Using full AD clinical trials databases for simulating and modeling future clinical trials

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

University of Southern California Keck School of Medicine, Los Angeles
University of Alabama at Birmingham, Birmingham
ISCTM Winter Meeting
February 17, 2015
Washington, DC
Open access journals to advance diagnosis, assessment, translational research, clinical interventions

http://www.alzheimersanddementia.com/trci

http://www.alzheimersanddementia.com/dadm
Overarching context

• Obstacles to translating pre-clinical research to clinical
• Imperative to do more trials with fewer patients; to “get a signal” earlier….many drugs and no validated targets
• Clinical trials often don’t turn out as planned
• 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”
• Lack of success of disease-modifying treatments led to recommendations to identify subgroups that are more likely to respond
  – Post-hoc analyses of clinical trials
  – Predictors of progression in observational studies
• We investigated effects of implementing several recently suggested changes for enrichment
Data sources

• Data from meta-database
  – (R01AG03756, Synthesis of Longer-Term Alzheimer Disease Studies)
  – Alzheimer’s Disease Cooperative Study
  – Alzheimer’s Disease Neuroimaging Initiative
  – Coalition Against Major Diseases

• More than 6500 participants
  – Not-impaired to MCI to AD
  – 6.7% African American
  – 8.2% Hispanic

Simulation methods

- Designed to mimic a typical clinical trial with treatment and placebo arm, 1:1 allocation
- Simulated typical clinical trials using resampling methodology
  - 50-400/arm, 12-24 months, 20-40% dropouts
- Treatment and placebo groups constructed by resampling from the entire meta-database
- Treatment effect added to scores in treatment group, with effect sizes of 0.15 to 0.55

Schneider et al, *Alzheimers Dement*, 2010
List of studies

- *Post hoc* ApoE
- Minority representation
- Biomarkers ApoE
- Biomarkers Aβ
- Age
- Adaptive design: sample size reestimation
Publications

- LS Schneider, RE Kennedy, G Wang, GR Cutter. Differences in Alzheimer Disease Clinical Trials Outcomes Based on the Age of the Participants. Neurology, in press.
Presentations/posters

• Testing Subgroup Analyses and Enrichment in AD Clinical Trials Using a Meta-database. Presented at the 2014 CTAD
• Differences in Alzheimer Disease Clinical Trials Outcomes Based on Age of Patients. Presented at the 2014 AAIC
• Use of Meta-database Data for Quality Control in Clinical Trials: An Example Using the MMSE. Presented at the 2014 AAIC
• Baseline cognitive severity does not predict rate of change in the ADAScog clinical trials. Lon S. Schneider, Richard E. Kennedy, Peter Wang, Gary R. Cutter. Presented at the 2014 AAIC
• Post-hoc Analyses of ApoE4 Effects in Clinical Trials: A Cautionary Note. Presented at 2013 AAIC
• Effect of ApoE Genotype Status on Targeted Clinical Trials Outcomes and Efficiency. Presented at the 2012 AAIC
• Alzheimer Disease Trials Simulations to Test New Research Criteria, Biomarkers, and other Proposed Methodological Improvements. Presented at the 2011 ACNP
• Amnestic MCI/ prodromal AD with more severe memory deficit is indistinguishable from diagnosed AD: Implications for the validity of clinical trials and biomarkers. 2011 AAIC (ICAD)
• Biomarker Positive and Negative Subjects in the ADNI: Clinical Characterization. 2011 Annual Meeting of the American Neuropsychiatric Association
• Biomarker Enrichment in Clinical Trials: Caveats from the ADNI Dataset Presented at the 2011 Annual Meeting, the Eastern North American Region of the International Biometrics Society
Trials Outcomes Analyzed by ApoE Status

- Rosiglitazone
- Bapineuzumab phase 2
- Tarenflurbil
- Bapineuzumab phase 3
- Future trials
  - Abeta antibodies
  - Pioglitazone (Takeda/Zinfandel)
  - Tramiprosate (Alzeon)
Bapineuzumab 201 Trial

Bapineuzumab Phase II data

Modified Intention to Treat

<table>
<thead>
<tr>
<th>Cognition</th>
<th>Change from Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bapineuzumab N=119</td>
<td>Placebo N=108</td>
</tr>
<tr>
<td>Rx difference at week 78 = 2.31</td>
<td>Rx difference over time</td>
</tr>
<tr>
<td>p=0.078</td>
<td></td>
</tr>
</tbody>
</table>

Completers

<table>
<thead>
<tr>
<th>Cognition</th>
<th>Change from Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bapineuzumab N=78</td>
<td>Placebo N=78</td>
</tr>
<tr>
<td>Rx difference at week 78 = 4.35</td>
<td>Difference increases over time</td>
</tr>
<tr>
<td>p=0.003</td>
<td></td>
</tr>
</tbody>
</table>

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

Bapineuzumab also showed a statistically significant impact on cognitive decline in ApoE4 non-carriers.

Bapineuzumab Phase II data in ApoE4 non-carriers

Modified Intention to Treat

<table>
<thead>
<tr>
<th>Cognition</th>
<th>Change from Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bapineuzumab N=47</td>
<td>Placebo N=32</td>
</tr>
<tr>
<td>Rx difference at week 78 = 5.0</td>
<td>Difference increases over time</td>
</tr>
<tr>
<td>p=0.026</td>
<td></td>
</tr>
</tbody>
</table>

Completers

<table>
<thead>
<tr>
<th>Cognition</th>
<th>Change from Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bapineuzumab N=36</td>
<td>Placebo N=21</td>
</tr>
<tr>
<td>Rx difference at week 78 = 7.3</td>
<td>Difference increases over time</td>
</tr>
<tr>
<td>p=0.003</td>
<td></td>
</tr>
</tbody>
</table>

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

Mean (+/-SE) Change From Baseline

Weeks

- Placebo (n=432)
- Bapineuzumab 0.5 mg/kg (n=658)

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

European Federation of Neurological Societies, Stockholm – September 11, 2012

APOE ε4 carriers


Change 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

Mean (+/-SE) Change From Baseline

Weeks

- Placebo (n=493)
- Bapineuzumab 0.5 mg/kg (n=314)
- Bapineuzumab 1.0 mg/kg (n=307)

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

European Federation of Neurological Societies, Stockholm – September 11, 2012

APOE ε4 non-carriers
Post-Hoc Analyses of ApoE

- ApoE is most common marker for post hoc analysis
- Recent commentary suggested analysis with <100 subjects/arm may lead to false positive findings
  - Placebo arm should have faster progression for ApoE4+, but opposite results in some trials
- We conducted simulation studies to evaluate the extent of this problem
  - Wider range of sample sizes
  - More diverse set of trials examined

Stone et al, *Pharmacogenomics J*, 2010
Simulations of ApoE4- trials

Kennedy et al, submitted
Post-Hoc Analyses of ApoE

• Up to 25% of MCI trials and 37% of AD trials had greater rate of progression in ApoE4-
  – Majority of these were not statistically significant
• Proportion of statistically significant results favoring ApoE4- decreased with sample size
  – <5% with 125 subjects/arm for AD
• Consistent with chance findings due to small sample size
• Post-hoc analysis of ApoE with smaller samples may give incorrect target for future trials

Increased Minority Participation

- Low minority enrollment in clinical trials may reflect provider/study bias as well as participant bias
  - Exclusion of comorbidities common in minorities
  - Concerns over dropout and retention
  - Increased variability on outcome measures

- We examined this issue across our meta-database (additional support from P30AG031054, UAB RCMAR)
  - Meta-analysis of rates of medical comorbidities
  - Simulations of outcomes with African American participation ranging from 20% to 80%

Increased minority participation

Kennedy et al, submitted
Increased minority participation

- African Americans had higher rates of several disorders (primarily cardiovascular and respiratory)
  - Based on small numbers for many disorders
  - May need to be considered in medication dosing

- However, increased participation of African Americans had little impact on trial outcome
  - Small (about 5%) decreases in power with smaller samples (less than 100 per group)
  - Nearly identical power with larger samples

Kennedy et al, submitted
Enrichment using biomarkers

- Some drug development programs target putative disease biomarkers, such as ApoE4 or CSF Aβ₁-42
- ApoE4 / Aβ₁-42 could serve to enrich prevention and symptomatic clinical trials for likely responders
- Conflicting data raise questions about utility of ApoE4 as inclusion criterion in routine clinical trials
- Aβ₁-42 is highly correlated with ApoE4 status
- We examined this issue across our meta-database
  - Simulations of outcomes with ApoE4+ ranging from 0% to 100% of the sample

Schneider et al, *Alzheimer Dement*, 2010
MCI CSF $A\beta_42$ positives (--) and negatives (--)

ADAS-cog

CDR-sb
Enrichment Using Biomarkers (ApoE)

Kennedy et al., *Alzheimer Dement*, 2014

Little difference in power with ApoE4 enrichment if no differential treatment effect present.
Enrichment Using Biomarkers (Aβ)

Little difference in power with Aβ\textsubscript{1-42} enrichment compared to clinical criteria

Schneider et al, Alzheimer Dement, 2010
Enrichment using biomarkers

• ApoE4+ enrichment generally resulted in less than 3% increase in power for each simulation condition
  – Use of enrichment should depend on demonstrated differential effect by ApoE4 genotype

• Recruitment time and/or number of sites would increase
  – About 60% of AD trial participants are ApoE4+

• Similar results apply with Aβ_{1-42}
  – About 70% of AD patients will meet Aβ_{1-42} cutoffs

• Enrolling all participants (with possible stratification on ApoE4 / Aβ_{1-42}) is more effective in most trials
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. Neurology® 2015;84:1-7
Outcomes based on age

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Subject characteristics at baseline and change in ADAS-cog scores over the durations of the trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age category, y</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>No. (%)</td>
<td>total = 2,793</td>
</tr>
<tr>
<td>Education</td>
<td>high school</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>Assigned to placebo</td>
<td>2,793</td>
</tr>
<tr>
<td>APO ε4 carrier</td>
<td>1,955</td>
</tr>
<tr>
<td>MMSE</td>
<td></td>
</tr>
<tr>
<td>CDR-SB</td>
<td></td>
</tr>
<tr>
<td>ADAS-cog</td>
<td></td>
</tr>
<tr>
<td>ADAS-cog change, 6 mo</td>
<td>2,157</td>
</tr>
<tr>
<td>ADAS-cog change, 12 mo</td>
<td>1,882</td>
</tr>
<tr>
<td>ADAS-cog change, 18 mo</td>
<td>1,100</td>
</tr>
<tr>
<td>ADAS-cog change, 24 mo</td>
<td>340</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.

a Significant after correction for multiple comparisons.

Schneider, Kennedy, et al, Neurology 2015
Outcomes based on age

Figure 1  Mean and predicted outcomes scores by age category

A  Mean ADAS-cog score  
B  Predicted ADAS-cog score

C  Mean CDR-sb score  
D  Predicted CDR-sb score

E  Mean MMSE score  
F  Predicted MMSE score

Figure 2  SDs on the ADAS-cog by age category

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, 0.82, p < 0.001; vs 76-80, F = 1.61, df = 178, 0.721, p < 0.001; vs 81-85, F = 1.75, df = 178, 0.545, p < 0.001; vs 86-105, F = 2.15, df = 178, 0.216, p < 0.001). ADAS-cog = Alzheimer’s Disease Assessment Scale-cognitive subscale.

Schneider, Kennedy, et al, Neurology 2015
Differences based on age

- Differences in change on the ADAS-cog between older and younger are substantially greater than those expected between experimental drugs and placebo in current trials or those between marketed cholinesterase inhibitors and placebo.

- The clinical interpretation of change on the ADAS-cog or MMSE differs depending on age.

- Considering effects of age on rates of change is particularly important regarding clinical practice and the outcomes of trials.
Effect of **Sample Size Re-estimation** in Adaptive Clinical Trials for Alzheimer’s Disease and Mild Cognitive Impairment

Richard E. Kennedy, MD, PhD, Guoqiao Wang, MS, Gary R. Cutter, PhD, Lon S. Schneider, MD
Sample size in clinical trials

- Clinical trials to slow progression of AD have been uniformly negative
- One explanation is smaller observed effects than expected
  - Progression in AD and MCI are highly heterogenous
- Inaccurate pre-trial estimates of effect could lead to inadequately powered trials
- Many phase 2/3 trials include < 400 patients or < 160 per group
- Adaptive trial designs may offer advantages over traditional designs in such circumstances

Adaptive trial designs

- Allow deviations from pre-specified trial design based on ongoing trial monitoring
  - Sample size re-estimation
  - Allocation to treatment groups ("play the winner")
  - Dose-finding regimens

- Several potential advantages
  - Decreased trial duration
  - Smaller, required number of subjects
  - Increased power to detect treatment effects

- Cost = complexity of trial execution and analysis

- Adaptive design can’t rescue poorly planned trial!

Sample size re-estimation

• Can increase the sample size and improve power based on initial data collected during the course of a trial

• Interim data analysis is used to estimate effect size in current sample

\[ N = \left| \frac{E_0}{E} \right|^a N_0 \]

Where \( N_0 \) and \( E_0 \) are initial sample size and effect size

\( E \) is interim effect size

\( a \) is a tuning parameter

\( N \) is re-estimated sample size

• Usually incorporates early stopping (trial already successful) and futility (too many subjects required)

Methods

• Data from meta-database of 18 ADCS studies + ADNI
  – 1418 subjects with mild AD, 1192 subjects with MCI
• Simulated trials were created using resampling methodology with simulated treatment effect
  – ADAS-cog as primary outcome measure
  – Sample sizes of 50, 100, 200, 300 and 400 per group
  – AD: 12 and 18 month long trials
  – MCI: 18 and 24 month long trials
  – Dropout rates of 20% and 40% in both groups
  – Effect sizes from 0.15 to 0.25 (i.e., small to medium) for treatment effect (or slowing of decline)

Kennedy et al, *Alzheimer Dementia*, 2013
Methods: sample size re-estimation

• Evaluated two methods for re-estimation
  – Effect size (requires adjustment of $\alpha$ level)
  – Variance (does not require adjustment of $\alpha$ level)

• Single re-estimation at 6 months or at 12 months

• Adjusted sample size is based on
  – Original sample size
  – Ratio of interim effect size / variance to pre-trial effect size / variance

• Increase in sample size is not necessary if significance is achieved by interim analysis

Methods: statistical analysis

• Primary analyses: Wilcoxon test of differences in ADAS-cog endpoints between treatment and control groups

• Secondary analyses: Mixed effects linear model (random coefficients / slope model) of difference in slope between treatment and control

• Power = proportion of 1000 simulated trials per scenario with p value ≤ 0.05

• Analyses in SAS 9.3 and R 2.15.1 (with nlme package 3.1-89)

Simulation Outcomes: AD (Effect Size)

AD 18 Month Trial, Effect Size 0.25

Power

Interim Sample Size at 6 Months

Before SSR
After SSR

50 100 200 300 400

AD 18 Month Trial, Effect Size 0.25

Power

Interim Sample Size at 12 Months

Before SSR
After SSR

50 100 200 300 400

Simulation Outcomes: MCI

Supplement figure 2: Power comparison by the timing of SSR.

Table 1: Increase in sample size after SSR by initial sample sizes.

<table>
<thead>
<tr>
<th>SSR method</th>
<th>Initial Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Increase in sample sizes</td>
<td>(mean(std))</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>SSR based on variances</td>
<td>43(18)</td>
</tr>
<tr>
<td>SSR based on effect size</td>
<td>166(219)</td>
</tr>
</tbody>
</table>

Comparison between SSR at 6-months based on variances and effect sizes

Sample size re-estimation: conclusions

• Sample size re-estimation in MCI due to AD and AD clinical trials can be effective in critical circumstances
  – For MCI: SSR at 6 months with 100 subjects / group added 23% and 24% power for 18 and 24 month trials
  – For AD: SSR at 6 months with 100 subjects / group added 16% power for 18 month trials

• Depends heavily on the number of subjects accumulated for interim analysis and true treatment effect size
  – Too few: imprecise estimate of effect size/variance leading to poor prediction of sample size
  – Too many: subjects are already enrolled, so no advantage over traditional clinical trial

• Sample size re-estimation in MCI and AD must incorporate longitudinal design
  – Recruitment period is shorter than trial duration, interim analysis does not have any complete data
  – Variance of outcome measure increases over time

• Depends on type of uncertainty in pre-trial estimates
  – Effect size: both mean and variance uncertain
  – Variance: mean known, variance uncertain

Conclusions and Discussion
Summary

- Discussed a range of scenarios
  - *Post hoc* ApoE
  - Minority representation
  - Biomarkers ApoE
  - Biomarkers Aβ
  - Age
  - Adaptive design: sample size reestimation
Summary

• Subgroup selections can have negative impact if not properly implemented
  – Incorrect subgroup selection based on post-hoc analysis of ApoE4 or Aβ
  – Unnecessary exclusion of minority participants
  – Unnecessary restriction of severity (e.g., MMSE)
  – Unnecessary restrictions on age

• Simulation studies can be a valuable tool for pursuing enrichment strategies
  – Cannot replace “real-world” clinical trials
  – Can guide selection among competing strategies to increase probability of success
  – Provide a reasonable way to manage design considerations in clinical trials, better than expert opinion, conventional wisdom
References


References


END