Keck School of Medicine of USC

Alzheimer's Therapeutic Research Institute

CONFIDENTIAL DOCUMENT

Statistical Analysis Plan (SAP)

Therapeutic effects of intranasally-administered insulin (INI) in adults with amnestic mild cognitive impairment (aMCI) or probable mild Alzheimer's disease (AD)

June 22, 2018
Author: Chung-Kai Sun, MS
From the ATRI Biostatistics Team

Contents

1 Introduction

This Statistical Analysis Plan (SAP) describes the planned analyses and reporting for the study *Therapeutic effects of intranasally-administered insulin (INI) in adults with amnestic mild cognitive impairment (aMCI) or probable mild Alzheimer's disease (AD)*.

The planned analyses identified in this SAP, will be included in future manuscripts. Exploratory analyses not necessarily identified in this SAP may be performed to support planned analyses. Any post-hoc or unplanned analyses not specified in this SAP will be clearly identified as such in the manuscripts for publication.

The following documents were reviewed when preparing this SAP:

- INI Clinical Research Protocol (INI_Protocol_v5.0_20161205.pdf)
- Case report forms (CRFs) for INI
- ATRI Statistical Analysis Plan SOP (ATRI SOP BI-002 Statistical Analysis Plan v1.0 20150908.pdf)

Readers of this SAP should read the Clinical Research Protocol for details on the conduct of this study and the operational aspects of clinical assessments and timing for completing a patient in this study.

2 Abbreviations

ABBREVIATION	DEFINITION
ADAS-Cog	Alzheimer's Disease Assessment Scale-Cognitive
ADCS-ADL MCI	Alzheimer's Disease Cooperative Study-Activities of Daily Living - MCI
AE	Adverse Event
AMCI	Amnestic Mild Cognitive Impairment
ANCOVA	Analysis of Covariance
APOE/APOE4	Apolipoprotein (APOE) Epsilon 4 (APOE4)
ATRI	Alzheimer's Therapeutic Research Institute
CDR-SB	Clinical Dementia Rating – Sum Of Boxes
CMRgl	Cerebral Metabolic Rate for glucose
CRF	Case Report Form
CSF	Cerebrospinal Fluid
DSMB	Data and Safety Monitoring Board
FDG	Fluoro Deoxy Glucose
FCSRT	Free And Cued Selective Reminding Test
mITT	modified Intention-To-Treat
MMSE	Mini- Mental State Examination
MRI	Magnetic Resonance Imaging
NPI	Neuropsychiatric Inventory
PBMC	Peripheral Blood Mononuclear Cell
PET	Positron-Emission Tomography
p-tau 181	phospho-tau 181
PK	Pharmacokinetic
SOP	Standard Operating Procedure
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SPM	Statistical Parametric Mapping
sROI	statistical region of interest

3 Study Design/Summary

This section is from the Clinical Research Protocol, INI_Protocol_v5.0_20161205.pdf

3.1 Overview

The study will consist of a multisite, randomized, double-blind trial comparing the effects of INI (20 IU bid for total daily dose of 40 IU) and placebo for 12 months, followed by a 6-month open-label period in which all participants will receive INI. Participants with aMCI or probable mild AD (n=240-300) will be enrolled. The primary outcome measure will consist of the ADAS-Cog12. Secondary measures will include a memory composite, an executive function test, CDR Sum of Boxes (CDR-SB), and the ADCS-ADL MCI. The ADAS-Cog12 will be administered at 3 month intervals to optimize imputation of missing data. Other cognitive and functional measures will be administered at 6 month intervals. MRI measures of entorhinal cortex and hippocampal atrophy will be obtained at screening and 12 months. CSF and plasma biomarkers, as well as PBMCs and APOE-ϵ4 allele carriage will also be assessed.

3.2 Study Aims

3.2.1 Primary Aim

To test the hypothesis that 12 months of treatment with INI (compared to placebo) in adults with aMCI and probable mild AD will improve performance on a global measure of cognition (ADAS- Cog12).

3.2.2 Secondary Aims

1. To test the hypothesis that 12 months of treatment with INI (compared to placebo) in adults with aMCI or probable mild AD will improve performance on a memory composite (Story Recall and FCSRT) and on daily function in adults with aMCI and probable mild AD.
2. To test the hypothesis that INI treatment reduces the rate of hippocampal and entorhinal atrophy as measured by MRI, and conduct exploratory analyses of other brain regions.
3. To test the hypothesis that INI will favorably alter CSF Aβ and the CSF Aβ/tau ratio, and will modulate inflammatory markers.
4. To examine whether baseline AD biomarker profile, APOE4 allele carriage and gender predict treatment response.
5. To determine whether further improvement occurs after 18 months of treatment.

3.3 Study Population

A total of 240-300 adults diagnosed with aMCI or probable mild AD will be enrolled in this trial. We expect to enroll no more than 50% of participants with probable mild AD and no more than 60% participants with aMCI diagnosis. To determine eligibility, all participants will undergo cognitive assessment, physical and neurological examination, ECG, clinical/safety laboratory assessment, and interviews of the participant and study partner.

3.4 Power and Sample Size Determination

Power calculations to estimate the sample sizes required to detect a difference between the rate of change in two groups were based on the two-sample t-test using the formula in Statistical Methods in Medical Research (Armitage and Berry 1994). Sample sizes per arm were estimated using ADNI 12-month change scores for the ADAS-Cog12 for the combined MCI and AD population, targeting power=.80, Type 1 error level=0.05, SD=5.72. Table 1 shows the range of estimates needed to detect absolute change score differences between groups ranging from 1.6-2.6 for dropout rates between 0.15-0.25. These estimates were derived with 6 month intervals between testing, and thus more frequent administration of the ADAS-Cog12 in

this trial should provide even greater power. Similarly, a 4-month pilot trial (Craft, Baker et al. 2012) dropout rate was <7%, and although a higher dropout rate is expected with a longer, multi-site trial, the excellent safety profile and incentive of the open-label extension at the end of the trial should maximize retention. The addition of a delayed recall score should increase sensitivity for aMCI participants. For MRI endpoints, sample sizes needed to detect a 25% slowing in mean rate of decline in the entorhinal cortex over a 12-month period ranged from 70 for AD to 280 for aMCI (Holland, Brewer et al. 2009). Given that our n=240, and the sample will be composed of both aMCI and AD, we should have adequate power.

Table 1: Sample Size Estimate

	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7
Dropout rate .15	431	382	341	306	276	250	228	209	192	177	163
Dropout rate .175	436	387	345	309	279	253	231	211	194	179	165
Dropout rate .20	441	391	349	313	282	256	233	214	196	181	167
Dropout rate .225	446	395	353	317	286	259	236	216	198	183	169
Dropout rate .25	451	400	357	320	289	262	239	218	201	185	171

3.5 Treatment Allocation

The two interventions are as follows:

- Group 1: active, 20 IU Humulin RU-100 bid for a total daily dose of 40 IU
- Group 2: placebo, matching placebo (sterile diluent)

Initially, eligible participants will be randomized on a 1:1 schedule to receive placebo or 20 IU bid of INI using the following randomization strategy:

- MMSE range (<=25 vs. >25)
- Site
- ApoE-e4 carrier status (yes or no)
- Gender (weight: 2)
- Age (<=70 vs. >70)

will be weighted in a covariate-adaptive randomization strategy to achieve optimal balance between treatment arms on important factors that may influence treatment outcome. After 12 months of double-blinded treatment with placebo or 20 IU bid of insulin, all participants will receive 20 IU INI bid in a 6-month, open-label extension.

4 Study Outcome Variables

4.1 Primary Outcome

Alzheimer's Disease Assessment Scale-Cognition 12 (ADAS-Cog12)

4.2 Secondary Outcomes

- CDR-SB
- memory composite (Immediate Paragraph Recall, Delayed Paragraph Recall and FCSRT)
- ADCS-ADL MCI

4.3 Safety Outcomes

- Adverse Events (AEs)
- Serious Adverse Events (SAE)
- Deaths
- Hospitalizations
- Vital Signs
- Standard Laboratory Tests

4.4 Biomarker and Imaging Outcomes

- MRI volumes (hippocampus, entorhinal cortex)
- Aβ42, CSF total tau, p-tau 181, Aβ42/tau, p-tau 181/Aβ42, Aβ42/Aβ40 ratio.

5 Sequence of Planned Analysis

5.1 Interim Analysis

AEs and SAEs are being monitored on a regular basis by a Data and Safety Monitoring Board (DSMB). There are no planned interim analysis being conducted for this study, but the DSMB may modify this during ongoing safety monitoring.

5.2 Final Analyses and Reporting

All final planned analyses identified in the research plan and in this SAP will be performed in two phases:

- blinded phase:
  After the last participant has completed all assessments scheduled for blinded phase (12 months) of the study and the database for this phase has been cleaned and locked. This analysis will only include visits up to and including 12 months.
- follow-up phase:
  After the last participant has completed assessments scheduled for the study (18 months) and the complete database has been cleaned and locked. This analysis will only include visits up to and including 18 months.

Any post-hoc exploratory analyses performed to provide support for planned analyses but not identified in this SAP will be documented and clearly identified as unplanned analysis in the manuscript.

6 Statistical Methods

6.1 Analysis Populations

- Intent-to-treat (ITT) population: All eligible individuals who are randomized.
- Modified ITT (mITT) population: Subjects who are randomized and have at least one post-baseline assessment.
- Compliance population: Subjects complete the study on treatment with $\geq$ 80% treatment compliance based on study diary.

6.2 Multiple-testing Strategy

We propose a multiple-testing strategy using a basic serial gatekeeping procedure to maintain over-all experiment-wise Type I error at α = 0.05 (two-sided). The four hypotheses regarding (i) ADAS-Cog12, (ii) CDR-SB, (iii) memory composite (Immediate Paragraph Recall, Delayed Paragraph Recall and FCSRT), (iv) ADCS-ADL MCI will be tested in order using a basic serial testing strategy. If, for example, $p_{ADAS} \geq 0.05$, no subsequent hypotheses will be declared statistically significant. Otherwise if $p_{ADAS} < 0.05$, the CDR-SB hypothesis is tested by comparing p_{CDR} to 0.05. This strategy continues until a p-value is ≥ 0.05, or the ADCS-ADL MCI hypothesis is declared statistically significant with $p_{ADL} < 0.05$.

Regardless of the outcome of the multiple testing strategy, p-values and 95% confidence intervals will be reported for all planned analyses.

6.3 Incomplete Follow-up/ Missing Data

- If any of the individual items for the clinical outcomes are missing, every effort will be made to obtain the score for the missing item or items.
- ADAS-Cog12 and CDRSB Scoring: Total scores for assessments with missing item scores will be imputed using a proration strategy. If the maximum score with the non-missing items represents $\geq$ 70% of maximum possible total score, the total score will be imputed. The score observed with the non-missing items will be prorated to the score with all items, as follows:

$$\text{Imputed score} = \text{Maximum score with all items} \times \frac{\text{Score observed with non-missing items}}{\text{Maximum score with non-missing items}}$$

  The imputed score will be rounded to the nearest integer. If the maximum score with the non-missing items represents < 70% of the maximum score with all items, the total score for the assessment at that visit will not be imputed and will be considered missing.
- ADCS-ADL MCI Scoring: The response of "Don't Know" for any individual question will be scored as 0 (the same score as response of "no").
- FCSRT scoring: Sum of total free and total cued recall score. The score is missing if cued recall is missing.
- Memory Composite (Immediate Paragraph Recall, Delayed Paragraph Recall and FCSRT): Each of the three component scores is divided by the baseline sample standard deviation of that component to form standardized Z scores. Three Z scores are summed to form the memory composite. If any of Immediate Paragraph Recall, Delayed Paragraph Recall and FCSRT is missing, the memory composite is missing.

6.4 Covariates

Covariates to be included are specified for each analysis, when applicable.

6.5 Data Management and Analysis Software

Most data manipulation, tables, figures, listings and analyses will be created using R programs and performed using R statistical software (http://www.r-project.org). R version used will be specified in the analysis reports.

7 Statistical Analysis

7.1 Study Flow

A CONSORT style flow diagram will illustrate patient progression through the trial from initial screening for eligibility to completion of the final primary outcome assessment. Number (percentage) will be included for subjects screened, screen failed (with reasons), randomized, study discontinued (with reasons), treatment discontinued (with reasons), analyzed, excluded from analysis (with reasons). The diagram will be separated by treatment groups starting from randomization.

7.2 Visit Windows

As subjects do not always adhere to the protocol visit schedule, the following rules are applied to assign actual visits to analysis visits. Listed below are the visit windows and the target days for each visit. The reference day is Study Day 1. If a subject has 2 or more actual visits in one visit window, the visit closest to the targetday will be used as the protocol visit for that visit window. Listed below (Table 2) are the visit windows and the target daysfor each visit defined in the protocol.

Table 2: Visit Windows for Analysis

Visit Code	Time Interval (days)	Target (days)
bl	(-30, 1)	1
m03	(2, 136)	91
m06	(137, 225)	181
m09	(226, 315)	271
m12	(316, 379)	366
m15	(380, 502)	458
m18	(503, 554)	541

7.3 Patient characteristics and baseline variables

Baseline demographics and clinical characteristics will be summarized by frequencies and percentages for categorical variables, continuous variables will be summarized by mean and standard deviation as well as quartiles. Group comparisons will be performed using Student t-tests or Wilcoxon Rank-Sum tests, as appropriate, for continuous variables. Fisher's Exact test will be used for categorical data. The following baseline demographic and clinical characteristics will be presented:

- Demographics
  - Age
  - Gender
  - Race
  - Ethnicity
  - Language
  - Marital status
  - Retirement
  - Living situation
  - Education
  - ApoE-e4 carrier status (yes or no)
- Clinical
  - ADAS-Cog12
  - Memory Composite (Story Recall, FCSRT)
  - MMSE
  - NPI
  - ADCS-ADL MCI
  - CDR-SB
  - Modified Hachinski
  - Logical Memory II Subscale (Wechsler Memory Scale)
  - Trail-making Test

7.4 Analysis of Primary outcome

Individuals in the active arm with Impel device will demonstrate slower rate of decline on ADAS-Cog12 in comparison to the placebo arm. Mixed Model repeated measures analysis of variance (MMRM) will be used to assess the difference between treatment groups in change from baseline for performance on the ADAS-Cog12. Participants will be included in the model if they have both a baseline and at least one post-baseline ADAS-Cog12. Change from baseline at each post-baseline visit for the ADAS-Cog12 will be the dependent variable. We will use a heterogeneous autoregressive AR(1) covariance structure. If the heterogeneous AR(1) covariance structure results in a lack of convergence, the following covariance structures will be used in sequence: heterogeneous compound symmetry covariance, homogeneous autoregressive AR(1) covariance, homogeneous compound symmetry covariance. Fixed effects in the model will include the following terms: (i) baseline ADAS-Cog12 score, (ii) treatment, (iii) visit, (iv) treatment-by-visit interaction, and (v) APOE-ϵ4 genotype (Y/N) (vi) gender (vii) baseline MMSE score (viii) baseline age. Visits (baseline, month 3, month 6, month 9 and month 12) will be treated as a categorical variable. The null hypothesis is that the treatment difference between active and placebo with Impel device for the ADAS-Cog12 at month 12 is equal to zero.

7.5 Analysis of Secondary Outcomes

Secondary Aims I

Individuals in the active arm with Impel device will demonstrate slower rate of decline on secondary outcomes including CDR-SB, memory composite (Story Recall and FCSRT), and ADCS-ADL MCI in comparison to the placebo arm. Mixed Model repeated measures (MMRM) models similar to the primary analysis will be used.

Secondary Aims V

To determine whether further improvement occurs after 18 months of treatment.

- The analysis will be the same as the primary analysis with data up to the month 18. The null hypothesis is that the treatment difference between active and placebo with Impel device for the ADAS-Cog12 at month 18 is equal to zero.
- Delayed-start analysis is to compare Δ_2, the treatment differences between active and placebo at the month 18 (delayed-start period) with Δ_1 that at the end of the placebo-controlled period with noninferiority margin as 50% of Δ_1 [1, 2].

7.6 Analysis of Safety Outcomes

To assess the safety and tolerability of treatment with INI over a 12 month compared to placebo . Tolerability and safety will be determined by serious adverse events, adverse events and treatment discontinuation rate of INI/placebo as determined by the participant and the physician. AEs and SAEs will be recorded and categorized by body system, event type, attribution, frequency, severity, and course. Differences for rates of AEs, SAEs or other abnormalities between groups will be assessed using Fisher's Exact tests. Changes in vital signs, and laboratory tests between groups will be assessed using a MMRM approach analogous to the primary analysis.

7.7 Analysis of Biomarker and Imaging Outcomes

Secondary Aims II and III

Individuals in the INI arm will demonstrate slower rate of decline on MRI volumes and CSF biomarkers in comparison to the placebo arm.
Analysis Plan: The rate of change in MRI volumes and CSF biomarkers will be compared between the two treatment groups using a linear mixed-effect model with MRI volumes and CSF biomarkers at each time point (including baseline) as the outcome variable. The model will include fixed effects for (i) time (as a continuous variable), (ii) time-by-treatment interaction, (iii) APOE-4 genotype (Y/N), (iv) gender (v) baseline MMSE score (vi) baseline age; and participant-specific random intercepts. The model will constrain the two treatment groups to have the same mean at baseline. The hypothesis will be tested using the p-value corresponding to the mean group difference in rate of change in MRI volumes and CSF biomarkers.

7.8 Exploratory Subgroup Analyses

Secondary Aims IV

All analyses described above will be repeated on the following subgroups to explore the differential treatment effects:

- Compliance population
- Amyloid Positive (< 600pg/ml) vs Negative (>= 600 pg/ml) based on Aβ42 MSD platform.
- Number of ApoE-e4 alleles (0, 1, 2).
- ApoE-e4 carrier status (Positive vs Negative).
- Gender Male vs Female.
- Screening MMSE median split.
- Clinical Diagnosis MCI vs AD.

Supporting analysis

Two sets of analyses will be conducted as the part of the primary, secondary and subgroup analysis.

- individuals randomized to Kurve device alone
- all individuals in the study with both Impel and Kurve device. Device and device by treatment interaction are included as fixed effects in the model. If the device by treatment interaction is not significant ($p > 0.05$), only device is included as a fixed effect.

References

[1] Liu-Seifert, Hong, Scott W. Andersen, Ilya Lipkovich, Karen C. Holdridge, and Eric Siemers. 2015. "A Novel Approach to Delayed-Start Analyses for Demonstrating Disease-Modifying Effects in Alzheimer's Disease." Edited by Rochelle E. Tractenberg. PLOS ONE 10 (3): e0119632. doi:10.1371/journal.pone.0119632.

[2] Liu-Seifert, Hong, Eric Siemers, Karen C. Holdridge, Scott W. Andersen, Ilya Lipkovich, Christopher Carlson, Gopalan Sethuraman, Sharon Hoog, Roza Hayduk, Rachelle Doody, Paul Aisen. 2015. "Delayed-Start Analysis: Mild Alzheimer's Disease Patients in Solanezumab Trials, 3.5 Years." Alzheimer's Dementia: Translational Research Clinical Interventions 1 (2): 111–21. doi:10.1016/j.trci.2015.06.006.

8 Approvals

Approved by	Title	Signature	Date
Suzanne Craft, PhD	Principle Investigator		
Paul Aisen, MD	Director, ATRI		
Rema Raman, PhD	Lead Biostatistician		

9 Revision History

Version		
Number	**Date**	**Revision Summary**
v1.0	20180205	Original
v1.1	20180607	Add a subgroup and correct typos