ADDENDUM TO STATISTICAL ANALYSIS PLAN

Title:	Statistical Analytical Plan for RVT-101-3001: A Phase 3, double blind, randomized study of RVT-101 versus placebo when added to existing stable donepezil treatment in subjects with mild to moderate Alzheimer's disease
Sponsor	Axovant Sciences Ltd.
Compound Name:	Intepirdine
Protocol Number	RVT-101-3001
Indication	Treatment of mild to moderate Alzheimer's disease in patients on stable therapy with donepezil
Development Phase	3
IND #	78,094
Version	Final SAP, Version: August 9, 2017 Addendum to SAP, Version: September 19, 2017
Axovant Sciences Study Director	[redacted]

This study is being conducted in compliance with good clinical practice, including the archiving of essential documents.

SIGNATURE PAGE

[redacted] Sep 19, 2017

Signature Date

[redacted] Sep 19, 2017

Signature Date

[redacted] Sep 19, 2017

Signature Date

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

Axovant Sciences, Inc., provided a request (June 12, 2017) to obtain European Medicines Agency (EMA) guidance regarding the ongoing Phase 3 (RVT-101-3001) study in patients with Alzheimer's disease (AD). The Phase 3 clinical trial will conclude in late September 2017. The purpose of the request was to reach agreement with EMA on the statistical and data analysis plans prior to study unblinding.

The Committee for Medicinal Products for Human Use (CHMP) has provided a response (Scientific Advice) to that June 12, 2017 request. That Scientific Advice, dated September 14, 2017, included specific input regarding analyses of the primary efficacy endpoints as well as other aspects of the statistical assessment.

The purpose of this addendum to the final statistical analysis plan (SAP), dated August 9, 2017, is to describe the ***additional*** analyses that will be performed in response to the Scientific Advice received from the EMA.

2. SUMMARY OF EMA SCIENTIFIC ADVICE

The focus of this addendum to the SAP is to respond to CHMP's response to Question 1 of the June 12, 2017 request. This section provides a summary of the Scientific Advice to that question.

Question 1

Does EMA concur with the statistical methodological approach to analyze Study RVT-101-3001? Specifically, our plan for the analysis of the primary endpoints, the hierarchical analysis of the key secondary endpoints, and the included plans for responder analyses?

{Following a summary of the study design, CHMP advised the following.}

The Applicant presents analysis methods for the co-primary and secondary endpoints that are comprehensively described in the protocol and the Statistical Analysis Plan. However, the target of estimation or scientific question addressed with the trial is not appropriately defined and discussed.

The Applicant proposes that inferences from treatment differences for the changes from baseline derived from a mixed model for repeated measures (MMRM) at Week 24 in the defined ITT population are appropriate, using all observed data for patients until the primary analysis time point and implicit imputations by the MMRM analysis for patients that withdraw from the study before the primary analysis time point. While this target of estimation would address a specific research question and estimate a treatment effect as if all patients adhered to treatment until the primary analysis time point, CHMP is of the opinion that this hypothetical target of estimation would provide a too optimistic estimate of an attainable treatment effect if a relevant number of patients withdraws from the study and the Missing At Random assumption would not be fulfilled. A considerable number of drop-outs has to be expected, withdrawal patterns could be different between treatment arms and the assumption of data only missing at random is not reasonable.

A more appropriate primary target of estimation could make use of information observed after a patient withdrew from randomized treatment. Patients starting a new medication or have their background medication changed after withdrawal should also be followed and included in the analysis, as a comparison of treatment policies would also be of interest. Additionally, multiple imputation methods would be an acceptable option if no observed data after treatment withdrawal are available. For imputation of missing data, information from patients receiving placebo should be considered to avoid a potential overestimation of the treatment effect. Reason for drop-out could be used and imputations for patients withdrawing due to lack of efficacy or adverse events could be treated differently.

An appropriate primary analysis should be accompanied with sensitivity analysis to test assumptions on properties of methods and data. It is noted that the set of analyses defined in the SAP partly address a different target of estimation (alternative analysis methods termed "Sensitivity Analyses of the Primary Efficacy Results" in section 8.3.2.2. on observed cases). Single imputation methods as in "Sensitivity Analysis from Imputation of Missing Efficacy Data for Primary Efficacy Endpoints" (section 8.3.2.4.) would not be acceptable and the outlined analysis that makes use of multiple imputation methods uses the complete trial population to

make covariate based predictions for imputation. While multiple imputation methods are considered appropriate, the dataset used to generate imputations would have to be justified.

From the EU Marketing Authorization Application perspective, an appropriate primary analysis demonstrating a relevant treatment effect will have to be defined and pre-specified in the Statistical Analysis Plan before unblinding the trial.

It is further noted that study 3001 uses a single-blind placebo run-in period after screening to select patients for randomization who do not improve by more than 3 points on the MMSE score after placebo treatment. Consequently, the target of estimation estimates a treatment effect in a selected subgroup of all patients eligible for treatment with the new product and it will have to be justified how the results from this specific trial design can be generalized for the whole population.

With regard to the definition of the ITT population, it is considered that the Intention-To-Treat population, as the definition already suggests, should consist of all patients randomized and the Applicant will have to justify that their definition (all randomized who took medication with one post-baseline assessment) will lead to the same conclusions on treatment effects.

Use of the proposed co-primary endpoints is acceptable and the strategy to control the type I error for the primary (both must be successful) and secondary endpoints with a hierarchical approach is considered adequate.

Methods for the (key) secondary and other secondary endpoints are acceptable. The defined responder analyses should not only be based on observed cases with assessments at week 24. Imputations should be used and an approach with missing data imputed as failures should be included. A responder criterion of 4 on the ADAS-Cog 11 should be generally implemented in addition to improvement of at least 3 points.

The assumptions for the sample size estimations for the co-primary endpoints and use of the treatment difference of 1.6 points on the ADAS-cog 11 and 2 points for the ADCS-ADL scales between active treatment and placebo are comprehensible appear technically adequate.

Sub-group analysis based on gender, region, race, MMSE severity, Donepezil dose is endorsed.

2.1. Summary of Additions to Statistical Analysis Plan in Response to Scientific Advice

The following is a summary of the additions to the SAP as well as explanation of the estimand and assumptions on missing data to the SAP.

1. A description of the estimand (based on ICH E9 R1 Addendum) is provided. A detailed description of intercurrent events is included (see Section 3.2.3).
2. A description of the estimated proportion of missing data due to overall dropout, due to missing at random and missing not at random reasons is provided (see Section 3.2.3.1).

In response to other comments in the CHMP scientific advice, the following additional analyses have been added.

1. Addition of a new imputation of missing data, using missing imputation methods as suggested by CHMP, is included. This new analysis will include:

a. The data will be imputed assuming missing not at random data. This imputation method will use information from patients receiving placebo (but not from patients receiving Intepirdine) in order to avoid a potential overestimation of the treatment effect. See Section 3.2.3.2.

b. For patients who discontinued due to adverse event or due to lack of efficacy, imputation is treated differently. Specifically, a scalar adjustment to the imputed values will be performed in order to assign an unfavourable outcome. See Section 3.2.3.3.

2. These imputed datasets will be used to run additional analyses using the MMRM and Wilcoxon tests for the ADAS-Cog and ADCS-ADL endpoints at Week 24. The analysis of these endpoints (and the responder outcomes) will ***also*** be performed on the Randomized population (in addition to the already-planned ITT population), and a comparison of the outcomes between the ITT population and Randomized population outcomes will be performed. See Section 4, Section 5.

Further, for clarification, the responder analysis as was originally in the SAP is re-iterated to address the comment from CHMP regarding treating missing data as failures.

1. The responder analysis (as is included in the final SAP) is confirmed to be performed with missing data imputed as failure (as well as being performed with observed cases only). See Section 5.2.

The following sections describe the additions to the SAP per the scientific advice.

3. TARGET OF ESTIMATION

CHMP provided the following advice on this topic:

> The Applicant presents analysis methods for the co-primary and secondary endpoints that are comprehensively described in the protocol and the Statistical Analysis Plan. *However, the target of estimation or scientific question addressed with the trial is not appropriately defined and discussed.*
>
> The Applicant proposes that inferences from treatment differences for the changes from baseline derived from a mixed model for repeated measures (MMRM) at Week 24 in the defined ITT population are appropriate, using all observed data for patients until the primary analysis time point and implicit imputations by the MMRM analysis for patients that withdraw from the study before the primary analysis time point. While this target of estimation would address a specific research question and estimate a treatment effect as if all patients adhered to treatment until the primary analysis time point, *CHMP is of the opinion that this hypothetical target of estimation would provide a too optimistic estimate of an attainable treatment effect if a relevant number of patients withdraws from the study and the Missing At Random assumption would not be fulfilled. A considerable number of drop-outs has to be expected, withdrawal patterns could be different between treatment arms and the assumption of data only missing at random is not reasonable.*
>
> A more appropriate primary target of estimation could make use of information observed after a patient withdrew from randomized treatment. Patients starting a new medication or have their background medication changed after withdrawal should also be followed and included in the analysis, as a comparison of treatment policies would also be of interest. Additionally, multiple imputation methods would be an acceptable option if no observed data after treatment withdrawal are available. *For imputation of missing data, information from patients receiving placebo should be considered to avoid a potential overestimation of the treatment effect. Reason for drop-out could be used and imputations for patients withdrawing due to lack of efficacy or adverse events could be treated differently.*

3.1. ICH E9 Addendum

The "ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials" (ICH E9 Addendum), see Appendix 1, is currently open for consultation and comments. However, the principles outlined in that document form the basis for our response to that aspect of the CHMP advice.

The ICH E9 Addendum indicates that an estimand defines in detail what needs to be estimated to address a specific scientific question of interest. A description of an estimand includes four attributes:

A. the population, that is, the patients targeted by the scientific question.

B. the variable (or endpoint), to be obtained for each patient, that is required to address the scientific question.

C. the specification of how to account for intercurrent events to reflect the scientific question of interest.

D. the population-level summary for the variable which provides, as required, a basis for a comparison between treatment conditions.

Together these attributes describe the estimand, defining the treatment effect of interest.

3.2. Estimand for Protocol RVT-101-3001

Consistent with the ICH E9 Addendum, definition of the attributes of the estimand (target of estimation) is provided in this section.

3.2.1. Population targeted by the scientific question

The population targeted by the scientific question is defined via the inclusion and exclusion criteria as part of the original protocol. Patients may be male or female, and must have a clinical diagnosis of AD in accordance with the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for AD.

A key aspect of eligibility is that subjects must have an MMSE score 12 to 24 inclusive at Screening and a Baseline MMSE score 10 to 26 inclusive. The difference between the Screening and Baseline MMSE score must be less than or equal to 3 points.

3.2.2. Variable (or endpoint), to be obtained for each patient, that is required to address the scientific question

The co-primary endpoints to be obtained for each patient in this study to address the scientific question are ADAS-Cog-11 score at Week 24 and ADCS-ADL score at Week 24.

3.2.3. The specification of how to account for intercurrent events to reflect the scientific question of interest

Premature discontinuation from the study is the primary potential intercurrent event that could occur. The following are the possible reasons for premature discontinuation that could be related to study treatment and that require an adjustment to the imputation strategy:

- Adverse Event
- Lack of efficacy

3.2.3.1. Basis for Use of the Multiple Imputation Strategy

The patterns of missing data for each primary efficacy endpoint will be presented, with the number (%) of patients in each treatment group with each pattern of missingness, as demonstrated in Table 1. These data will be examined in order to determine if there was any general difference in the pattern of missing data between the treatment groups, as well as to explore the assumption of missingness at random.

Table 1: Examples of Patterns of Missingness (Study RVT-101-3001)

Pattern	Treatment Group	Number (%) of Patients	Baseline	Week 3	Week 6	Week 12	Week 18	Week 24
1	Intepirdine	xx (xx.x%)	X	X	X	X	X	.
	Placebo	xx (xx.x%)	X	X	X	X	X	.

Pattern	Treatment Group	Number (%) of Patients	Baseline	Week 3	Week 6	Week 12	Week 18	Week 24
2	Intepirdine	xx (xx.x%)	X	X	X	X	.	.
	Placebo	xx (xx.x%)	X	X	X	X	.	.
Etc.								

In addition, a comparison of dropout due to AE will be performed between the treatment groups, along with a graphical display of the time to discontinuation due to AE by treatment group.

A concern indicated by CHMP was relating to considerable dropout. However, the actual dropout in this study is estimated to be no higher than 11%. An examination of the patterns of missing data for the co-primary endpoints demonstrates that missing data are rare, and that collection of the endpoints was, in general, thorough, and that over 87% of patients had at least one of their Week 24 co-primary efficacy endpoints collected. Further, dropout (discontinuation) due to adverse event is estimated to be less than 4%, and dropout due to lack of efficacy less than 1%. Therefore, the overwhelming majority (~95%) of Week 24 values will either be collected or will be attributed to MAR. These outcomes reflect a high level of study oversight and management of patient visits as well as patient retention to minimize the effects of missing values on the primary targets of estimation. Further, the methods proposed for the primary analysis (MMRM) are recognized as robust under these assumptions.

Data collection beyond premature (early) discontinuation was not performed in this study, while the number of patients with missing values at Week 24 is low, patients discontinuing early did not have their later efficacy data collected. Relating to this, CHMP indicated:

- "A more appropriate primary target of estimation could make use of information observed after a patient withdrew from randomized treatment. Patients starting a new medication or have their background medication changed after withdrawal should also be followed and included in the analysis, as a comparison of treatment policies would also be of interest."

As collection of post-drop data was not performed, the alternative option to address these missing values will be pursued. To address this situation, CHMP suggested:

- "Additionally, multiple imputation methods would be an acceptable option if no observed data after treatment withdrawal are available. For imputation of missing data, information from patients receiving placebo should be considered to avoid a potential overestimation of the treatment effect. Reason for drop-out could be used and imputations for patients withdrawing due to lack of efficacy or adverse events could be treated differently."

Therefore, to account for the intercurrent events described above, an additional analysis via multiple imputation of those missing data for those patients experiencing certain intercurrent events (discontinuation due to AE or due to lack of efficacy) is planned. The ICH E9 Addendum indicates:

- "…for a numerical variable, experiencing an intercurrent event might be ascribed an extreme unfavourable value and a suitable summary measure selected."

Details of the details of this additional multiple imputation strategy are provided in the following section. This analysis will be performed IN ADDITION to the final SAP.

3.2.3.2. Description of Additional Multiple Imputation Strategy

Missing on-treatment data for the primary endpoints will be imputed utilizing multiple imputation methods. Multiple imputation provides a useful strategy for analyzing data sets with missing values. Instead of filling in a single value for each missing value, Rubin's [Rubin 1976, Rubin 1987] multiple imputation strategy replaces each missing value with a set of plausible values that represent the uncertainty about the correct value to impute. Once a monotonic missing data pattern is established, the imputation method will specifically impute data at each time point ***based on the distribution data only from the placebo-treated patients. Thus, all Week 24 values will be imputed using only the placebo-based distribution.***

To examine the primary endpoint with a multiple imputation method that does not require the MAR assumption, a pattern-mixture approach will be utilized for the imputation. The first step will be to apply a Markov chain Monte Carlo (MCMC) method [Schafer 1997] that assumes multivariate normality will be used to impute all missing values to make the imputed (resulting) data sets have strictly monotone missing patterns. The resulting monotone missing pattern will then, in a second imputation step, be used to impute the remaining missing values; specifically, a regression-based pattern-mixture method for continuous variables will be applied.

The SAS software system will be used to perform this imputation. Because the imputation of missing data is a key aspect to the analysis of the data, explicit details regarding this imputation are provided via sample SAS code that is intended to demonstrate the application of these strategies. Variable definitions are:

- TRT=treatment group (1=Intepirdine, 0=Placebo)
- Eff1 is the first time-point for the variable score
- Eff2 is the second time-point for the variable score
- Eff_last is the last time point for the variable score. Additional variable scores (between Eff1 and Eff_last) would be included in this model according to the time points for collection.

The first step will be to impute partially to obtain a monotone missing data pattern. ***This imputation will be done independent of treatment.***

```
proc mi data=DATAIN out=DATAIN_MONO nimpute=100 seed=123;
    var Eff1 Eff2 …. Eff_last;
    mcmc chain=multiple impute=monotone;
    run;
```

The second step will be to impute the remaining (monotone) missing data that is MAR for each of the 100 imputed datasets from the first step. ***This step will assume that the distribution of missing values are based on the distribution of placebo outcomes ONLY*** (and thus reduces the risk of overestimating the treatment effect).

Ratitch and O'Kelly (Ratitch 2011) describe an implementation of the pattern-mixture model approach that uses a control-based pattern imputation. That is, an imputation model for the missing observations in the treatment group is constructed not from the observed data in the treatment group but rather from the observed data in the control group. This model is also the imputation model that is used to impute missing observations in the control group.

Following the imputation of the data to create a monotone missing pattern, missing values for the earliest missing value (visitnum1) are imputed BY VISIT. The first visit to be imputed would be the Week 3 time-point, then the Week 6, Week 12, Week 18, and finally Week 24. The Week 24 time-point is the key timepoint of interest.

The process is repeated sequentially, ie, by visit, until the monotone missing data pattern is completely filled in (imputed), and thus there may be up to 5 calls to the PROC MI procedure to generate the final imputed dataset.

The following sample SAS code demonstrates the general methodology for imputing the first visit (Week 3), second visit (Week 6) data, and third visit (Week 12). Note that the output dataset from the first imputation is used as the input dataset for the second imputation, and the output dataset from the second imputation is used as the input dataset for the third imputation. Note that minor alterations to the SAS code may be performed.

```
  **First procedure imputes missing values for the first visit (Week 3);
  proc mi data=DATAIN_MONO out=DATAREG1 seed=465 nimpute=1;
      class trt;
      monotone regression(eff1);
      mnar model (eff1 / modelobs=(trt='0')); **indicates only observations
      where treatment was "placebo" are used to derive the imputation model
      for the outcomes;
      var eff1;
      run;
  ** Second procedure imputes missing values for the second visit (Week 6)
  proc mi data=DATAREG1 out=DATAREG2 seed=465 nimpute=1;
      class trt;
      monotone regression(eff2);
      mnar model (eff1 eff2 / modelobs=(trt='0')); **indicates only
      observations where treatment was "placebo" are used to derive the
      imputation model for the outcomes;
      var eff1 eff2;
      run;
** Third procedure imputes missing values for the Third visit (Week 12)
  proc mi data=DATAREG2 out=DATAREG3 seed=465 nimpute=1;
      class trt;
      monotone regression(eff3);
      mnar model (eff1 eff2 eff3 / modelobs=(trt='0')); **indicates only
      observations where treatment was "placebo" are used to derive the
      imputation model for the outcomes;
      var eff1 eff2 eff3;
      run;
```

This process is continued for all visits through Week 24.

Note that the regressions do not include the subset of observations from the Intepirdine treatment arm. **The imputed data at each time point will be based on the distribution data only from the placebo-treated patients**, and thus information only from patients receiving placebo are considered to avoid potential overestimation of the treatment effect. This is consistent with the CHMP advice:

- "… multiple imputation methods would be an acceptable option if no observed data after treatment withdrawal are available. *For imputation of missing data, information from patients receiving placebo should be considered to avoid a potential overestimation of the treatment effect. Reason for drop-out could be used and imputations for patients withdrawing due to lack of efficacy or adverse events could be treated differently."*

However, as suggested by CHMP, for those patients discontinuing due to either AE or to lack of efficacy, imputation is treated differently. This is also specifically suggested in the ICH E9 Addendum, "…for a numerical variable, experiencing an intercurrent event might be ascribed an extreme unfavourable value and a suitable summary measure selected." This will be accomplished with an additional imputation step, as described in the following section.

3.2.3.3. Imputation for Intercurrent Events of Discontinuation due to AE or Lack of Efficacy

For patients with intercurrent events of discontinuation due to adverse event OR due to lack of efficacy, imputation will follow the same strategy as above, with the difference that the imputed Week 24 values (after following the steps in the section above) will be adjusted and therefore ascribed an extreme unfavourable value.

Specifically, for these patients who discontinued due to adverse event or due to lack of efficacy at Week 24, the missing values will be scaled to "worse" as compared to patients with MAR Week 24 data. Ie, a scalar value of 10% worsening, using a patient's baseline to define the 10% scalar value, will be added to the patients Week 24 imputed value in order to 'worsen' the outcome.

The scalar will be determined as follows. Patients are expected to have an average ADAS Cog 11 score of ~25, and thus a 10% worsening would average approximately 2.5 points. Worsening of 2.5 points in the ADAS-Cog 11 was similar for the placebo arms in those clinical trials for many approved drugs (eg, donepezil, rivastigmine, memantine), and thus 10% of a patient's baseline value is a clinically relevant and meaningful adjustment (worsening) to the patients who discontinued due to adverse event or due to lack of efficacy.

For example, assume a patient discontinues prematurely due to an adverse event and the patient is therefore missing the Week 24 ADAS Cog 11 value. Imputation of missing values for this patient will be performed as follows:

- The patients missing values after discontinuation will *FIRST* be imputed accorded to the same procedures as for all patients with missing values.
- Then, the Week 24 imputed value will be FURTHER scaled (worsened) as follows:
  - The patients baseline value for the variable will be identified (eg, assume the patient has an ADAS Cog 11 score of 27 at baseline). And note that increases in the ADAS Cog-11 indicate a worsening of function, and thus increases from baseline would indicate a poor treatment outcome.
  - The baseline score will be multiplied by 10%, eg, 27 * 0.1 = 2.7. This 2.7 would be this individual patients' scalar adjustment to the Week 24 ADAS Cog 11 score.
  - Therefore, the imputed Week 24 score for that patient will be *increased (worsened) by 2.7*. Thus, if, for example, the patients Week 24 value is imputed by the primary imputation methods to be 25, *a further scalar increase of 2.7 will be attributed for that patient,* ie, the resulting imputed score for analysis for that patient will be 25+2.7 = 27.7.

Thus, in this manner, patients who discontinue due to adverse event or due to lack of efficacy will, as a group, have a worse distributional outcome for each co-primary endpoint than other patients with missing values at random. A 10% worsening in the co-primary endpoints is considered sufficient to ensure that these patients who discontinued due to adverse event or due to lack of efficacy are treated accordingly with unfavorable outcomes.

3.2.4. The population-level summary for the variable which provides, as required, a basis for a comparison between treatment conditions

The population-level summary for the co-primary variables in this study are ADAS-Cog-11 score change from baseline to Week 24 and ADCS-ADL score change from baseline to Week 24. The basis for the comparison will be made using the Randomized population and the MMRM analysis on the imputed dataset that specifically attributes distributionally 'worse' outcomes to those patients with an intercurrent event of dropout due to either adverse event or lack of efficacy.

4. ANALYSIS POPULATIONS

The Intent-to-Treat (ITT) Population was defined as:

- All subjects randomized to treatment who have taken at least one dose of double-blind investigational product and who have at least one baseline (and one post baseline) primary efficacy assessment for at least one of either the ADAS-Cog-11 or the ADCS-ADL. This will be the primary population used for the efficacy analysis.

Note that the inclusion of 'and one post baseline' assessment was added to the original definition.

The Randomized Population was defined as all patients who are randomized.

CHMP advice suggested the following:

- "With regard to the definition of the ITT population, it is considered that the Intention-To-Treat population, as the definition already suggests, should consist of all patients randomized and the Applicant will have to justify that their definition (all randomized who took medication with one post-baseline assessment) will lead to the same conclusions on treatment effects."

Therefore, an analysis of the ***Randomized Population*** will be performed for the co-primary endpoints using the multiple imputation strategy as indicated above, and the responder analyses as described for those outcomes.

A comparison of the conclusions regarding treatment effect between the Randomized Population and the ITT population will be performed. This analysis comparison will allow for justification as to whether the ITT population (as defined) leads to the same conclusions from the Randomized population.

5. DESCRIPTIONS OF ADDITIONAL STATISTICAL ANALYSES

The key analysis method of these additional imputed data (as described in the previous sections) will be using the mixed model repeated measures (MMRM) approach. A summary of this method is described below.

Treatment comparisons between the intepirdine and placebo groups in ADAS-Cog-11 and ADCS-ADL change from baseline to Week 24 will be analyzed for the dataset using a mixed model for repeated measures with restricted maximum likelihood estimation, an unstructured covariance matrix, and the Kenward-Roger approximation for denominator degrees of freedom.

The same procedures for analysis of the imputed data (as were proposed in the original SAP) would then be performed, as follows:

- Each of the 100 imputed datasets will be analyzed using the MMRM model. The statistical model will be fitted with terms for treatment group, visit, treatment by visit interaction, baseline score, baseline MMSE, baseline score by visit interaction, and Region. In this way, the test of the efficacy endpoint at Week 24 will be obtained once for each of the 100 imputed datasets.
- The 100 resulting treatment effect parameters and standard errors from these will be combined to provide a distribution of parameters (and standard errors) upon which the sensitivity analysis will be concluded. PROC MIANALYZE will be used for this summary of the analysis.

Inferences will be drawn from treatment differences for the changes from baseline derived from the MMRM models at Week 24. As additional supportive information, treatment differences for each post baseline visit will also be derived. The estimated treatment difference for "intepirdine 35mg – Placebo" at each visit will be displayed in the summary of statistical analysis together with the 95% confidence interval and the associated p-value.

Least Squares Means for each visit will also be presented with the standard error and the number of subjects contributing to the Least Squares Means. Least Squares Means and estimated treatment differences for each visit and the associated 95% confidence interval will be displayed graphically.

Note that while the primary determination of the effects of treatment will remain based on the indicated analyses in the final SAP (August 9, 2017), comparison of those outcomes to the additional analyses above will be made to assess the robustness of the outcomes.

5.1. Sensitivity Analyses of the Primary Efficacy Results

CHMP as the following advice:

- "An appropriate primary analysis should be accompanied with sensitivity analysis to test assumptions on properties of methods and data. It is noted that the set of analyses defined in the SAP partly address a different target of estimation (alternative analysis methods termed "Sensitivity Analyses of the Primary Efficacy Results" in section 8.3.2.2. on observed cases). Single imputation methods as in "Sensitivity Analysis from Imputation of Missing Efficacy Data for Primary Efficacy Endpoints" (section 8.3.2.4.) would not be acceptable and the outlined analysis that makes use of multiple

imputation methods uses the complete trial population to make covariate based predictions for imputation. While multiple imputation methods are considered appropriate, the dataset used to generate imputations would have to be justified."

In response to this advice, the following sensitivity analysis will be performed for the co-primary efficacy endpoints using the 100 imputed datasets:

- Analysis using a non-parametric (distribution-free) Wilcoxon Rank Sum Test of the treatment differences will be run for the Week 24 data, and conclusions will be based on the combined outcomes from the 100 imputed datasets. *Sample* SAS code is as follows. The actual SAS code may be modified as appropriate.

```
Proc npar1way data=dat;
  Class treat;
  Var outcome;
  ods select WilcoxonScores;
  Run;
```

5.2. Responder Analyses

CHMP had the following advice:

- "The defined responder analyses should not only be based on observed cases with assessments at week 24. Imputations should be used and an approach with missing data imputed as failures should be included. A responder criterion of 4 on the ADAS-Cog 11 should be generally implemented in addition to improvement of at least 3 points."

The analyses of responder (defined below) are consistent with that advice, as shown below.

The responder analyses will be performed based on the observed data (subjects with Week 24 assessment for the endpoint of interest), ***as well as analyzed assuming all subjects with missing Week 24 are treatment failures (non-responder).***

The following responder definitions will be defined and each will be analyzed using a logistic regression model as described above as well as a CMH test. The percentage of responders and the difference in proportions compared to placebo will be presented. This analysis is planned for both the ITT population as currently defined as well as the Randomized population.

1. ***ADAS-Cog-11 Improvement by at least 3 points at Week 24***
   a. ***Repeated for by at least 4 points at Week 24***
   b. Repeated for 'no change or improvement' vs 'worsening'
2. ADCS-ADL No change/improvement vs worsening at Week 24
3. CIBIC+ No change/improvement vs worsening at Week 24
4. ADAS-Cog-11/CIBIC+/ADCS-ADL composite, simultaneously meeting the criteria for:
   a. ADAS-Cog-11 Improvement of at least 3 points at Week 24
   b. CIBIC+ No change/improvement at Week 24

c. ADCS-ADL No change/improvement at Week 24

5. ADAS-Cog-11/ADCS-ADL composite, simultaneously meeting the criteria for:

   a. ADAS-Cog-11 Improvement by at least 3 points at Week 24
   b. ADCS-ADL No change/improvement at Week 24

6. ADAS-Cog-11/CIBIC+/ADCS-ADL composite, simultaneously meeting the criteria for:

   a. ADAS-Cog-11 Improvement of at least 4 points at Week 24
   b. CIBIC+ No change/improvement at Week 24
   c. ADCS-ADL No change/improvement at Week 24

7. ADAS-Cog-11/ADCS-ADL composite, simultaneously meeting the criteria for:

   a. ADAS-Cog-11 Improvement by at least 4 points at Week 24
   b. ADCS-ADL No change/improvement at Week 24

6. REFERENCES

Ratitch, B. and O'Kelly, M. (2011), "Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures," in Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group), SP04, Nashville.

APPENDIX 1. ICH E9 (R1) ADDENDUM ON ESTIMANDS AND SENSITIVITY5 ANALYSIS IN CLINICAL TRIALS TO THE GUIDELINE ON STATISTICAL 6 PRINCIPLES FOR CLINICAL TRIALS