint.

Clinical trials designed to involve participants at the MCI due to AD stage can be improved. Clearly, through the use of AD-specific biomarkers, the clinical characteristics of the participants can be stratified and a more homogeneous group of participants can be enrolled. If the intervention is targeted at amyloid, amyloid positivity on PET scanning or CSF Aβ42 levels can be used to select participants. It is uncertain as to whether these biomarkers will overcome some of the other sources of variability encountered in the trials described, but that is a distinct possibility. The implementation of continuous outcome measures rather than clinical states may improve performance of trials, but clinical meaningfulness needs to be established. However, attention to other factors such as *APOE*4 carrier status, language, culture, and comorbidities still needs to be entertained in designing trials in this state of the disease. Finally, these factors are not peculiar to trials of MCI due to AD but have implications for other chronic disorders of a gradually progressive nature. Several important lessons have been derived from analyzing the previously conducted trials.

AUTHOR CONTRIBUTIONS

Dr. Petersen: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data. Dr. Thomas: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data. Dr. Aisen: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data. Dr. Mohs: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data. Dr. Carrillo: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data. Dr. Albert: obtaining data, reviewing data, reviewing manuscript.

ACKNOWLEDGMENT

The authors thank the FNIH staff, Judy Siuciak, Steve Hoffmann, Dorothy Jones-Davis, and Rosa Canet-Aviles for their assistance with this project.

STUDY FUNDING

Supported by NIH and FDA (P50 AG016574, U01 AG006786, U01 AG024904, P50 AG005146, U19 AG033655, T32 AG027668). The data described within this document represent the work of the Biomarkers Consortium Project "Placebo Data Analysis in Alzheimer's Disease and Mild Cognitive Impairment Clinical Trials." This project was submitted to the Biomarkers Consortium Neuroscience Steering Committee for execution and was managed by a Biomarkers Consortium Project Team. In addition to the NIH and FDA, participating and funding organizations include Alzheimer's Association (HAT-09-136580), Abbott, AstraZeneca, Boehringer-Ingelheim, Bristol-Myers Squibb, Eisai, Inc, Eli Lilly and Company, Merck and Company, Pfizer Inc., and Takeda Global Research and Development Center, Inc.

DISCLOSURE

R. Petersen is a consultant for Roche, Inc., Merck, Inc., Genentech, Inc., Biogen, Inc., and Eli Lilly and Company and receives royalties from Oxford University Press for *Mild Cognitive Impairment*. R. Thomas has nothing to disclose. P. Aisen has served as a consultant to the following companies: NeuroPhage, Elan, Eisai, Bristol- Myers Squibb, Eli Lilly, Merck, Roche, Amgen, Genentech, Abbott, Pfizer, Novartis, AstraZeneca, Janssen, Medivation, Ichor, Toyama, Lundbeck, Biogen Idec, iPerian, Probiodrug, Somaxon, Biotie, Cardeus, Anavex, AbbVie, and Cohbar. R. Mohs was a full-time employee and shareholder of Eli Lilly and Company. M. Carrillo is a full-time employee of the Alzheimer's Association. M. Albert is a consultant for Biogen Idec, Eli Lilly, and Genentech, Inc. Go to Neurology.org for full disclosures.

Received May 6, 2015. Accepted in final form January 5, 2017.

REFERENCES

1. Sperling RA, Jack CR Jr, Aisen PS. Testing the right target and right drug at the right stage. Sci Transl Med 2011;3:111cm33.
2. Jack CR Jr, Albert MS, Knopman DS, et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:257–262.
3. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004;256:183–194.
4. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment: beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004;256:240–246.
5. Albert MS, DeKosky ST, Dickson D, et al. 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. Alzheimers Dement 2011;7:270–279.
6. Petersen RC, Thomas RG, Grundman M, et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. N Engl J Med 2005;352:2379–2388.
7. Winblad B, Gauthier S, Scinto L, et al. Safety and efficacy of galantamine in subjects with mild cognitive impairment. Neurology 2008;70:2024–2035.
8. Feldman HH, Ferris S, Winblad B, et al. Effect of rivastigmine on delay to diagnosis of Alzheimer's disease from mild cognitive impairment: the InDDEx study. Lancet Neurol 2007;6:501–512.
9. Thal LJ, Ferris SH, Kirby L, et al. A randomized, double-blind, study of rofecoxib in patients with mild cognitive impairment. Neuropsychopharmacology 2005;30:1204–1215.
10. Doody RS, Ferris SH, Salloway S, et al. Donepezil treatment of patients with MCI: a 48-week randomized, placebo-controlled trial. Neurology 2009;72:1555–1561.
11. Grundman M, Petersen RC, Bennett DA, et al. Alzheimer's association Research roundtable meeting on mild cognitive impairment: what have we learned? Alzheimers Dement 2006;2:220–233.
12. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984;141:1356–1364.
13. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
14. Wechsler DA. Wechsler Memory Scale–Revised. San Antonio: Psychological Corporation; 1982.
15. Mathews M, Abner E, Caban-Holt A, Dennis BC, Kryscio R, Schmitt F. Quality of education and memory test performance in older men: the New York University paragraph recall test normative data. Curr Alzheimer Res 2013;10:776–783.
16. Rey A. l'Examen Clinique en Psychologie. Paris: Presses Universitaires de France; 1964.
17. Sanan DA, Weisgraber KH, Russell SJ, et al. Apolipoprotein E associates with beta amyloid peptide of Alzheimer's disease to form novel monofibrils: isoform apoE4 associates more efficiently than apoE3. J Clin Invest 1994;94:860–869.
18. Kozauer N, Katz R. Regulation of drugs for early Alzheimer's disease. New Engl J Med 2013;369:288.
19. Farias ST, Mungas D, Reed BR, Harvey D, DeCarli C. Progression of mild cognitive impairment to dementia in clinic- vs community-based cohorts. Arch Neurol 2009;66:1151–1157.