Differences in Alzheimer disease clinical trial outcomes based on age of the participants

ABSTRACT

Objective: We tested the a priori hypothesis that older participants differ in rates of decline on cognitive outcomes compared with younger participants, and examined the potential effect of age distributions on individual clinical trial outcomes.

Methods: From a meta-database of 18 studies from the Alzheimer’s Disease Cooperative Study and the Alzheimer’s Disease Neuroimaging Initiative, we included a cohort of 2,793 participants for whom there were baseline demographic data and at least one postbaseline cognitive assessment on the Alzheimer’s Disease Assessment Scale–cognitive subscale (ADAS-cog), Clinical Dementia Rating–Sum of Boxes (CDR-SB), or Mini-Mental State Examination (MMSE). We used mixed-effects models (random coefficient models) to estimate change on the outcomes across 7 age groups ranging from younger than 61 years to older than 85 years after adjusting for education.

Results: Significant worsening occurred in all age groups on all outcomes over time. The 4 older groups, aged 71 years and older, showed slower rates of decline on the ADAS-cog than the younger groups (p = 0.001). The older groups scored 2–3, 2–5, and 4–6 points better than the younger groups at 12, 18, and 24 months, respectively. There were similar differences across age groups for the MMSE, but not for the CDR-SB.

Conclusions: The differences in change on the ADAS-cog between older and younger participants are substantially greater than differences expected between experimental drugs and placebo in current trials or differences between marketed cholinesterase inhibitors and placebo. The clinical interpretation of change on the ADAS-cog or MMSE differs depending on age. Until predictors of decline are better understood, considering effects of age on rates of change is particularly important regarding clinical practice and outcomes of trials.

GLOSSARY

AD = Alzheimer disease; ADAS-cog = Alzheimer’s Disease Assessment Scale–cognitive subscale; CDR-SB = Clinical Dementia Rating–Sum of Boxes; MMSE = Mini-Mental State Examination.

Analyses of several Alzheimer observational studies1–3 and clinical trials4 suggest that older participants decline less on cognitive outcomes than younger participants, although this finding is not uniform.1,5 This may be attributable to selection biases of who enrolls in trials; it also may be attributable to the pathogenesis and virulence of Alzheimer disease (AD) reflected by age at onset. Nevertheless, any age effect may have resulted in an attenuation of measurable treatment effects or decreased likelihood to detect differences between drug and placebo. Some clinical trial protocols constrain the lower and upper age limits for study entry, thus affecting the distribution of younger and older participants and possibly the trial outcomes.6 It is not clear, however, how robust any age-associated effect may be, how individual trials may be affected, or how this affects clinical meaning.7 We assessed the extent of this phenomenon using pooled clinical trials data.

From the Departments of Psychiatry and Neurology (L.S.S.), Keck School of Medicine of USC, Los Angeles, CA; Division of Gerontology, Geriatrics, and Palliative Care, Department of Medicine (R.E.K.), and Department of Biostatistics (G.W., G.R.C.), University of Alabama at Birmingham.

Presented, in part, at Alzheimer’s Association International Conference, Copenhagen, Denmark, July 15, 2014.

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.
METHODS We selected participants from a meta-database consisting of 18 studies from the Alzheimer’s Disease Cooperative Study and the Alzheimer’s Disease Neuroimaging Initiative conducted from 1995 to 2012 to analyze the decline on the Alzheimer’s Disease Assessment Scale-cognitive subscale (ADAS-cog), Clinical Dementia Rating—Sum of Boxes (CDR-SB) scale, and Mini-Mental State Examination (MMSE) over time. Participant selection criteria for the analysis were the selection criteria for the respective studies. Additional inclusion criteria were (1) diagnosis of mild to moderate AD dementia, and (2) at least one assessment on the ADAS-cog, CDR-SB, or MMSE. We analyzed the 10 studies meeting these requirements. All diagnoses of AD were based on National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer’s Disease and Related Disorders Association criteria, with the additional requirement of a minimal severity based on clinical ratings. These were a CDR of ≥2 for the SL trial and MMSE scores between 14 and 26 (DHA, HC), between 12 and 28 (CE), between 12 and 26 (LL), between 13 and 26 (PR, NS), between 10 and 24 (HU), between 12 and 26 (VN) (Alzheimer’s Disease Neuroimaging Initiative). We assessed outcomes at 6-month intervals over 2 years, with the a priori hypothesis that older participants would differ in rates of decline on cognitive outcomes compared with younger participants. Based on the sample size, participants were divided into 5-year age categories of 48–60, 61–65, 66–70, 71–75, 76–80, 81–85, and 86–105 years; the groups younger than 55 years were merged with the 55–60 group, and the groups older than 90 years merged with the 86–90 group, because of the small number of participants in those age ranges. We used mixed-effects models (random coefficient models) to compare the rate of decline in the outcomes scores between the reference group 60 years and younger and each of the remaining age groups, adjusting for education. The mixed-effects model was selected because it utilizes data from all participants (rather than just completers), minimizes bias, and better controls for type I error in the presence of missing data.13 The slope (rate of decline) and intercept (baseline score) were modeled as independent. We also conducted a sensitivity analysis restricting participants to those in the placebo arms of the parent studies.

RESULTS Of the 5,990 participants available in the meta-database, 2,799 from the 10 studies met inclusion criteria. Six participants were excluded because of missing baseline demographic data or cognitive assessment, yielding 2,793 participants for analysis. The number of participants in each age category at each follow-up time is shown in table e-1 on the Neurology® Web site at Neurology.org. The age ranges allowed by the included studies were from 50 years in 5 studies, 60 in 1, no lower age limit in 3, and a range from 55 to 90 in 1 trial. Individuals in the older age groups had fewer years of formal education, were less likely to be married at entry, and were more likely to be Caucasian than the younger age groups (table 1). Unadjusted baseline ADAS-cog scores differed among age groups (p = 0.002) with lower (better) scores in the

Table 1 Subject characteristics at baseline and change in ADAS-cog scores over the durations of the trials

<table>
<thead>
<tr>
<th>Age category, y</th>
<th>No. 48-60</th>
<th>61-65</th>
<th>66-70</th>
<th>71-75</th>
<th>76-80</th>
<th>81-85</th>
<th>86-105</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. (% total = 2,793)</td>
<td>2,793</td>
<td>1,793 (6.8)</td>
<td>192 (8.9)</td>
<td>354 (12.7)</td>
<td>583 (20.9)</td>
<td>722 (25.8)</td>
<td>546 (19.5)</td>
<td>217 (7.8)</td>
</tr>
<tr>
<td>Education &lt; high school</td>
<td>2,793</td>
<td>6</td>
<td>10</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>21</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2,793</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>0.95</td>
</tr>
<tr>
<td>Married</td>
<td>2,793</td>
<td>83</td>
<td>85</td>
<td>83</td>
<td>81</td>
<td>71</td>
<td>59</td>
<td>42</td>
</tr>
<tr>
<td>Caucasian</td>
<td>2,793</td>
<td>85</td>
<td>92</td>
<td>87</td>
<td>91</td>
<td>91</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>Female</td>
<td>2,793</td>
<td>58</td>
<td>58</td>
<td>54</td>
<td>58</td>
<td>59</td>
<td>61</td>
<td>63</td>
</tr>
<tr>
<td>Assigned to placebo</td>
<td>2,793</td>
<td>35</td>
<td>42</td>
<td>44</td>
<td>44</td>
<td>42</td>
<td>49</td>
<td>45</td>
</tr>
<tr>
<td>APO-ε4 carrier</td>
<td>1,955</td>
<td>53</td>
<td>65</td>
<td>74</td>
<td>74</td>
<td>66</td>
<td>57</td>
<td>48</td>
</tr>
<tr>
<td>MMSE</td>
<td>2,791</td>
<td>18.6 ± 5.7</td>
<td>18.7 ± 5.6</td>
<td>19.2 ± 5.5</td>
<td>19.9 ± 4.7</td>
<td>19.6 ± 4.5</td>
<td>19.6 ± 4.2</td>
<td>19.3 ± 3.8</td>
</tr>
<tr>
<td>CDR-SB</td>
<td>2,191</td>
<td>6.4 ± 3.1</td>
<td>6.4 ± 3.2</td>
<td>6.5 ± 3.2</td>
<td>6.3 ± 3.1</td>
<td>6.6 ± 3.1</td>
<td>6.7 ± 3.1</td>
<td>7.3 ± 2.9</td>
</tr>
<tr>
<td>ADAS-cog</td>
<td>2,793</td>
<td>27.8 ± 13.3</td>
<td>28.2 ± 12.3</td>
<td>26.4 ± 12.1</td>
<td>24.4 ± 10.4</td>
<td>25.6 ± 10.5</td>
<td>25.3 ± 10.1</td>
<td>26.3 ± 9.1</td>
</tr>
<tr>
<td>ADAS-cog change, 6 mo</td>
<td>2,157</td>
<td>3.0 ± 5.4</td>
<td>2.39 ± 5.52</td>
<td>2.91 ± 5.74</td>
<td>2.51 ± 5.86</td>
<td>2.46 ± 5.35</td>
<td>2.49 ± 5.65</td>
<td>1.40 ± 5.02</td>
</tr>
<tr>
<td>ADAS-cog change, 12 mo</td>
<td>1,882</td>
<td>7.04 ± 7.42</td>
<td>6.21 ± 7.37</td>
<td>6.24 ± 7.14</td>
<td>5.16 ± 7.11</td>
<td>4.68 ± 6.33</td>
<td>3.63 ± 6.63</td>
<td>2.96 ± 5.56</td>
</tr>
<tr>
<td>ADAS-cog change, 18 mo</td>
<td>1,100</td>
<td>12.09 ± 9.17</td>
<td>9.33 ± 9.00</td>
<td>8.83 ± 7.70</td>
<td>7.58 ± 8.47</td>
<td>6.92 ± 7.91</td>
<td>5.42 ± 7.68</td>
<td>4.38 ± 6.37</td>
</tr>
<tr>
<td>ADAS-cog change, 24 mo</td>
<td>340</td>
<td>15.4 ± 9.3</td>
<td>16.9 ± 10.3</td>
<td>14.2 ± 8.4</td>
<td>9.8 ± 9.6</td>
<td>8.7 ± 6.7</td>
<td>7.5 ± 7.0</td>
<td>7.7 ± 6.6</td>
</tr>
</tbody>
</table>

Abbreviations: ADAS-cog = Alzheimer’s Disease Assessment Scale-cognitive subscale; CDR-SB = Clinical Dementia Rating—Sum of Boxes; MMSE = Mini-Mental State Examination.

Summary data for the sample consists of frequencies (percentages) for categorical variables and means ± SDs for continuous variables. Groups were compared with χ² tests for categorical variables and analysis of variance for continuous variables. See table e-1 for changes in CDR-SB and MMSE scores over the durations of the trials.

*Significant after correction for multiple comparisons.
group of participants aged 71 to 75 years, but without a clear trend across age groups (table 1).

ADAS-cog scores increased (i.e., worsened) over time in all age groups ($p < 0.001$). The rate of decline became progressively smaller with older age groups relative to the age 48–60 group, although this only achieved statistical significance for those older than 65 (table 2, figure 1). Significant differences in baseline were only observed for the 71–75, 76–80, and 81–85 age groups. Post hoc tests showed that the 4 oldest age categories, 71 years and older, had significantly slower rates of decline compared with the 70 years and younger group ($\chi^2 = 55.47, df = 1, p < 0.001$). The older age groups worsened about 3–5 points by 12 months, 4–7.5 points at 18 months, and 7.5–10 points by 24 months. The 70 years and younger groups worsened by about 6–7 points at month 12, 9–12 points at month 18, and 14–15 points at month 24. These outcomes were confirmed in the analytic sample that received only placebo in the original trials (figure e-1).

The SDs of the changes by age were also systematically smaller compared with the younger participants, indicative of less heterogeneity in the rate of decline among the older participants (all $p < 0.001$, figure 2).

There were similar differences across age groups and over time for the MMSE (table 1, figure 1, and table e-2). Finally, CDR-SB scores also increased (i.e., worsened) over time in all age groups ($t = 11.95, df = 5,283, p < 0.001$; data not shown); however, there were no significant differences across age groups (all $p > 0.37$) (figure 1).

**DISCUSSION** Younger participants with mild to moderate AD dementia who were enrolled in clinical trials show substantially greater worsening on the ADAS-cog and MMSE over the course of 12 to 24 months than older participants. The differences in rate of change on the ADAS-cog between older and younger participants are substantially greater than differences expected between experimental drugs and placebo in current trials or from the observed effects of marketed cholinesterase inhibitors compared with placebo. The 4- to 6-point difference based on age category across 24 months of clinical trials duration is 2 to 3 times the differences planned for in the current design of clinical trials of experimental drugs for mild or moderate dementia due to AD,$^6$ or the actual differences in marketed cholinesterase inhibitors trials compared with placebo.$^{24,25}$

Varying the distributions or mixtures of ages in a trial, therefore, may have profound effects on the outcomes of such trials even if the percentage reduction is the same. Thus, a 40% slope reduction in those 71 years and older would correspond to a reduction of 3 or 4 points on the ADAS-cog in a 24-month trial, while the reduction in those 70 years and younger would be 5 or 6 points. Such differences in scores (with corresponding differences in effect sizes) may be critical in determining the success of a clinical trial. Nevertheless, if a drug is differentially effective because of age, then the effect may not be observed if age by treatment interactions are not considered or planned. Such adjustments for age, or stratifying analyses by age group, would potentially alleviate this effect but would require that age is not systematically associated with disease progression (based on cognitive outcome measures) and treatment response. For example, if younger age is associated with more rapidly progressive disease that is treatment-resistant, while older age is more slowly progressive but treatment-responsive, a larger trial with an older age group would be needed to detect treatment effects. Considering age in sample selection for a trial may be particularly important with respect to the underlying neuropathology, availability of participants, trials outcomes, and external validity.

Of participants enrolled in these trials, 13.3% and 26.0% were younger than 66 and 71 years,
respectively; participants older than 80 years comprised 26.7% of the sample. AD dementia in those younger than 70, however, is a fairly uncommon condition with a prevalence of less than 1% in the 65 to 69 age range, increasing substantially to approximately 18% in the 80 to 84 age range. Therefore, younger participants are relatively overrepresented and older participants underrepresented in the pooled studies. Although they showed greater worsening on the ADAS-cog, few of the younger patients would have
2.15, df 5 months at baseline (SDs were calculated for each age category at each time point of baseline, 6, 12, 18, and 24 months at baseline (< 60 vs. 71–75, F = 1.64, df = 178, 582, p < 0.001; vs 76–80, F = 1.61, df = 178, 721, p < 0.001; vs 81–85, F = 1.75, df = 178, 543, p < 0.001; vs 86–105, F = 2.15, df = 178, 216, p < 0.001). ADAS-cog = Alzheimer’s Disease Assessment Scale-cognitive subscale.

SDs were calculated for each age category at each time point of baseline, 6, 12, 18, and 24 months at baseline (< 60 vs. 71–75, F = 1.64, df = 178, 582, p < 0.001; vs 76–80, F = 1.61, df = 178, 721, p < 0.001; vs 81–85, F = 1.75, df = 178, 543, p < 0.001; vs 86–105, F = 2.15, df = 178, 216, p < 0.001). ADAS-cog = Alzheimer’s Disease Assessment Scale-cognitive subscale.

fulfilled proposed criteria for rapidly progressive AD dementia that include in part an annual decline of 6 MMSE points,28 and does not appear to be particularly associated with younger age.29

Because AD is multidetermined and pleomorphic, relatively older participants compared with younger participants in a trial may represent those with additional or different risks or pathology,29 including, for example, vascular disease, hippocampal sclerosis, TDP-43 (TAR DNA-binding protein 43) proteinopathy, and other comorbid non-Alzheimer neuropathology, or they may represent those in whom clinical AD and pathology is more complex.30 Thus, one explanation for the smaller degree of change in cognitive scores in older participants is that misclassified older participants may decline less. It is also possible that participants who are available for trials at relatively older ages may decline more slowly because of survival bias, with the more rapidly declining participants already removed from the prevalent pool.

Potential limitations to these analyses include that the datasets for the selected trials span 2 decades, and secular changes in participants, selection, trials conduct, test administration, and site differences31 may have influenced the results. Because the sites participating were mostly members of the Alzheimer’s Disease Cooperative Study, however, it is probable that these are lesser sources of error than with other pooled databases. Another limitation is that the ADAS-cog is a poorly metricized scale, nonlinear regarding its change over longer periods of time, and not an interval scale. For example, individual items, such as orientation, word recall, word recognition, and naming, contribute disproportionately to total ADAS-cog change scores,32 and the extent to which older and younger patients change on those items may differ. The lack of differences seen with the CDR-SB may imply that the cognitive and functional components of this measure are less sensitive than the ADAS-cog or MMSE to these issues, the range of the scale is constricted, or that the noncognitive components of the scale are less sensitive to the effects of age. It is also possible that the scale is simply unable to identify the more subtle changes in this group.

The decreased variability in the ADAS-cog and MMSE by age suggests it is more difficult to predict the course in younger compared with older individuals in clinical studies. Therefore, setting an absolute threshold for improvement, e.g., change of 2 to 4 points in the ADAS-cog, would correspond to a smaller effect size in younger individuals than in older individuals and create potentially more misclassification in younger relative to older patients with significant change.

As a clinical research tool, absolute cutoff scores on the ADAS-cog have been recommended to estimate change for individual patients or to explain the effects of drugs. For example, 2 or 4-point differences on the ADAS-cog—at least over 6 to 18 months—have been considered to be clinically meaningful,33–35 and 3-point differences have been interpreted as the equivalent of a 6-month change in clinical course. Results here suggest that 2- to 4-point differences or less may in fact miss a number of older individuals showing a significant change that is less than this threshold. Clinical interpretation of change on both the ADAS-cog and MMSE should therefore be considered within the context of age and baseline scores, and perhaps with other measures.

These findings have implications as well for mild cognitive impairment and prevention trials because they are more likely to include younger participants and participants with Alzheimer neuropathology.36 Composite scales currently used in secondary prevention or prodromal AD trials include the orientation items from the MMSE and word lists similar to that used in the ADAS-cog. Our results suggest that any positive effects from interventions more likely would be detected in younger study participants because slowing of clinical progression would be more evident. However, our results also raise concerns that outcomes from trials involving younger individuals may not necessarily apply to older individuals seen in the clinic, which reflect the majority of patients with AD and ages of greatest incidence. While the use of biomarker enrollment criteria (such as amyloid imaging) may result in a more homogeneous sample and more accurate classification of Alzheimer pathology across age groups, it is not certain that biomarker
requirements would lead to better extrapolation from younger to older ages. Thus, consideration should be given to including older adults in trials, despite the perceived problems. An additional implication is that defining successful prevention should vary by the age of the participants.

Although these findings need to be replicated in other databases, it appears older individuals with AD dementia enrolled in clinical trials show substantially less cognitive worsening measured with the ADAS-cog or MMSE than younger individuals, and this needs to be accounted for in clinical trial designs. The clinical interpretation of change on the ADAS-cog may also differ depending on age. Until predictors or markers of decline are better understood, considering age in sample selection may be particularly important regarding clinical management and therapeutic trial outcomes.

Comment:
Age effects on clinical trial results in Alzheimer dementia

Therapeutic trials in Alzheimer disease (AD) are notoriously difficult and have produced no new approved treatments in the past decade. Trial success often depends on decline in a placebo arm, and population characteristics that diminish placebo decline reduce the chance of detecting positive therapeutic effects of a drug. This study of data from multiple trials of AD dementia demonstrates that age can have such an effect: older participants showed less decline than younger participants by differences that were both meaningful and greater than expected.

Age effects were limited to cognitive test results. Baseline imbalances likely accounted for some of the effect on decline since older patients had better initial cognitive scores, a predictor of slower progression. Differences in baseline scores would have been even more pronounced if age-adjusted. Such considerations are worth exploring, but this hardly blunts the key point that ignoring age effects on progression in study populations can negatively influence trial results.

These effects may be mitigated by several strategies, including stratification, lowering an age threshold, and more rigorous exclusion of comorbid conditions, each with some tradeoff in increased trial burden or consequences. Another mitigation is improvement of diagnostic accuracy. The recent negative bapineuzumab and solanezumab studies enrolled an unexpectedly high percentage of patients who may also differ depending on age. Until predictors or markers of decline are better understood, considering age in sample selection may be particularly important regarding clinical management and therapeutic trial outcomes.

REFERENCES

H. Robert Brashear, MD

From Janssen Research and Development, LLC, Fremont, CA.

STUDY FUNDING
Supported by NIH R01 AG037561.

DISCLOSURE
L. Schneider reports being an editor on the Cochran Collaboration Dementia and Cognitive Improvement Group, which oversees systematic reviews of drugs for cognitive impairment and dementia; receiving a grant from the Alzheimer’s Association for a registry of dementia and cognitive impairment trials; within the past 3 years receiving grant or research support from NIA, Baxter, Eli Lilly, Forum, Genentech, Lundbeck, Merck, Novartis, Pfizer, and Tauftids; and having served as a consultant for or receiving consulting fees from AC Immune, Allons, AstraZeneca, Avraham Pharmaceutical, Ltd., Baxter, Biogen Idec, Ceretec, Cypco, Elan, Eli Lilly, Forum, GlaxoSmithKline, Johnson & Johnson, Lundbeck, Merck, Pfizer, Roche, Servier, Takeda, Toyama, and Zinfandel. R. Kennedy reports receiving grant support from NIA, National Institute of Neurological Disorders and Stroke, NHLBI, NIDDK, and the Department of Education. G. Wang reports receiving grant support from NIA. G. Cutters reports receiving grant or research support from Participation of Data and Safety Monitoring Committees: all of the following organizations are focused on medical research: Atopek, Biogen Idec, Cleveland Clinic, GlaxoSmithKline Pharmaceuticals, Gilead Pharmaceuticals, Modigenetech/Postor, Merck/Omo Pharmaceuticals, Merck, Neuren, Realesio, Sanofi-Aventis, Teva, Vivus, NHLBI (Bone Marrow Transplant Protocol Review Committee), National Institute of Neurological Disorders and Stroke, NMSS, and NICHD (OPRU oversight committee). Consulting, speaking fees, and advisory boards: Alexion, Alkyn, Bayer, Celgene, Corrado Biosciences, Consortium of MS Centers (grant), DsoGenix, Klein Buendel Incorporated, MedImmune, Novartis, Nuron Biotech, Receptors, Spinifex Pharmaceuticals, and Teva Pharmaceuticals. Dr. Cutters is employed by the University of Alabama at Birmingham and President of Pythagoras, Inc., a private consulting company located in Birmingham, AL. Go to Neurology.org for full disclosures.

AUTHOR CONTRIBUTIONS
Lon S. Schneider, MD: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis, and interpretation of data. Richard E. Kennedy, MD, PhD: revising the manuscript for content, including medical writing for content, study concept and design, analysis, and interpretation of data. Guoguo Wang, MS: revising the manuscript for content, analysis, and interpretation of data. Gary R. Cutters, PhD: revising the manuscript for content, study concept and design, analysis, and interpretation of data.

STUDY FUNDING
Supported by NIH R01 AG037561.


