Intepirdine as adjunctive therapy to donepezil for mild-to-moderate Alzheimer's disease: A randomized, placebo-controlled, phase 3 clinical trial (MINDSET)

Lang, Frederick M; Mo, Yi; Sabbagh, Marwan; Solomon, Paul; Boada, Merce; Jones, Roy W; Frisoni, Giovanni; Grimmer, Timo; Dubois, Bruno; Harnett, Mark; Friedhoff, Sarah R; Coslett, Shari; Cummings, Jeffrey L

Publication DOI: 10.1002/trc2.12136

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	Version: August 9, 2017
Axovant Sciences Study Director	Ilise Lombardo, MD Telephone: 646-495-8197 Email: ilise.lombardo@axovant.com

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

SIGNATURE PAGE

Lawrence Friedhoff MD, PhD
Chief, Research & Development

Signature Date: 9 August 2017

Ilise Lombardo, MD
Senior Vice President Clinical Research

Signature Date: 9AUG2017

Divakar Sharma
Vice President Biostatistics

Signature Date: 9th Aug 2017.

TABLE OF CONTENTS

LIST OF TABLES

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
AE	Adverse Event
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
CMH	Cochran Mantel-Haenszel
CRF	Case Report Form
CSR	Clinical Study Report
DOB	Date of Birth
ITT	Intent-to-Treat Population
LLN	Lower Limit of Normal
MedDRA	Medical Dictionary for Regulatory Activities Terminology
MMRM	Mixed model for repeated measures
N	Total Sample Size
OC	Observed Cases
PCS	Potential Clinical Significance
PP	Per-Protocol Population
SD	Standard Deviation
SAE	Serious Adverse Event
SAS	Statistical Analysis System
ULN	Upper Limit of Normal

2. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the analyses to be performed following the completion of Study 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.

Study measurements and assessments, planned statistical methods, and derived variables are summarized in this plan. Planned tables, figures, and listings are specified. The SAP was finalized prior to locking the study database and the blind being broken (unblinding the study).

2.1. Nomenclature

For purposes of this SAP, as well as in the analysis tables, figures, and listings, the study drug is referred to as "Intepirdine". The study drug is also referred to (eg, in the protocol) as RVT-101.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Study Objectives

3.1.1. Primary Objective

The primary objective of this study is:

- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on cognitive function as measured by the Alzheimer's Disease Assessment Scale – Cognitive Subscale 11 items (ADAS-Cog-11) after 24 weeks of treatment.
- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on activities of daily living as measured by the Alzheimer's Disease Cooperative Study - Activities of Daily Living (ADCS-ADL) scale after 24 weeks of treatment

3.1.2. Secondary Objectives

Key secondary objectives include:

- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on global clinical assessment of change as measured by Clinician's Interview-Based Impression of Change – plus caregiver interview (CIBIC+) after 24 weeks of treatment
- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on subject dependency as measured by the Dependence Scale (DS) total score after 24 weeks of treatment

Other secondary objectives include:

- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on neuropsychiatric symptoms as measured by the Neuropsychiatric Inventory (NPI) total score after 24 weeks of treatment
- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on the incidence of falls during 24 weeks of treatment
- To assess the effect of intepirdine versus placebo as an adjunct to stable donepezil therapy on different domains of neuropsychiatric symptoms as measured by the NPI after 24 weeks of treatment
- To assess the effects of intepirdine versus placebo as an adjunct to stable donepezil therapy on an analysis of responders based on prespecified efficacy evaluations
- To measure intepirdine plasma concentrations and donepezil plasma concentrations in study subjects
- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on cognition as measured by the ADAS Cog-13 (ADAS-Cog-11 plus delayed recall and digit cancellation count tests) after 24 weeks of treatment

- To assess how baseline Mini Mental State Examination (MMSE) score severity affects efficacy outcome measures after 24 weeks of treatment
- To estimate the pharmacokinetic (PK) parameters of intepirdine and explore relationships to efficacy or safety endpoints, as appropriate
- To assess the safety and tolerability of intepirdine as an adjunct therapy to stable donepezil treatment

3.1.3. Tertiary Objectives

Tertiary objectives of this study are:

- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on healthcare utilization as measured by the Resource Utilization in Dementia Lite (RUD Lite)
- To assess the effects of intepirdine versus placebo as adjuncts to stable donepezil therapy on quality of life as measured by the EuroQOL five dimensions questionnaire (EQ-5D)

3.2. Study Endpoints

3.2.1. Primary Endpoints

The co-primary endpoints in this study are:

- ADAS-Cog-11 score change from baseline to Week 24
- ADCS-ADL score change from baseline to Week 24

3.2.2. Secondary Endpoints

Key secondary endpoints will be tested in a sequential manner as follows:

- CIBIC+ score at Week 24
- DS Total Score changes from baseline to Week 24

Other secondary endpoints include:

- NPI total score change from baseline to Week 24, analyzed as 'No Change/ Improvement' versus 'Worsening (ie, Progressors)'
- Incidence of falls through Week 24 (as collected through adverse event reporting)
- ADAS-cog score change from baseline; ADCS-ADL score change from baseline; CIBIC+ score; all assessed by MMSE baseline score
- ADCS-ADL Basic, Instrumental, and Independence score changes from baseline to Week 24
  - Communication and Engagement Factor and Outside Activities Factor changes from baseline to Week 24

- NPI total, caregiver distress, Psychosis domain score changes from baseline to Week 24 and Progressors analyses of the Psychosis Domain and each of the 12 NPI domain scores
- Responders, using the following definitions:
    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
- Progressors in Dependence Level, defined as patients worsening from baseline,
- ADAS-Cog-13 (ADAS-Cog-11 plus delayed recall and total digit cancellation) score change from baseline to Week 24
- ADAS-Cog-11 and ADCS-ADL Time to progression, defined as a worsening (from baseline) by 1 point (or more).
- Measurement of concentrations of intepirdine and donepezil in plasma
- Steady state area under the concentration-time curve (AUCτss), peak concentration (Cmax-ss), and minimum concentration (Cmin-ss) of intepirdine in plasma
- Occurrence of adverse events (AEs) and changes in physical examinations, vital signs measurements, electrocardiograms (ECGs), routine laboratory assessments, and Columbia Suicide Severity Rating Scale (C-SSRS)

3.2.3. Tertiary Endpoints

The tertiary endpoints in this study are:

- RUD Lite score change from baseline to Week 24
- EQ-5D score change from baseline to Week 24

3.3. Descriptions and Scoring of Efficacy Endpoints

3.3.1. Mini-Mental State Evaluation (MMSE)

While the MMSE is not specifically an endpoint, it is used to measure baseline disease severity, and therefore is included in the descriptions in this section.

The MMSE [Folstein 1975] consists of 11 tests of orientation, memory (recent and immediate), concentration, language and praxis. Scores range from 0 to 30, with lower scores indicating greater cognitive impairment. The MMSE score is the sum of the 11 items; if any item is missing then the total score will be set to missing.

3.3.2. Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-Cog)

The 11-item and 13-item ADAS-Cog [Rosen 1984; Mohs 1997] assesses a range of cognitive abilities including memory, comprehension, orientation in time and place, and spontaneous speech. Most items are evaluated by tests, but some are dependent on clinician ratings on a 5-point scale.

- The ADAS-Cog 11 total score range is from 0 to 70, with a higher score indicating more severe cognitive impairment.
- The ADAS-Cog 13 is the ADAS-Cog 11 with two additional items:
  - Delayed Word Recall and Number Cancellation.
  - Scores for the ADAS-Cog-13 range from 0 to 85 with higher scores indicating greater dysfunction.

If a question was not completed for cognitive reasons, the worst possible score will be assigned for the question.

In order to calculate the ADAS-Cog-11 total, the following scores are first computed:

1. The score for Word Recall is calculated as the mean number of words not recalled over the three trials (Trial 1, 2 and 3). Note that if data is available for only two of the trials then the calculated score for Word Recall will be the average of the two trials, and if a score for only one trial is recorded then this will be set to be the score for Word Recall. The score for Word Recall will be set to missing only if data is missing for all three trials.
2. The results from Naming Objects and Fingers are scored as follows:

   0 = 0 – 2 items named incorrectly
   1 = 3 – 5 items named incorrectly
   2 = 6 – 8 items named incorrectly
   3 = 9 – 11 items named incorrectly
   4 = 12 – 14 items named incorrectly
   5 = 15 – 17 items named incorrectly

The ADAS-Cog-11 total score is the sum of the calculated scores for the following items:

1. Word Recall
2. Commands

3. Constructional Praxis
4. Naming Objects and Fingers
5. Ideational Praxis
6. Orientation
7. Word Recognition
8. Remembering Test Instructions
9. Comprehension
10. Word Finding Difficulty
11. Spoken Language Ability

Most items are evaluated by tests, but some are dependent on clinician ratings. The total score ranges from 0-70 with higher scores indicating greater dysfunction.

When a score is missing due to non-cognitive reasons for one of the questions, the total score will be calculated as a weighted average of the scores provided for the remaining ten questions as follows (and see Table 2):

- Imputed Total= Observed Total Score * (1+ Maximum Score of the missing value / Sum of the Maximum Score of the non - missing values)

Table 2: Calculation of Imputed ADAS-Cog-11 Total Score (Study RVT-101-3001)

Missing Question	Imputed Total formula
Word Recall	Observed Total*(1+10/60)
Orientation	Observed total*(1+8/62)
Word Recognition	Observed total*(1+12/58)
All other questions	Observed total*(1+5/65)

In cases where 2 or more questions are missing, the total score will not be imputed and will be set to missing.

ADAS-COG-13 will be sum of all the ADAS-COG-11 item scores plus the item scores of Delayed Word Recall and Number Cancellation. The imputed total score of ADAS-Cog-13 when there is only one item is missing due to non-cognitive reasons is presented in Table 3.

Table 3: Calculation of Imputed ADAS-Cog-13 Total Score (Study RVT-101-3001)

Missing Question	Imputed Total formula
Word Recall	Observed Total*(1+10/75)
Orientation	Observed total*(1+8/77)
Word Recognition	Observed total*(1+12/73)
Delayed Word Recall	Observed Total*(1+10/75)
All other questions	Observed total*(1+5/80)

3.3.3. Alzheimer's Disease Cooperative Study - Activities of Daily Living (ADCS-ADL)

The ADCS-ADL (Alzheimer's Disease Cooperative Society - Activities of Daily Living) scale [Galasko 1997] measures functional impairment in terms of activities of daily living. The ADCS-

ADL is an interviewer-administered informant-based scale where the informant (caregiver) responds to 23 activities of daily living questions about the subject. There are 4 scores derived: the Total score, Basic score, Instrumental score, and the Independence Score.

The questions range from basic to instrumental activities of daily living and take approximately 20 minutes to complete. The Total score ranges from 0-78 and a higher score signifies greater functional ability; the lower the score the greater (worse) the impairment.

The questionnaire is split into two types of questions, an initial question relating to whether a subject has completed a particular activity and then a follow-on question which scores how much assistance the subject has required if they have performed that particular activity.

- The Total score is calculated by adding up the responses for each of the individual activities. Note if a caregiver responds to the question relating to whether the subject performed a particular activity as no or don't know this will contribute a zero to the Total score.
- The Basic score will also be calculated as the sum of questions 1-6b and ranges from 0-22 (activities included in the basic score are: eating, walking, using the toilet, bathing, grooming and dressing)
- The Instrumental score in the, sum of questions 7-23, ranges from 0-56 (the Instrumental score includes: using the telephone, watching television, conversations, clearing dishes, personal belongings, making drinks, making snacks, taking rubbish out, getting out and about, shopping, keeping appointments, being left alone, current events, reading, writing, pastimes/hobbies, household chores).
- The Independence Score is calculated by re-scoring all 23 individual questions for each activity, such that a given question is given a '1' if the MAXIMUM (best) score is observed, and a '0' otherwise. Thus, each question contributes a maximum of 1 unit to the total, with a total range for the Independence Score ranging from 0 (worst) to 23 (best).

If there are any missing scores within the Basic activities questions, the Basic total score will not be imputed and therefore will be set to missing.

If there is only one missing score within the Instrumental activities questions, the Instrumental score will be calculated as a weighted average of the scores provided for the remaining 16 activities as follows (and see Table 4):

- Imputed Total = Observed Total Score x (1+ Maximum Score of the missing value/ Sum of the Maximum Score of the non -missing values)

Table 4: Calculation of Imputed ADCS-ADL Instrumental Score (Study RVT-101-3001)

Missing Question	Imputed Instrumental formula
8a, 8b, 8c, 16b, 18a,18b, 18c,19a,19b,19c, 20a, 20b	Observed total*(1+1/55)
9, 10, 11, 12, 14, 16a, 17, 21, 22b	Observed total*(1+3/53)
13, 15, 23b	Observed total*(1+4/52)
7	Observed total*(1+5/51)

The total ADCS-ADL score will be calculated by adding up the Basic total score and the Instrumental imputed total score.

The Total Independence score will give any missing score from either Basic or Instrumental questions a value of zero.

In cases where 2 or more questions are missing, the total Instrumental score will not be imputed (ie, it will be set to missing) except for the following situation:

The following is the English version of Questions 18a, 18b and 18c:

18a. Was {P} left AWAY FROM HOME, for 15 minutes or longer, during the day?

18b. Was {P} left AT HOME, for an hour or longer, during the day?

18c. Was {P} left AT HOME, for less than 1 hour, during the day?

In Spain and Korea, because Questions 18b and 18c were incorrectly translated as "left AWAY FROM HOME" rather than "left AT HOME", scores of Questions 18b and 18c are not correct as intended when Question 18, "During the past 4 weeks, was {P} **ever left on his/her own**?", has a response of "Yes". Therefore, for all subjects' visits in Spain and Korea, when Question 18=Yes, Questions 18b and 18c will be considered missing and the total Instrumental Score will be normalized as follows in cases where no other Instrumental ADCS-ADL question is actually missing:

Observed total*(1+2/54)

Correct scores of Questions 18b and 18c in Spain and Korea will also be collected separately for visits that occur after accurate translations are deployed to sites.

The Independence score will assign a '0' to any question that is missing.

In addition, two Factor Scores are predefined for this study. These Factor Scores are specifically noted to be of relevance to patients with AD, particularly with progression of the disease.

- Communication and Engagement Factor Score, defined as the sum of the following questions (where all questions are needed to calculate the Factor Score):
  - Watching television (Questions 8a to 8c)
  - Conversation (Question 9)
  - Keeping appointments (Question 17)
  - Current events (Questions 19a to 19c)
  - Reading (Questions 20a and 20b)
  - Writing (Question 21)
  - Hobbies (Question 22).
- Outside Activities Factor Score, defined as the sum of the following questions (where all questions are needed to calculate the Factor Score):
  - Getting out and about (traveling, Question 15)
  - Shopping and paying (Questions 16a and 16b)

– Being left alone (Questions 18a to 18c).

3.3.4. Dependence Scale (DS)

The DS measures the amount of assistance patients with dementia require in performing daily activities [Brickman 2002]. The caregiver answers questions about the dependency of the subject. The scale consists of 13 items. A total score will be calculated with a range from 0 to 15 with higher scores indicating greater dependency.

In addition, a Dependence Level score ranging from 0 to 5 based on responses to the DS items will be analyzed. The Dependence Level score of 0 indicates no dependence, while a score of 5 indicates complete dependence in self-care activities. This will be used for a Progressor-type analysis, with any score increasing is considered progressing (worsening).

No imputation for missing values is performed; thus, if an item is missing, the corresponding domain is not calculated (ie, is set to missing).

3.3.5. Clinician's Interview Based Impression of Change Plus Care Giver Interview (CIBIC+)

The CIBIC+ assessment [Schneider 1997] measures the global functioning of the subject through structured interviews by an investigator with both the subject and caregiver. The change from baseline is recorded on a 7-point scale with a score of 4 indicating no change, scores above 4 indicating worsening, and scores below 4 indicating improvement.

The CIBIC+ scale is a single item instrument, if the CIBIC+ is not done at a visit or the result is missing at a visit, the result will be set to missing.

3.3.6. Neuropsychiatric Inventory (NPI)

The NPI is a behavior rating scale composed of a 12-item structured interview of the caregiver that is scored from 0 to 144 (the higher the score, the greater the psychiatric disturbance). It assesses 12 behavioral and psychological disturbances (domains) occurring in dementia patients. These 12 domains are:

- Delusions
- Hallucinations
- Dysphoria
- Anxiety
- agitation/aggression
- euphoria
- disinhibition
- irritability
- apathy
- aberrant motor activity

- night time behavior disturbance
- eating abnormalities.

For each domain there are four scores assigned:

- Frequency
- Severity
- Total (frequency x severity).
- Caregiver distress

Scores are assigned to questions as demonstrated in Table 5. No imputation for missing values is performed; thus, if an item is missing, the corresponding domain is not calculated (ie, is set to missing).

Table 5: Scores for NPI Individual Questions (Study RVT-101-3001)

Category	Numeric values assigned
Frequency is rated as:	1=Occasionally - less than once per week 2=Often - about once per week 3=Frequently - several times per week but less than every day 4=Very frequently - daily or essentially continuously present
Severity is rated as:	1=Mild - produces little distress in the patient 2=Moderate - more disturbing to the patient but can be redirected by the caregiver 3=Severe - very disturbing to the patient and difficult to redirect
Caregiver Distress is rated as:	0=no distress 1=minimal 2=mild 3=moderate 4=moderately severe 5=very severe or extreme.

There are three overall scores that will be analyzed from the NPI:

- A Psychosis Domain will be defined as the sum of the Delusions total and the Hallucinations total.
  - Thus, the Psychosis Domain score can range from 0 to 24. This will be then dichotomized as 'No Change/ Improvement' versus 'Worsening (ie, Progressors)' from baseline for each patient.
- The NPI Total score will be calculated by adding the total scores of the 12 domain scores together (and thus a range of 0 to 144). The caregiver distress score is not included in the total NPI score.
- Caregiver Distress is defined as the sum of the caregiver distress items from the 12 domains, and thus scores can range from 0 to 60

3.3.7. EuroQol-5D (EQ-5D)

The EQ-5D is a standardized measure of health status that provides a measure of health-related quality of life that is widely used in clinical trials [Rabin 2001]. For this study the EQ-5D will be a caregiver proxy assessment. The assessment will be completed by the caregiver and will assess the caregiver's impressions of how the subject would rate his/her own quality of life.

The EQ-5D questionnaire consists of 2 components: the EQ-5D descriptive system and the EQ visual analogue scale (EQ-VAS). The EQ-VAS records overall health status on a 20 cm vertical line with a score of 0 (worst health one can imagine) to 100 (best health one can imagine).

The EQ-5D comprises the following five dimensions:

- Mobility
- self-care
- usual activities
- pain/discomfort
- anxiety/depression.

Each dimension has 5 response levels:

- no problem
- slight problems
- moderate problems
- severe problems
- extreme problems.

The number and percentage by response level will be summarized by visit and treatment for each domain. The response levels will be dichotomized to show number and percentage of subjects without problems (reported no problems) and with problems (slight problems, moderate problems, severe problems, and extreme problems).

The index values of EQ-5D-5L Health Status Today and attack EQ-5D-5L will be calculated using the crosswalk link function based on the published EQ-5D-5L Crosswalk Index Value set. For countries where the index value sets were not yet available, regional values will be selected.

3.3.7.1. RUD-Lite

The RUD-Lite scale [Wimo 1998] is designed to assess caregiver burden and provide pharmacoeconomic data related to AD. The RUD-Lite (Version 3.3) includes both baseline and follow-up questions, and includes two domains:

- Personal ADL relates to assisting subjects with basic activities of daily living (eg, grooming).
- Instrumental ADL relates to assisting subjects with instrumental activities of daily living (eg, driving to store).

Table 6 demonstrates how these measures will be calculated for caregiver time spent during the last 30 days caring for the subject:

Table 6: RUD-Lite Domains (Study RVT-101-3001)

Category	Formula from Questions
Personal ADL (Activities of Daily Living)	2A (hours and minutes) x 2B (days)
Instrumental ADL	3A (hours and minutes) x 3B (days)

In addition to the categories above, a Total score will be calculated as follows:

- Total Score = Personal ADL + Instrumental ADL

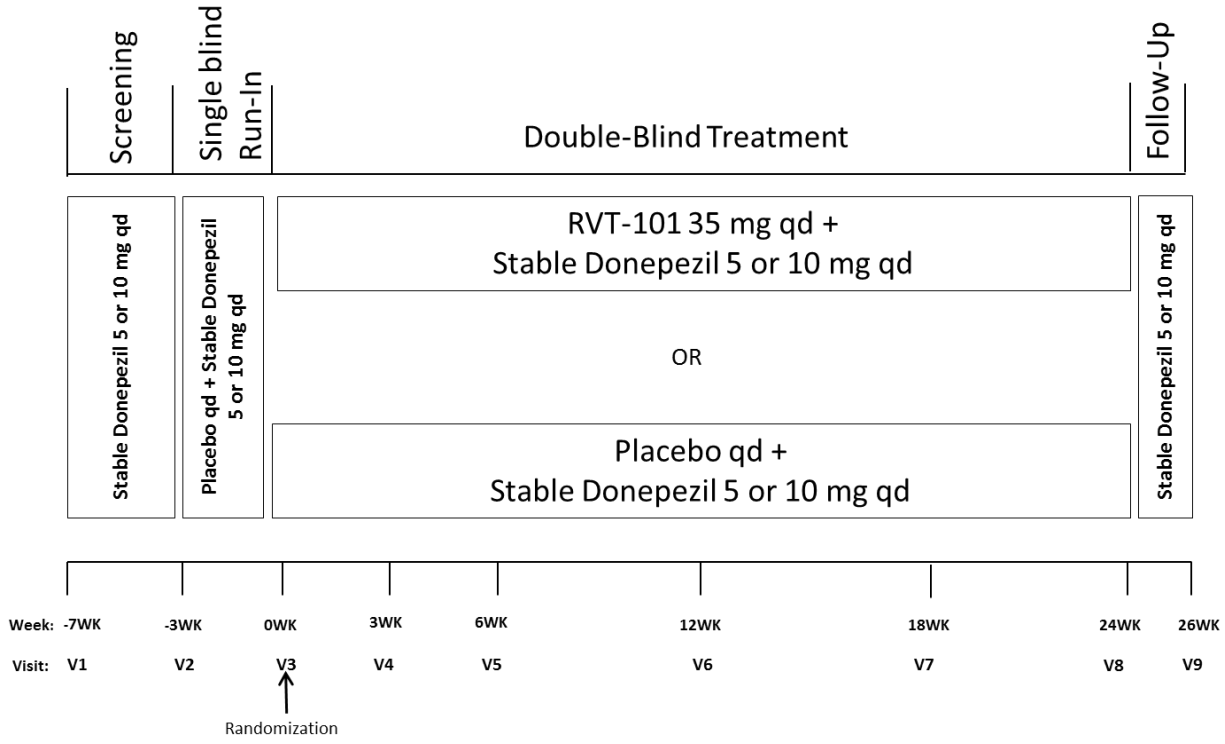
4. STUDY DESIGN

4.1. Summary of Study Design

This is a multi-center, double-blind, randomized, placebo-controlled, parallel-group study in patients with mild to moderate Alzheimer's disease who are on stable donepezil therapy. The efficacy and safety of intepirdine at a dose of 35 mg will be evaluated over a 24 week treatment period. Approximately 1150 subjects will be enrolled. The randomization ratio will be 1:1 (intepirdine:placebo) and enrolled patients will be stratified by baseline MMSE score. The primary endpoints will be measured after 24 weeks of treatment. Study participation will last approximately 33 weeks: 0 to 4 weeks for Screening, a 3 week Single-Blind Placebo Run-In Period to evaluate baseline status, a 24 week randomized Double-Blind Treatment Period, and a 2-week Follow Up (post-treatment) Period.

Figure 1 provides a study scheme demonstrating the key design elements and scheduled study visits.

Figure 1: Study Schema (Study RVT-101-3001)



4.2. Definition of Study Drugs

There are two study drugs being given:

- Intepirdine (Test therapy): Intepirdine 35 mg given as an adjunct to a stable regimen of donepezil (5 mg or 10 mg)
- Placebo (Reference): Matching placebo, also given as an adjunct to a stable regimen of donepezil (5 mg or 10 mg)

4.3. Sample Size Considerations

4.3.1. Sample Size Justifications

The primary comparisons of interest are to compare intepirdine to placebo at Week 24 for change from baseline in both ADAS-Cog-11 and ADCS-ADL. A sample size of 435 subjects per treatment group will allow a difference of 1.6 points between placebo and active treatment in the change from baseline in ADAS-Cog score to be detected with 95% power and a 0.05 significance level assuming an underlying standard deviation (SD) of 6.5. This sample size will also allow a difference of 2 points between placebo and active treatment in the change from baseline in ADCS-ADL score to be detected with 90% power and a 0.05 significance level assuming an underlying SD of 9.

Both endpoints need to achieve a significance level of 0.05 to maintain an overall 5% significance level. Hence, the multiplicity issue associated with the trial having two primary endpoints is addressed as efficacy will only be concluded if both endpoints are significant.

Assuming a drop-out and missing data of approximately 25% up to Week 24, a total of approximately 1150 subjects will be randomized in a 1:1 ratio to intepirdine or placebo.

4.3.2. Sample Size Re-estimation

No sample size re-estimation is planned.

4.4. Randomization/Stratification

Subgroups using MMSE score at baseline will be examined (see Section 6.6). Stratification levels will be defined as presented in Table 7.

Table 7: MMSE Stratification Levels (Study RVT-101-3001)

Baseline MMSE Score	**10-15**	**16-20**	**21-26**
Target for recruited population	≤30%; ≥20%	≤60%; ≥40%	≤30%; ≥20%

4.5. Clinical Assessments

The schedule of clinical assessments is presented in Table 8. Detailed descriptions of these assessments can be found in the study protocol.

Table 8: Schedule of Events (Study RVT-101-3001)

Study Period:	**Screening**	**Run-in**	**Baseline**	**Treatment**					**Follow-up**	**ET**
Study Visit Number:	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**ET**
Study Week:	**W(-7)**	**W(-3)**	**W0**	**W3**	**W6**	**W12**	**W18**	**W24**	**W26**	
Study Day: relative to Baseline unless specified	**Up to 28 days before V2**	**21 + 5 days before V3**	**0**	**21 ± 3**	**42 ± 3**	**84 ± 5**	**126 ± 5**	**168 ± 5**	**14 to 19 days after last dose of IP**	
Informed consent	X									
Inclusion and exclusion criteria	X		X							
Demography	X									
Medical history	X									
Concomitant medications review	X	X	X	X	X	X	X	X	X	X
Urine drug screen	X									
C-SSRS	X		X	X	X	X	X	X	X	X
Randomization			X							
Dispense investigational product		X	X		X	X	X			
Dispense study-supplied donepezil		X	X		X	X	X			
Assess investigational product and study-supplied donepezil compliance			X	X	X	X	X	X		X
Physical exam/ current medical conditions	X	X	X	X	X	X	X	X	X	X
Complete neurological exam	X		X			X			X	
MRI or CT	X									
12-lead ECG	X	X	X	X	X	X	X	X	X	X
Vital signs	X	X	X	X	X	X	X	X	X	X
Review adverse events		X	X	X	X	X	X	X	X	X
Serum OR urine pregnancy test	X								X	X
Hep B, and Hep C screen	X									
TSH, vitamin B_{12}, syphilis serology	X									
Serum chemistry	X	X	X	X	X	X	X	X	X	X
Hematology	X	X	X	X	X	X	X	X	X	X
Urinalysis	X	X	X	X	X	X	X	X	X	X
Blood alcohol content	X	X	X	X	X	X	X	X		
intepirdine level					X	X	X	X		

Study Period:	**Screening**	**Run-in**	**Baseline**	**Treatment**					**Follow-up**	**ET**
Study Visit Number:	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**ET**
Study Week:	**W(-7)**	**W(-3)**	**W0**	**W3**	**W6**	**W12**	**W18**	**W24**	**W26**	
Study Day: relative to Baseline unless specified	**Up to 28 days before V2**	**21 + 5 days before V3**	**0**	**21 ± 3**	**42 ± 3**	**84 ± 5**	**126 ± 5**	**168 ± 5**	**14 to 19 days after last dose of IP**	
Donepezil level			X			X		X		
Hachinski Ischaemia Scale	X									
ADAS-Cog		X	X	X	X	X	X	X		
MMSE	X		X							
ADCS-ADL		X	X	X	X	X	X	X		
CIBIS			X							
CIBIC+						X	X	X		
NPI			X			X	X	X		
DS			X					X		
RUD Lite			X					X		
EQ-5D			X					X		

5. ANALYSES

5.1. Interim Analyses

No unblinded interim efficacy or safety analyses are planned.

5.2. Final Analyses

There is only one analysis planned for this study. This final analysis will be performed upon completion of the following items:

- The database has been locked according to the study data management plan.
- The list of patients excluded from Per Protocol population has been identified.

6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

6.1. General Table and Individual Subject Data Listing Considerations

Tables, listings and figures will be prepared in accordance with the current International Conference on Harmonization Guidelines. The information and explanatory notes in the "footer" or bottom of each table and listing will include the following information:

- Date of output generation
- SAS® program name, including the path that generates the output
- Any other output specific details that require further elaboration

All hypothesis tests and confidence intervals will be two-sided at an alpha level of 5%.

Version 9.1 or higher of the SAS system will be used to analyze the data as well as to generate tables, figures, and listings.

In general, tables will be formatted with a column displaying findings for all subjects combined. The summary tables will clearly indicate the number of subjects to which the data apply, and "unknown" or "not performed" will be distinguished from "missing" values data.

The treatment groups will be referred to in the tables, listings, and figures with the following conventions:

- Intepirdine 35 mg
- Placebo

Data analyses will be performed by, or under the direct supervision of, Axovant Sciences.. Summaries will be presented by treatment and time. Unless stated otherwise, descriptive summaries will include n, mean, standard deviation, coefficient of variation (CV), median, minimum, and maximum for continuous variables, and n and percent for categorical variables.

In addition, the geometric mean with associated 95% CI, between-subject CVs (CVb), and SD of the logs based on the geometric mean, will be included for the log-transformed PK parameters.

The between-subject %CV of the log transformed variable will be calculated as:

- $\%CVb = 100 * \sqrt{(e^{SD\wedge 2} - 1)}$, with SD of the log-transformed data

Deviations from the analyses as described in this SAP will be identified in the final Clinical Study Report. When differences exist between the protocol-described analysis and the SAP, the SAP will take precedence.

All data collected in the CRF, as well as laboratory data, will be provided in data listings, sorted (in general) by patient and chronological date order. For date fields, Study Day will be calculated and presented in these listings.

6.2. Analysis Populations

It is intended that a complete accounting of patients for the analysis populations will be provided, from the Screening Population through the Per-Protocol Population.

6.2.1. Screening Population

All patients who are screened (signed an informed consent) will be included in the Screening Population.

6.2.2. Placebo Run-In

All patients entering the Placebo Run-in will be included in the Placebo Run-In population. This population will be used to provide an accounting of the disposition of patients during this phase of the study.

6.2.3. Randomized Population

The Randomized Population will include all patients who are randomized.

6.2.4. Safety Population

Safety population will consist of all subjects who were randomized and took at least one dose of double-blind investigational product.

6.2.5. ITT Population

The Intent-to-Treat (ITT) Population will consist of 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.

6.2.6. Per-Protocol (PP) Population

The Per-Protocol (PP) Population will consist of those members of the ITT Population who have no major protocol violations. This population will be used for confirmatory analysis of the two primary efficacy variables.

Major protocol violations will be defined as follows:

- Subject had >7 consecutive days interruption in double-blind study medication
- Subject had >7 consecutive days interruption or change in donepezil dose during the study
- Treatment with excluded concomitant medications.
- Significant study drug administration errors.
- No documented history of at least 4 months of ongoing donepezil therapy for AD.
- No documented history of at least 2 months of stable donepezil therapy of 5 or 10 mg
- Invalid week 0 efficacy assessment.
- Subject does not have MMSE score 12-24 inclusive at Screening
- Subject does not have MMSE score 10-26 inclusive at baseline.
- MMSE score at baseline not within +/- 3 points of the Screening value

- Evidence of other causes of dementia.
- History or presence of significant medical or psychiatric condition.
- Subject does not live with (or have substantial Periods of contact with) a regular caregiver.
- Subject is <80 % compliant or > 120% compliant between Visit 7 (Week 18) and Visit 8 (Week 24) with either study drug or donepezil. Patients who fail to return their bottle (and thus compliance cannot be determined) will be assumed to have been compliant for purposes of the Per Protocol population.

6.2.7. Completers (CS) Population

The study completers (CS) Population will consist of those members of the ITT Population who have completed the study. This population will be used for supportive analysis of the primary efficacy variable and other cognition and efficacy endpoints.

6.2.8. Pharmacokinetics (PK) Population

The PK population will include all subjects in the Safety Population who undergo plasma PK sampling and have at least one post-baseline evaluable PK concentration result.

6.3. Multiple Comparisons and Multiplicity

This is a Phase 3 study in which a number of efficacy endpoints will be assessed. There are two pre-specified co-primary endpoints, the ADAS-Cog-11 and the ADCS-ADL, tested at Week 24. Both endpoints need to achieve a significance level of 0.05 to maintain an overall (study-wise) 5% significance level. Hence, the multiplicity issue associated with the trial having two primary endpoints is addressed as efficacy will only be concluded if both endpoints are significant.

Further, the pre-specified key secondary endpoints will be tested in a sequential manner. The comparisons for the pre-specified key secondary efficacy endpoints will be performed at the 5% level of significance using a stepwise gate-keeping testing procedure in which testing will proceed to the next ordered tests only if the prior test was found to be significant (p-value $\leq$ 0.05). This method will be used to control the type I error rate for these key secondary endpoints.

Analysis of each key secondary variable will be performed, with the nominal p-values reported. The interpretation of those p-values (and conclusions regarding hypothesis testing) will then be performed using the sequential testing strategy, as follows:

1. CIBIC+ score at Week 24. If significant (p-value $\leq$ 0.05), proceed to next test.
2. Dependence Scale (DS) Total Score changes over time. The DS is collected at only one time point post-baseline. Thus, this analysis will be performed using BOTH, an observed case and LOCF. The ANCOVA will have effects for treatment, baseline DS score, baseline MMSE, and Region (as defined in Section 6.6). The observed case analysis will be considered the primary assessment of this variable, with the LOCF analysis supportive.

6.4. Baseline Definition

Baseline or pre-dose assessment is the last available assessment prior to time of first dose of double-blind treatment, unless otherwise specified. If assessments fall on the same day as the date of first dose, those assessments will be assumed to be baseline assessments. If there are multiple assessments collected on the same date, the average of these assessments will be used.

6.5. Derived and Transformed Data

6.5.1. Baseline Age

Age (in years) will be determined at the baseline visit (V3) using the date of the baseline visit as recorded in the CRF.

Some birth dates may be incomplete. If the month and the day are missing, we will impute January 1, if the day is missing, we will impute the 1st of the month.

6.5.2. Study Day

Study day for the assessment date of interest should be calculated relative to the first dose date of double-blind investigational product using the appropriate formula below:

- If the assessment date of interest is on or after the first dose date of double-blind investigational product:
  - Study day = assessment date – first dose date of double-blind investigational product + 1
- If the assessment date of interest is before the first dose date of double-blind investigational product:
  - Study day = assessment date – first dose date of double-blind investigational product

6.5.3. Handling of Incomplete Dates

Where necessary for the calculation of derived variables, partial dates will be completed using the earliest calendar date based on the partial date provided. This rule is valid for all partial dates with the exception of the following, for which the date is left as missing:

- Start and stop dates of adverse events
- Start and stop dates of concomitant medication
- Start and stop dates of double-blinded medication

Completely missing dates will not be replaced and the corresponding derived variables will be set to missing.

6.5.4. Visit Windows

Actual times will be used in the listings, individual concentration-time profiles, and in the calculation of PK parameters.

There are 4 study periods during this study:

- The **Screening Period** is defined as the period of time prior to the subject receiving the first dose of single-blind placebo run-in medication.
- The **Single-Blind (or Placebo) Run-In Period** is defined as the period from the first dose of single-blind run-in medication and ends on the last dose date of single-blind run-in medication. The assessments taken on the date of the first dose of randomized double-blind study medication will be assumed to have been completed before the first dose of study medication and will be slotted to this period.
- The **Double-Blind Treatment Period** starts on the date of first dose of randomized double-blind investigational product and ends on the last dose date +7 days. If the first dose of randomized investigational product is taken on the same day as the baseline assessments, it will be assumed that the baseline assessments were completed before the first dose of investigational product was taken.
- The **Post-Treatment Period**. Assessments are regarded as Follow-Up if the assessment occurred >7 days after the last dose date of investigational product.

Data collected at assessment visits provide information of the status at that point in time (e.g. efficacy measures, vital signs, laboratory parameters, etc.) and may provide biased results if the visit is attended early or late, in which case the subject will have received more or less treatment than scheduled. For this reason, visits will be slotted with similar investigational product exposure.

Post-baseline assessments from the Double-Blind Treatment Period data will be visit slotted according to the time intervals shown based on the days relative to the date of first dose of double blind investigational product (Day 1). Early withdrawal data that is assigned to Double-Blind Treatment Period will be slotted using the assessment windows.

Data will be phased first. All data that slots into pre-treatment period will be assigned a visit of Screening. All data that occurs in Single-Blind Run-In Period will be assigned a visit of Run-in or Baseline depending on medication start/stop dates and data that slots into the Follow Up Period will be assigned a visit of Follow Up.

Table 9 provides the windows (relative to start of double-blind treatment) that visit data are to be slotted into.

Table 9: Windows for Slotting Visits (Study RVT-101-3001)

Days relative to start of double-blind treatment*	Target Day	Visit Slot
Days 2 - 34	21	Week 3
Days 35 - 69	42	Week 6
Days 70 – 104	84	Week 12
Days 105 - 140	126	Week 18
Days 141 - 196	168	Week 24
*Date of assessment – date of first dose + 1		

For data where two or more assessments are slotted into the same visit interval, the data recorded closest to the target visit day will be used for the summary tables and analyses. If two assessments are equally close to the target day, the later of the two assessments will be used in tabulations and analyses; data from the earlier assessment(s) will be listed only.

- Note that for safety data (except blood pressure or heart rate assessments), the closest assessment is always the assessment used, however, for efficacy, if the closest assessment is not evaluable, the other assessment would be used. For example, for ADAS-Cog, if the total score could not be calculated for the closest assessment, the other assessment would be used.
- If on one of the dates only one of the sitting or standing blood pressure assessments is present then the assessments recorded on the other date, which slots to the same visit interval, will be used in tabulations and analyses. Therefore orthostatic blood pressure will always be calculated using sitting and standing measurements from the same date, these assessments however may not be the closest to the target day. A similar rule applies for Heart rate assessments.

For data where two or more assessments are slotted to follow up (and thus after double-blind treatment has ended), then the data recorded closest to target day (14 days after last dose of double-blind medication) will be used for summaries.

6.5.5. Missing Start and Stop Dates for Prior and Concomitant Medication and Adverse Events

For analyses of AEs and concomitant medication usage, a complete date must be established in order to correctly identify the AE or medication as occurring during treatment or not. For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the subject data listings (ie, no imputed values will be displayed in data listings).

Missing start day, but month and year present:

- If the date of the first dose of double–blinded study medication is in the same month and year as the occurrence of the AE/concomitant medication, the start day of the event/concomitant medication will be assigned to the day of the first dose of the double-blinded medication.
- Otherwise the start day will be set to the 1st day of the month.

Missing start day and month, but year present:

- If the first dose of double–blinded study medication in the same year as the occurrence of the AE/concomitant medication, the start day and month will be assigned to the date of the first dose of the double-blinded medication.
- Otherwise the start day and month will be set to January 1st.

Missing end day, but month and year present:

- The end day will be set to the last day of the month.

Missing end day and month, but year present:

- The end day and month will be set to the maximum of the date of study termination or the date equivalent to 30 days after the last dose of the double-blinded medication.
- However, if the study termination year and year for the date of the last dose of the double-blinded medication +30 days are greater than the event/concomitant medication year, the day and month are to be set to December 31st.

6.6. Subgroups

While this study is not powered for subgroup analyses, select efficacy and safety endpoints will be analyzed by defined subgroups. Additional subgroup may be added as the data dictate. If the total number of subjects in a subgroup is < 10% of the total number of randomized subjects, the analysis for that subgroup may not be performed.

The following subgroups are planned for safety analysis:

- Disease Severity based on baseline MMSE
  - 10-19 and 20-26
  - Mild (16-26) and Moderate (MMSE of 10-20)
- Geographic Region
  - US, Non-US English (UK, Canada, Australia), West Europe (Germany, Spain, Italy, France), East Europe (Czech Republic, Poland, Slovakia, Croatia, Serbia, Bulgaria), and Other (Rest of World)
- Age (years)
  - <74 and ≥74
  - <65 and ≥65
- Gender
  - Female and Male
- Race
  - White and Non-White
- Ethnicity
  - Hispanic and Not Hispanic
- Stable Dose of Donepezil
  - 5 mg and 10 mg

The following subgroups are planned for efficacy analysis:

- Disease Severity based on baseline MMSE
  - 10-19 and 20-26
  - Mild (16-26) and Moderate (MMSE of 10-20)

- Geographic Region
  - US, Non-US English (UK, Canada, Australia), West Europe (Germany, Spain, Italy, France), East Europe (Czech Republic, Poland, Slovakia, Croatia, Serbia, Bulgaria), and Other (Rest of World)
- Age
  - <74 and ≥74
- Gender
  - Female and Male
- Race
  - White and Non-White
- Stable Dose of Donepezil
  - 5 mg and 10 mg

Analysis of the primary endpoints and the CIBIC+ will be run by these subgroups.

7. STUDY POPULATION

7.1. Subjects Disposition

The total number of subjects that participated in the study will be summarized, including whether they were randomized or not. A tabulation will be provided for the number of subjects screened, screen-failed prior to the Placebo Run-in period. The denominator for those outcomes for this analysis will be based on the number of screened subjects.

The number subjects entering the Placebo Run-in, number subjects not completing the Placebo Run-in (and reasons for not completing) will be presented, using the subjects entering the Placebo run-in as the denominator.

The number patients randomized, as well as the number of patients in the analysis populations (Safety, ITT, etc.) will be presented for the Randomized Population.

A second analysis will be performed, using the Safety population as the denominator, of the number of patients included in the Safety, ITT, Per Protocol, Completers and pharmacokinetic populations; this will be summarized by treatment group.

- A tabulation of the number of subjects randomized in each country will be provided.
- An overall summary of the number and percentage of subjects who completed or withdrew prematurely from the study will be displayed by treatment group, for the overall population as well as by MMSE strata. Reasons for premature withdrawal will be presented in the order they are displayed in the CRF. Subjects who withdrew prematurely from the study will be listed by treatment group and subject.
- A cumulative distribution plot of the time to early withdrawal will be plotted and will include both treatment groups in order to allow for a visual inspection of the time to discontinuation. A similar plot will be provided for those patients who discontinue due to adverse event. The denominator for this plot will be the underlying population (and thus 'censoring' will not be applied).

7.2. Screen Failures

Screen failure data will be tabulated by reason patients were found to not be eligible for the study.

7.3. By-Center

A tabulation of the number of patients enrolled by country will be presented. This tabulation will also include the pooled center counts (number of patients in each of the pooled/geographic region-based centers).

7.4. Protocol Deviations

Protocol deviations will be identified prior to unblinding, and the Per Protocol population will be determined based on a review of those deviations. A list of all protocol deviations will be tabulated and listed. In addition, a listing of subjects for whom the treatment blind was broken during the study will also be provided, if appropriate.

7.5. Demographic and Baseline Characteristics

Demographic data will be summarized for the ITT Population and the Safety Population.

- All demographic data [eg, age, gender, race, baseline MMSE, donepezil dose and duration of donepezil treatment prior to randomization (as captured from concomitant medications)] will be tabulated, along with the defined subgroups that will be used for safety or efficacy analyses (as listed in Section 6.6).
- Alzheimer's disease diagnosis information will be summarized in a similar fashion.
- Other characteristics collected on the CRF (eg, significant worsening in the past 6 months) will be summarized by the number and percentage of subjects, or as appropriate.
- Baseline values for the primary and secondary efficacy endpoints (eg, ADAS-Cog-11, ADCS-ADL, DB Total Score) will be included.
- Details of the primary caregiver (sex, relationship to patient, live with the patient) will be included in a tabulation.

7.6. Subject Inclusion and Exclusion Criteria

Patients not meeting specific eligibility criteria questions will be listed.

7.7. Medical History and Medical Conditions Present at Entry

Past and current medical conditions will be collected and coded using the Medical Dictionary for Regulatory Activities (MedDRA), using Axovant Sciences coding conventions.

A tabulation displaying medical history will be provided and will form the basis for discussion regarding the medical history of patients enrolled in this study.

7.8. Prior and Concomitant Medications

Medication verbatim text will be coded using the WHO ATC classification (http://www.whocc.no/filearchive/publications/1_2013guidelines.pdf).

Medications will be tabulated as:

- Prior medications
- Concomitant medications
  - Single-blind run-in
  - Double-blind period
- New-onset concomitant medications during the double-blind period

Medications received prior to the date of first dose of single-blind study medication are considered as prior medications. Medications will be considered as concomitant if the start date of the medication is on or after the date of first intake of investigational product or if the start date is prior to the first date of investigational product but the medication is ongoing during the treatment period in the study.

Concomitant medications will be further split by concomitant medications taken during the Single-Blind Placebo Run-In Period and concomitant medications taken during the Double-Blind Treatment Period. It should be noted that a concomitant medication can be counted in more than one period.

New-onset medications (those starting only after the date of first dose of double-blind study medication) will also be presented.

Note that multiple drug use (by preferred name) will be counted once only per subject in the tabulations. No inferential statistics are planned.

Use of non-study medications will be summarized (number and percentage of subjects) by treatment and WHO preferred name.

A listing of non-medication therapy will be provided.

7.9. Extent of Exposure

Duration of exposure for the Double-Blind Treatment Period will be calculated as follows:

- Duration of exposure in days = (Double-Blind Treatment Period stop date – Double-Blind Treatment Period start date) + 1

Total dose will be calculated as:

- Total Dose = duration of exposure × dose.

Duration of exposure and Total dose will be tabulated descriptively. Exposure also will be categorized as:

- 0 to 2 weeks
- >2 to 4 weeks
- >4 to 8 weeks
- >8 to 12 weeks
- >12 to 16 weeks
- >16 to 24 weeks
- >24 weeks

7.10. Treatment Compliance

Compliance with study drug and with donepezil (separately) will be assessed over the Double-Blind Treatment Period using drug dispensing records. Treatment compliance during the double-blind period will be computed by determining the number of tablets taken relative to the number of tablets expected. Patients are expected to take one tablet per day of each.

Treatment compliance based on the drug accountability will be calculated as follows:

- compliance (%) = (number of tablets taken)/(number of tablets expected)*100
  - number of tablets taken = the number of tablets dispensed – the number of tablets returned

- number of tablets expected = (date of the last dose-date of the first dose+1)

Summary of treatment compliance will be presented by treatment group, between Week 18 and 24 as well as overall. The number and percentage for compliance expressed as a categorical variable (<80%, ≥80% to ≤120%, and >120%) will be also presented by treatment group, for Week 18 to Week 24 as well as overall. Note that patients missing a bottle (ie, they do not return a bottle) will have compliance set to MISSING.

For each bottle of double-blind medication, 50 tablets are included. For each bottle of donepezil is 90 tablets.

7.11. Overdose

A listing of any overdose information will be provided. This will be included in the same section of the data listings as the compliance listings.

8. EFFICACY

8.1. General Considerations

Tabulation of the primary and secondary endpoints will generally be presented for the ITT, Per-Protocol, Completers Population, and subgroups as identified in Section 6.6. The primary efficacy assessments will be performed on the Observed Cases (OC) dataset. The observed cases dataset is the data that is analyzed "as is", with no imputation for missing values.

Sensitivity analyses will be performed for the primary efficacy endpoint using a multiple imputation approach as described in Section 8.3.2.3.

For the total scores for ADAS-Cog-11 and ADCS-ADL, if an endpoint has not been completed at the baseline visit or has too many missing items to be able to calculate the imputed total score or has not been completed at all, the imputation rules as shown in Table 10 will be applied.

Table 10: Imputation Rules for Pre-Treatment (Baseline) Total Scores for ADAS-Cog-11 and ADCS-ADL (Study RVT-101-3001)

Week -3 Total Score	Week 0 Total Score	Baseline Total Score
x	x	Week 0
.	x	Week 0
x	.	Week -3
.	.	Missing

8.2. Statement of the Null and Alternate Hypothesis

This study will examine the following primary hypothesis:

- H_o: There is ***NO*** difference between treatment with intepirdine and treatment with placebo with respect to changes over 24 weeks in the ADAS-Cog-11 score or in the ADCS-ADL score in patients with Alzheimer's disease.
- H_a: There ***IS*** a difference between treatment with intepirdine and treatment with placebo with respect to changes over 24 weeks in the ADAS-Cog-11 score and in the ADCS-ADL score in patients with Alzheimer's disease.

The primary comparisons of interest will be performed at the 5% level of significance. All hypothesis tests will be two-sided.

8.3. Analysis of the Primary Efficacy Endpoint

8.3.1. Primary Endpoints

The co-primary endpoints in this study are ADAS-Cog-11 score change from baseline to Week 24 and ADCS-ADL score change from baseline to Week 24.

8.3.2. Primary Efficacy Analysis

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 OC dataset using a mixed model for repeated measures (MMRM) with restricted maximum likelihood estimation, an

unstructured covariance matrix, and the Kenward-Roger approximation for denominator degrees of freedom.

- This model corrects for dropout and accounts for the fact that measurements taken on the same subject over time tend to be correlated, by using all available information on subjects within the same covariate set to come up with an estimate of the treatment effect for a dropout free population.
- For this primary analysis, no imputation of the missing values will be made or required and the data used in the analysis will be the actual observed responses at each visit.
- 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 (as defined in Section 6.6).

The region by treatment interaction term will be evaluated at the 10% level of significance; if the interaction term is found to be significant, it will be included in the MMRM model. Note, this applies to ALL MMRM models, EXCEPT the by-region Subgroup analyses, where, because the subgroup analyses by Region only include 1 region, the interaction makes no sense. The same MMRM model is also used for subgroups for the efficacy endpoints, as appropriate.

Primary 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 using the MMRM models.

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.

8.3.2.1. Further Details of the MMRM Analysis

The mixed model for repeated measures (MMRM) model will be fitted using SAS PROC MIXED procedure with restricted maximum likelihood estimation, an unstructured covariance matrix, and the Kenward-Roger approximation for denominator degrees of freedom. The unstructured covariance is the least restrictive and generally performs well with limited number of repeated measures per patient and puts no parameters on the data with respect to the covariance structure assumptions.

- Only in the unlikely circumstance that there are convergence problems with the MMRM analysis will other covariance structures be examined to resolve the convergence issue, ie, we would evaluate other additional variance-covariance structures, including compound symmetry (CS), heterogeneous compound symmetry (CSH), and auto-regressive [AR(1)]. In this eventuality, the Akaike’s Information Criterion (AIC) will be used to determine the optimal variance-covariance structure matrix for the primary comparisons.

8.3.2.2. Sensitivity Analyses of the Primary Efficacy Results

The following sensitivity analyses will be performed for the primary efficacy endpoint:

- An ANCOVA model, using observed cases, will be run for the Week 24 data. SAS Proc Mixed will be used for this analysis.
- Repeat the primary MMRM analysis (using the same model as was used for the primary assessment of ADCS-ADL, ie, excluding or including the region by treatment interaction as appropriate) for both the co-primary endpoints in the following two scenarios' regarding the ADCS-ADL translation errors observed for Question 18b and 18c, in Spain and Korea:
  - For all subject visits in Spain and Korea, when Question 18 is answered "Yes", Questions 18b and 18c will both be considered as zero when calculating total Instrumental Score in cases where no other Instrumental ADCS-ADL question is missing.
  - For all subject visits in Spain and Korea, when Question 18 is answered "Yes", then assign the total Instrumental Score as missing.
- Analysis using a Cochran-Mantel-Haenszel test with modified ridit scores to assess the differences in row mean scores between the treatment groups (CMH-RMS). *Sample* SAS code is as follows, where MMSEbase is the MMSE baseline stratum, treat is the treatment variable, and outcome (change from baseline) is the endpoint. This analysis will be performed for the Week 24 observed cases data. The actual SAS code may be modified as appropriate.

```
Proc freq data=dat;
  Tables MMSEbase*treat*outcome / cmh rms;
  Run;
```

- Analysis using a non-parametric (distribution-free) Wilcoxon Rank Sum Test and Hodges-Lehmann estimates of the treatment differences. This analysis will be performed for the Week 24 observed cases data. Confidence intervals based on the Wilcoxon Rank Sum Test will also be provided. *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 HodgesLehmann
  Run;
```

8.3.2.3. Description of Missing Data for the Primary Efficacy Endpoints

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 11. 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 11: 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	.
2	Intepirdine	xx (xx.x%)	X	X	X	X	.	.
	Placebo	xx (xx.x%)	X	X	X	X	.	.
Etc.								

8.3.2.4. Sensitivity Analysis from Imputation of Missing Efficacy Data for Primary Efficacy Endpoints

8.3.2.4.1. Last Observation Carried Forward

An ANCOVA model will be run on the Week 24 data for which any missing values at Week 24 are imputed from last non-missing value carried forward (LOCF) for a given patient. *Sample* SAS code is as follows (where basescore is the baseline score for the endpoint and center is the pooled geographic region variable [US, non-US English, West Europe, East Europe, and Rest of World]).

```
Proc Mixed data=dat;
  Class treat MMSEbase center;
  Model outcome=treat mmsebase center basescore;
  Run;
```

8.3.2.4.2. Multiple Imputation

Missing 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.

The assumption that data are missing at random (MAR) is common to the above analyses. To examine the primary endpoint with an analysis 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 or 0)
- 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 in order to obtain a monotone missing data pattern.

```
proc mi data=DATAIN out=DATAIN_MONO nimpute=100 seed=123;
    var TRT 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.

Following the imputation of the data to create a monotone missing pattern, missing values for the earliest missing value (visitnum1) are imputed BY VISIT (thus, one imputation at a time), using the covariates basescore, MMSEbase, center, time_from_diagnosis (of AD), gender, and previous score (ie, the non-missing score from previous 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. Thus, the process is repeated sequentially, 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) and second visit (Week 6) data. Note that the output dataset from the first imputation is used as the input dataset for the second imputation.

```
**First procedure imputes missing values for the first visit (Week 3);
proc mi data=DATAIN_MONO out=DATAREG1 seed=465 nimpute=1;
    by _Imputation_;
    var visitnum1 basescore MMSEbase center time_from_diagnosis gender
    previous;
    monotone regression(visitnum1);
    run;
** Second procedure imputes missing values for the second visit (Week 6)
proc mi data=DATAREG1 out=DATAREG2 seed=465 nimpute=1;
    by _Imputation_;
    var visitnum1 visitnum2 basescore MMSEbase center time_from_diagnosis
    gender  previous;
    monotone regression(visitnum2);
    run;
```

Note that the regressions do not include the treatment variable, and thus the imputed data at each time point will be based on the distribution of all data in the model rather than treatment-group specific distributions.

- A total of 100 imputed datasets will be created using these imputed data.
- Each of the 100 imputed datasets will then be analyzed using the MMRM model with effects for treatment, mmsebase, center, and basescore. 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 may be used for this summary of the analysis.

8.3.2.5. Additional Analyses of Primary Efficacy Endpoints

The primary endpoints will be analyzed in the additional manner:

- A cumulative distribution plot of changes in scores from baseline to Week 24 will be provided, with percent of patients on Y-axis and change score on X-axis.
- Time to progression, defined as a worsening (from baseline) by 1 point (or more). Worsening must be observed in 2 consecutive visits (eg, Week 6 and 12) in order to be defined as progression. Thus, the worsening must be by 1 point or more from baseline for two consecutive visits. A log-rank statistic will be provided.
- For the efficacy subgroup analyses, the same MMRM model will be used as for the primary analysis of the variable of interest. Thus, if region by treatment interaction is not significant (at 0.10) for the primary efficacy endpoint overall, then it will not be included in the model for the subgroup analyses for that variable.

- However, for the REGION subgroups, the region by treatment interaction term will necessarily be excluded.

8.4. Analysis of the Secondary Efficacy Endpoints

Continuous variables will be analyzed in a fashion similar to that of the primary endpoint, using the MMRM methodology on the observed cases dataset. The MMRM model will include the same terms as the analysis of the primary efficacy endpoints.

For those endpoints that are collected only once on-treatment (DS, RUD Lite, and EQ-5D), as there are no repeated measurements, an ANCOVA will be used; however, it is intended that the PROC MIXED SAS procedure will be used in all cases to perform the analysis, as it is capable of both the ANCOVA and MMRM analyses. The region by treatment interaction term will be evaluated at the 10% level of significance; if the interaction term is found to be significant, it will be included in the MMRM model.

Categorical outcomes will be analyzed using a logistic regression approach to test the effect of treatment on the outcomes. The model will include terms for treatment, baseline score, baseline MMSE, and Geographic Region. The odds ratios for each treatment group compared to placebo will be displayed with the 95% confidence intervals.

8.4.1. Key Secondary Endpoints

The key secondary endpoints in this study will be analyzed in the following manner:

- CIBIC+
  - Observed values at each time point using MMRM as the primary approach, with similar subgroups as for the primary endpoints. The statistical model will be fitted with terms for treatment group, visit, treatment by visit interaction, baseline MMSE, and Region (as defined in Section 6.6). The region by treatment interaction term will be evaluated at the 10% level of significance; if the interaction term is found to be significant, it will be included in the MMRM model. Note, that as no baseline CIBIC+ score was collected, there is no baseline covariate.
  - Frequency Distribution of CIBIC+ at Week 24 (including tabulation and bar chart), with percentage of patients on Y-axis and CIBIC+ rating categories (Markedly Improved through Markedly Worse on the X-axis).
  - Analyzed, for Week 24, using a CMH row means scores (modified ridit scores) test, a Hodges-Lehmann test, and a Wilcoxon Rank-Sum Test (see Section 8.3.2.2 for sample SAS code that is generally applicable for this analysis).
- Dependence Scale (DS) Total Score changes over time. The DS is collected at only one time point post-baseline. Thus, this analysis will be performed using BOTH, an observed case and LOCF. The ANCOVA will have effects for treatment, baseline DS score, baseline MMSE, and Region (as defined in Section 6.6). The region by treatment interaction term will be evaluated at the 10% level of significance; if the interaction term is found to be significant, it will be included in the MMRM model.

8.4.2. Other Secondary Endpoints

The other secondary endpoints in this study will be analyzed in the following manner:

- Dependence Scale (DS)
  - Progressors in Dependence Level, defined as patients worsening from baseline, analyzed as a categorical outcome (using logistic regression)
- NPI Total score as a shift from baseline, defined as 'No Change/ Improvement' versus 'Worsening (ie, Progressors)' (Table 12). This variable will be analyzed as a categorical outcome (using logistic regression stratified by the baseline MMSE).

Table 12: NPI Total Score Progressors Analysis (Study RVT-101-3001)

NPI Total Score Outcome	Placebo	Intepirdine 35 mg
No Change/Improvement	XX (%)	XX (%)
Progressors	XX (%)	XX (%)

-
- Falls
  - The incidence of falls will be compared, where MedDRA SOC=Injury, poisoning, and procedural complications, PT=Fall will be selected. A logistic regression model controlling for baseline stratification of MMSE will be used for this analysis. In the Phase 2 study AZ3110866, the incidence of falls in the intepirdine 35 mg treatment group was numerically lower than that in the placebo group. The pre-specified analysis of the incidence of falls in RVT-101-3001 is to assess whether the observed reduction of falls can be replicated in a larger study.
- ADAS-Cog-11 and ADCS-ADL
  - Time to progression, defined as a worsening (from baseline) by 1 point (or more). Worsening must be observed in 2 consecutive visits (eg, Week 6 and 12) in order to be defined as progression. Thus, the worsening must be by 1 point or more from baseline for two consecutive visits. A log-rank statistic will be provided.
- ADCS-ADL Basic, Instrumental, and Independence subscores
  - Changes over Time
- ADCS-ADL Communication and Engagement Factor and Outside Activities Factors
  - Changes over Time
- ADAS-Cog-13
  - Changes over Time
- The NPI scores will be analyzed as follows:
  - NPI Total score as changes over Time
  - Caregiver distress as changes over Time

  - NPI Domain score of the following Domains as changes over Time
    - Agitation/aggression
    - Anxiety
    - Apathy
    - Dysphoria
    - Irritability
  - Progressors analysis of the Psychosis Domain score
  - Progressors analysis for each of the 12 NPI domain scores.
- RUD Lite, the Total, Personal, and Instrumental domains
  - Changes from baseline to end of study
- EQ-5D
  - Total score changes from baseline to end of study (including EQ-VAS)
  - The number and percentage of patients by each response level will be summarized by visit and treatment for each of the 5 domains. This will be analyzed via a CMH RMS test stratified by baseline MMSE strata levels.
  - The response levels will be dichotomized to show number and percentage of subjects without problems (reported no problems) and with problems (slight problems, moderate problems, severe problems, and extreme problems). This will be analyzed via a logistic regression model.

8.4.3. Responder Analyses

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. The analyses will be performed based on the observed data (subjects with Week 24 assessment for the endpoint of interest).

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

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

9. SAFETY AND TOLERABILITY

The safety analysis will be descriptive in nature, and will be presented for the Safety Population. All safety data collected and captured in the eCRF will be included in data listings sorted by domain, patient and time point, or as appropriate. Mean observed values at pre-treatment and on-treatment (as well as post-treatment) will generally be tabulated by protocol-specified time points, while the number of patients with potentially clinically significant values at pre-treatment and at each endpoint will be presented. The last non-missing baseline value will be used as the baseline value for that parameter.

Adverse events will be classified using the most up-to-date MedDRA coding dictionary. Tabulations will include an overall incidence of at least one adverse event, incidence within body system, and incidence by preferred term. Each patient may only contribute once (ie, first occurrence) to each of the incidence rates, regardless of the number of occurrences.

Generally, safety data will be presented for the Placebo Run-In Period, for the Double-Blind Treatment Period, and for the post-treatment Follow-up Period. AEs will be presented for the Screening period as well.

9.1. Adverse Events

Adverse events will be tabulated by MedDRA system organ class (SOC) and preferred term, unless otherwise specified.

- Adverse events occurring prior to the first dose of single-blind study medication (ie, prior to the Placebo Run-In period) will be presented as Screening events.
- Adverse events occurring prior to the first dose of double-blind study medication will be presented as Placebo Run-In events.
- Adverse events occurring on or after the first dose of double-blind study medication, and within 7 days of the last dose of double-blind study medication, will be referred to as On-Treatment AEs (OTAEs).
- Events occurring after that will be presented as Post-Treatment. While the Post-Treatment period is defined as up to 30 days post-last-dose, it is possible some events may be reported that start after that 30 day period. These events will be included in the analysis.

AEs during the double-blind treatment period will be presented overall as well as by the subgroups as defined in Section 6.6 for OTAEs, SAEs, and OTAEs resulting in death.

Tabulations of the incidence of AEs will be presented as follows:

- Screening (using the Screening Population as the denominator).
- Single-Blind Placebo Run-In (using the Placebo Run-In Population as the denominator)
  - AEs
  - AEs resulting in discontinuation from study

  - SAEs. SAEs will be tabulated (separately) as All SAEs, Non-Fatal SAEs and Fatal SAEs.
  - Deaths
- Double-Blind Treatment Period (using the Safety Population as the denominator)
  - OTAEs
  - OTAEs by MedDRA System Organ Class Only
  - OTAEs occurring in ≥2%, 5%, and 10% in EITHER treatment arm
  - OTAEs judged by the investigator to be at least possibly related to study medication
  - OTAEs by maximum intensity (Grade 1 mild, Grade 2 moderate, and Grade 3 severe)
  - OTAEs resulting in discontinuation from study medication
  - SAEs. SAEs will be tabulated (separately) as All SAEs, Non-Fatal SAEs and Fatal SAEs.
  - SAEs judged by the investigator to be at least possibly related to study medication
  - OTAEs resulting in Death
  - OTAEs resulting in Death judged by the investigator to be at least possibly related to study medication
  - An analysis of Adverse Reactions (defined as events occurring with higher incidence in the Intepirdine arm) will be provided via a tabulation of these OTAEs.
- Post-Treatment Period (using the Safety Population as the denominator)
  - AEs
  - SAEs
  - Deaths

The time to first onset of OTAEs will be presented using the following bins of onset, where the bins are defined from the start dose of study medication. As the bins are defined by exposure, events occurring after the date of last dose are not included in this analysis.

- 0 to 2 weeks
- >2 to 4 weeks
- >4 to 8 weeks
- >8 to 12 weeks
- >12 to 16 weeks
- >16 to 24 weeks

- >24 weeks

Time to first onset for select OTAEs will be presented via Kaplan-Meier plots and a log-rank statistic will be calculated (stratified by randomization baseline MMSE score). These events will include:

- OTAE resulting in discontinuation from study medication
- SAE
- Death
- Falls (a pre-specified secondary endpoint)
  - SOC=Injury, poisoning, and procedural complications, PT=Fall. PT's of fracture and contusion may also be investigated.
- The most common events (based on incidence of 2% or more, subject to adjustment depending on actual observed incidence in the study)

The Incidence of Adverse Events Occurring after Last Dose of Double-Blind Study Medication by Time to Onset of Event (1 to 14 days, >14 to 30 days, and >30 days) will be presented. This will include all events that occur with a start-date that is 1 or more days AFTER the date of the last dose of study medication.

9.2. Clinical Laboratory

Clinical laboratory parameters will be presented using three methods:

1. Tabulations of observed values over time. Group-mean plots (mean and standard error) over time will also be provided. Reference lines for the upper and lower normal values will be provided: where there are difference reference ranges for males and females, the highest (or lowest) value will be used as the reference line.
2. Tabulations of the incidence of potentially clinically significant (PCS) values. A focus will be on new-onset PCS values, ie, patients with pre-existing PCS values at pre-treatment will not be considered to have new-onset values on-treatment. For purposes of this analysis, the most extreme (highest and lowest) value AT ANY TIME for a parameter (for a given patient) will be used. Thus, a patient may contribute to both a low PCS value for a parameter as well as a high PCS value for that same parameter. PCS laboratory values are presented for serum chemistry in Table 14 and for hematology in Table 15.
3. Shifts from baseline to on-treatment, where values are categorized as low, normal, or high according to the lab normal values. For purposes of this analysis, the most extreme on-treatment values will be used (most extreme low and most extreme high). Shifts from baseline will be based on laboratory normal ranges as provided by the central laboratory (see Appendix 1).

In addition, the incidence of OTAEs relating to hematology and serum chemistry laboratory parameters during the Double-Blind Treatment Period will be presented. These laboratory AEs will be identified from a review of the adverse events and select SOC and PTs will be included.

Table 13: Serum Chemistry Potentially Clinically Significant Criteria (Study RVT-101-3001)

Parameter	Gender	Low PCS	High PCS
Alanine aminotransferase (U/L)	Male	NA	133
	Female	NA	100
Albumin (g/L)	Male/Female	29	NA
Alkaline phosphatase (U/L)	Male	NA	323
	Female	NA	246
Aspartate aminotransferase (U/L)	Male	NA	118
	Female	NA	103
Direct bilirubin (umol/L)	Male/Female	NA	7.7
Total bilirubin (umol/L)	Male/Female	NA	30.8
Calcium (mmol/L)	Male/Female	1.99	2.91
Creatinine (umol/L)	Male	NA	174
	Female	NA	147
GGT (U/L)	Male	NA	136
	Female	NA	93
Creatinine clearance (mL/min)	Male/Female	29	NA
Random glucose (mmol/L)	Male/Female	2.9	NA
Potassium (mmol/L)	Male/Female	3.4	5.6
Sodium (nmol/L)	Male/Female	129	151

Table 14: Haematology and Differentials Potentially Clinically Significant Criteria (Study RVT-101-3001)

Parameter	Gender	Low PCS	High PCS
WBC (10^9/L)	Male/Female	2.9	100.1
Hemoglobin (g/L)	Male	99	196
	Female	99	181
Platelets (10^9/L)	Male/Female	74	NA
Absolute neutrophil count (10^9/L)	Male/Female	1.4	NA
Calculated absolute neutrophil count (10^9/L)	Male/Female	1.4	NA
Absolute lymphocyte count (10^9/L)	Male/Female	0.7	5.1
Calculated absolute lymphocyte count (10^9/L)	Male/Female	0.7	4.1

9.3. Vital Signs

Vital sign data (blood pressure [BP], pulse, and weight) will be summarized by treatment, visit, and planned time and listed by subject, visit, treatment, planned time and actual date and time. Change from baseline will also be summarized. The incidence of PCS values will be presented, with a focus on new-onset PCS values.

PCS ranges for these parameters are provided in Table 15.

Table 15: List of Potentially Clinically Significant Ranges for Vital Sign Parameters (Study RVT-101-3001)

Systolic blood pressure:

Low: < 90 and decrease ≥30 mm Hg
High: ≥ 180 and increase ≥ 20 mm Hg
Diastolic blood pressure:
Low: < 50 and decrease ≥20 mm Hg
High: >90 and increase ≥20 mm Hg
Pulse
Low: <50 bpm and Decrease ≥30 bpm
High: >100 and Increase ≥30 bpm
Weight
≥ 10% increase from baseline
≥ 10% decrease from baseline

9.4. Physical Examination

Physical examination data will be listed by patient and time point.

9.5. ECGs

ECG data will be summarized by treatment, visit, and planned time and listed by subject, visit, treatment, planned time and actual date and time. Change from baseline will also be summarized.

The ECG analysis will include a careful review of QTcF values. As part of this review, a summary of the number (percent) of patients with QTcF values in the following ranges will be provided: ≤450 msec, >450 to ≤480 msec, >480 to ≤500 msec, and >500 msec. This will be performed by visit as well as at ANY time during the double-blind treatment period (where the highest QTcF value will be used for that assessment).

The overall Investigator interpretation of ECG by visit (shifts from baseline) will be tabulated. The incidence of PCS values will be presented, with a focus on new-onset PCS values.

Table 16: List of Potentially Clinically Significant Ranges for ECG Parameters (Study RVT-101-3001)

QTcF Interval
High: ≥500 msec only
High: Increase ≥ 60 msec only
High: ≥500 and Increase ≥60

9.6. Pregnancy

If any female subjects or female partners of male subjects become pregnant during the study, a listing will be provided.

9.7. Suicidality

The following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

- Category 1 – Wish to be Dead
- Category 2 – Non-specific Active Suicidal Thoughts
- Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5 – Active Suicidal Ideation with Specific Plan and Intent
- Category 6 – Preparatory Acts or Behavior
- Category 7 – Aborted Attempt
- Category 8 – Interrupted Attempt
- Category 9 – Actual Attempt (non-fatal)
- Category 10 – Completed Suicide
- Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

Endpoints based on the above categories are defined below.

- Suicidal **ideation**: A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS.
- Suicidal **behavior**: A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS.
- Suicidal **ideation or behavior**: A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.
- Self-injurious behavior without suicidal intent.

A shift table to present changes from pre-treatment to any time post-first-dose will be assessed, with the number (and percent) of subjects showing a worsening from baseline on any of the four endpoints.

The C-SSRS was collected at screening (and the assessment was based on patient recall from the past year), while on-treatment collection was based on outcomes 'since the prior visit'. Thus, the analysis will be done in two ways:

1. Shifts from screening, where C-SSRS is relative to patient recall of the past year (using Screening value C-SSRS, and is shifts to each post-screening time point).

2. Shifts from baseline, where C-SSRS is collected relative to patient recall since the PREVIOUS visit, using Baseline value C-SSRS and is shifts to each post-baseline time point).

Table 17 presents the general manner in which these data will be tabulated.

Table 17: Sample Table Demonstrating How C-SSRS Outcomes will be Presented (Study RVT-101-3001)

Treatment	Baseline Category	Treatment Category		
		No suicidal ideation or behavior n (%)	Suicidal ideation n (%)	Suicidal behavior n (%)
Drug Name (N=xxx)	No suicidal ideation or behavior	x (%)	x (%)	x (%)
	Suicidal Ideation	x (%)	x (%)	x (%)
	Suicidal Behavior	x (%)	x (%)	x (%)
Comparator Name (N=xxx)	No suicidal ideation or behavior	x (%)	x (%)	x (%)
	Suicidal Ideation	x (%)	x (%)	x (%)
	Suicidal Behavior	x (%)	x (%)	x (%)

Notes: N = number of patients with a baseline and post-baseline C-SSRS assessment, n = number of patients in category. % = 100*n/N.
Baseline refers to [specify definition]
Suicidal Ideation includes any one of the five suicidal ideation events (Categories 1-5). Suicidal behavior includes any one of the five suicidal behavior events (Categories 6-10).
Each patient is counted in one cell only. Patients with both Suicidal Ideation and Suicidal Behavior are included in the Suicidal Behavior category.

A listing of all C-SSRS data will be provided, while a listing of patients who demonstrated a worsening on any one of the four endpoints will also be provided. These listings will include the Study Day (relative to first dose of double-blind treatment) the C-SSRS was collected.

10. PHARMACOKINETIC /PHARMACODYNAMIC ANALYSES

Plasma intepirdine and donepezil concentrations will be listed and summarized by treatment group and visit and included in the CSR.

Intepirdine population PK analyses will be described in a separate population PK/PD analysis plan and reported separately to the main CSR. Population PK analyses may pool data from the current study and prior Phase 1 and Phase 2 studies, as appropriate, and evaluate any impact of donepezil on the PK of intepirdine.

No population PK analysis is planned for donepezil. Donepezil concentrations collected before and after initiation with intepirdine will be descriptively compared, and compared to literature steady state values (mean ± SD) [Tiseo 1998] to assess a pharmacokinetic interaction:

- 5 mg Cmin=21.4 ± 3.8 and Cmax= 34.1 ± 7.3 ng/mL
- 10 mg Cmin=38.5 ± 8.6 and Cmax=60.5 ± 10.0 ng/mL

In addition, the following analyses of donepezil concentrations will also be performed:

- Comparison of donepezil concentrations

  The donepezil concentration by visit will be compared between subjects taking intepirdine and subjects taking placebo.

- Comparison of donepezil trough concentrations

  In patients with a value of ≥20 hours between time of last donepezil dose and collection of the PK blood sample, the mean donepezil concentration will be compared by visit between subjects taking intepirdine and subjects taking placebo

10.1. Exposure-Efficacy Analyses

The relationship between intepirdine exposure and the ADAS-Cog-11 score and ACDS-ADL endpoints will be assessed using a longitudinal population PK/PD model, similar to that previously described [Holford 1992]. Intepirdine population PK/PD analyses will be described in a separate population PK/PD analysis plan and reported separately to the main CSR. This analysis also will include endpoint data from the prior Phase 2 study AZ3110866 and may include donepezil trough concentrations as a covariate.

10.2. Exposure-Safety Analyses

Intepirdine exposure-safety analyses will initially consist of a series of exploratory plots (scatter plots, box-plots, or similar) to evaluate any relationship between intepirdine or donepezil exposure and frequency or severity of a select group of AEs or laboratory findings (to be determined after initial evaluation of the AE profile of the subjects enrolled in the current study). The Week 12 (Visit 6) intepirdine and donepezil pre-dose trough concentration will be used for these exploratory plots. Further model-based analyses may be conducted based on the findings of these exploratory graphical outputs, and incorporate other exposure metrics (such as $AUC_{\tau ss}$, $C_{max\text{-}ss}$, and $C_{min\text{-}ss}$).

11. CHANGES FROM THE PROTOCOL TO THIS SAP

Additional details and specifications have been included in this SAP to allow for better understanding of the intended methods. Additional analyses (including subgroup assessment) for select variables have been added. Further, key secondary endpoints were identified and multiplicity and alpha control were addressed through a hierarchical testing strategy for these endpoints.

12. REFERENCES

Brickman AM, Riba A, Bell K, et al.. Longitudinal assessment of patient dependence in Alzheimer disease. Arch Neurol. 2002;59(8):1304-1308.

Folstein MF, Folstein SE, McHugh PR. “Mini-Mental State”: A practical method for grading the cognitive state of patients for the clinician. J Psychiatry Res. 1975;12:189-198.

model. Technical report 97-04, Dept. of Statistics, The Pennsylvania State University.

Holford NH, Peace KE. Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. *Proceedings of the National Academy of Sciences.* 1992;89(23):11466-11470. Available at: http://www.pnas.org/content/89/23/11466.long. Accessed 9-feb-2017.

Mohs RC, Knopman D, Petersen RC, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord. 1997;11(Suppl 2):S13–21.

Rabin R, deCharro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-343.

Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer’s disease. Am J Psychiatry. 1984;141:1356-1364.

Rubin D, Schenker N. Multiple Imputation for Interval Estimation from Simple Random Samples with Ignorable Nonreponse. Journal of the American Statistical Association. June 1986, Vol 81, No 394.

Rubin D. Inference and Missing Data. Biometrika, Vol. 63, No. 3 (Dec., 1976), pp. 581-592

Schafer, J.L. (1997) Imputation of missing covariates under a multivariate linear mixed

Schneider L, Olin J, Doody R, et al. and the ADCS. Validity and reliability of the Alzheimer’s disease cooperative study-clinical global impression of change. Alzheimer Dis Assoc Disord. 1997;11(2):S22-S32.

Tiseo PJ, Rogers SL, Friedhoff LT. Pharmacokinetic and pharmacodynamics profile of donepezil HCl following evening administration. Br J Clin Pharm. 1998;46:Suppl 1:13-18.

Wimo A, Wetterholm AL, Mastey V, Winblad B. Evaluation of the resource utilization and caregiver time in Anti-dementia drug trials - a quantitative battery. In: Wimo A, Jonsson B, Karlsson G, Winblad B, editors. The Health Economics of dementia. London: John Wiley & Sons; 1998. Pp1. 465-499.

APPENDIX 1. CENTRAL LABORATORY REFERENCE RANGES

	SI Reporting of Reference Ranges											
Parameter	Age	Sex	Units	Ref Range Low	Ref Range High	Flag L at	Flag H at	Flag LL at	Flag HH at	Call Alert Low at	Call Alert High at	Methodology
BIOCHEMISTRY												
Alanine aminotransferase	Adult 18y	M	U/L	0	44	NA	45	NA	133**	NA	221**	Mod IFCC, H Bergmeyer
Alanine aminotransferase	Adult 18y	F	U/L	0	33	NA	34	NA	100**	NA	166**	Mod IFCC, H Bergmeyer
Albumin	Adult 18y	M/F	g/L	35	52	34	53	29**	71	19**	71	BCG Dye Binding
Alkaline phosphatase	Adult 18y	M	U/L	53	129	52	130	NA	323**	NA	646**	Mod IFCC AMP
Alkaline phosphatase	Adult 18y	F	U/L	42	98	41	99	NA	246**	NA	491**	Mod IFCC AMP
Aspartate aminotransferase	Adult 18y	M	U/L	14	39	13	40	NA	118**	NA	196**	Mod IFCC, H Bergmeyer
Aspartate aminotransferase	Adult 18y	F	U/L	14	34	13	35	NA	103**	NA	171**	Mod IFCC, H Bergmeyer
Bicarbonate	>=1y	M/F	mmol/L	20	31	19	32	15	40	NA	NA	Enzymatic/PEPC
Bilirubin, direct	All	M/F	umol/L	0.0	5.1	NA	5.2	NA	7.7**	NA	15.4**	Vanadate Oxidation, endpoint
Bilirubin, total	>=1mnth	M/F	umol/L	5.1	20.5	5.0	20.6	NA	30.8**	NA	61.6**	Vanadate oxidation
Blood Urea Nitrogen	Adult 18y	M/F	mmol/L	3.2	8.2	3.1	8.3	NA	25.0	NA	NA	Urease with GLDH
Calcium	Adult 18y	M/F	mmol/L	2.15	2.55	2.14	2.56	1.99**	2.91**	1.74**	3.11**	Cresolphthalein complexone
Chloride	Adult 18y	M/F	mmol/L	99	109	98	110	74	131	NA	NA	Ion-Selective Electrode, Diluted
Creatinine	Adult 18y	M	umol/L	62	115	61	116	NA	174**	NA	346**	Jaffe, Alkaline picric acid, Rate
Creatinine	Adult 18y	F	umol/L	44	97	43	98	NA	147**	NA	292**	Jaffe, Alkaline picric acid, Rate
Gamma glutamyl transferase	Adult 18y	M	U/L	0	54	NA	55	NA	136**	NA	271**	Modified IFCC/L.M. Shaw, et al
Gamma glutamyl transferase	Adult 18y	F	U/L	0	37	NA	38	NA	93**	NA	186**	Modified IFCC/L.M. Shaw, et al
eCreatinine Clearance(Cockcroft Gault Calc)	Adult 18y	M/F	mL/Min	60	NA	59	NA	59**	NA	29**	NA	Calculation
Glucose, Random	All	M/F	mmol/L	3.3	7.8	3.2	7.9	2.9**	20.0	2.1**	NA	Hexokinase, G-6-PD
Potassium (K)	Adult 18y	M/F	mmol/L	3.5	5.1	3.4	5.2	3.4**	5.6**	2.9**	6.1**	Ion-Selective Electrode, Diluted
Protein, total	Adult 18y	M/F	g/L	64	83	63	84	39	101	NA	NA	Weichselbaum-Biuret
Sodium (Na)	Adult 18y	M/F	mmol/L	136	145	135	146	119	151**	129**	156**	Ion-Selective Electrode, Diluted

**LL/HH flagging values are consistent with CTCAE Version 4 Grade 2 values.
Low and High Call alerts are consistent with CTCAE Version 4 Grade 3 values.

central laboratory

	SI Reporting Reference Ranges											
Parameter	Age	Sex	Age	Sex	Ref Range High	Flag L at	Flag H at	Flag LL at	Flag HH at	Call Alert Low at	Call Alert High at	Methodology
Hematology												
White blood cells	Adult 18y	M/F	10^9/L	4.5	11.0	4.4	11.1	2.9**	22.0	1.9**	100.1**	Direct Current; Radiofrequency; Light scatter
Red blood cells	Adult 18y	M	10^{12}/L	4.50	5.90	4.49	5.91	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Red blood cells	Adult 18y	F	10^{12}/L	3.80	5.20	3.79	5.21	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Hemoglobin	Adult 18y	M	g/L	130	175	129	176	99**	196**	79**	216**	Direct Current; Radiofrequency; Light scatter
Hemoglobin	Adult 18y	F	g/L	115	160	114	161	99**	181**	79**	201**	Direct Current; Radiofrequency; Light scatter
Hematocrit	Adult 18y	M	L/L	0.416	0.541	0.415	0.542	0.210	0.600	0.210	0.600	Direct Current; Radiofrequency; Light scatter
Hematocrit	Adult 18y	F	L/L	0.364	0.489	0.363	0.490	0.210	0.600	0.210	0.600	Direct Current; Radiofrequency; Light scatter
Mean corpuscular volume	Adult 18y	M/F	fL	83.0	104.0	82.9	104.1	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Mean corpuscular hemoglobin	Adult 18y	M/F	pg	26.0	34.0	25.9	34.1	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Platelets	Adult 18y	M/F	10^9/L	130	400	129	401	74**	1000	49**	NA	Direct Current; Radiofrequency; Light scatter
Absolute Count												
Absolute Neutrophil Count	Adult 18y	M/F	10^9/L	1.8	7.7	1.7	7.8	1.4**	15.4	0.9**	NA	Direct Current; Radiofrequency; Light scatter
Absolute Lymphocyte Count[1]	Adult 18y	M/F	10^9/L	1.0	4.8	0.9	4.9	0.7**	5.1**	0.4**	20.1**	Direct Current; Radiofrequency; Light scatter
Absolute Monocyte Count	Adult 18y	M/F	10^9/L	0.1	0.8	0.0	0.9	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Absolute Eosinophil Count	Adult 18y	M/F	10^9/L	0.0	0.5	NA	0.6	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Absolute Basophil Count	Adult 18y	M/F	10^9/L	0.0	0.2	NA	0.3	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Differential, Automated												
Percent Neutrophils	Adult 18y	M/F	%	40.0	70.0	39.9	70.1	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Percent Lymphocytes	Adult 18y	M/F	%	22.2	43.6	22.1	43.7	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Percent Monocytes	Adult 18y	M/F	%	2.0	12.0	1.9	12.1	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Percent Eosinophils	Adult 18y	M/F	%	0.0	4.5	NA	4.6	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter
Percent Basophils	Adult 18y	M/F	%	0.0	1.8	NA	1.9	NA	NA	NA	NA	Direct Current; Radiofrequency; Light scatter

1. Grade 2 CTCAE Absolute Lymphocyte Count value falls within ACM Normal Range.

**LL/HH flagging values are consistent with CTCAE Version 4 Grade 2 values. Low and High Call alerts are consistent with CTCAE Version 4 Grade 3 values.

	Age	Sex	SI Reporting Reference Ranges									
HEMATOLOGY			Units	Ref Range Low	Ref Range High	Flag L	Flag H	Flag LL	Flag HH	Call Alert Low	Call Alert High	Method
Differential, Manual												
Percent Neutrophils	All	M/F	%	36	66	35	67	NA	NA	NA	NA	Microscopic examination
Band form neutrophils (Band cells; band forms)	All	M/F	%	0	7	NA	8	NA	NA	NA	NA	Microscopic examination
Percent Lymphocytes	All	M/F	%	21	51	20	52	NA	NA	NA	NA	Microscopic examination
Percent Monocytes	All	M/F	%	4	15	3	16	NA	NA	NA	NA	Microscopic examination
Percent Eosinophils	All	M/F	%	0	6	NA	7	NA	NA	NA	NA	Microscopic examination
Percent Basophils	All	M/F	%	0	2	NA	3	NA	NA	NA	NA	Microscopic examination
Percent Metamylocytes	All	M/F	%	0	0	NA	1	NA	NA	NA	NA	Microscopic examination
Percent Myelocytes	All	M/F	%	0	0	NA	1	NA	NA	NA	NA	Microscopic examination
Percent Promyelocytes	All	M/F	%	0	0	NA	1	NA	NA	NA	NA	Microscopic examination
Percent Blasts	All	M/F	%	0	0	NA	1	NA	NA	NA	NA	Microscopic examination
Calculated Absolute Neutrophils	All	M/F	10^9/L	1.8	7.7	1.7	7.8	1.4**	15.4	0.9**	NA	Microscopic examination
Calculated Absolute Lymphocytes[1]	All	M/F	10^9/L	1.0	4.8	0.9	4.9	0.7**	4.1**	0.4**	20.1**	Microscopic examination
Calculated Absolute Monocytes	All	M/F	10^9/L	0.1	1.1	0.0	1.2	NA	NA	NA	NA	Microscopic examination
Calculated Absolute Eosinophils	All	M/F	10^9/L	0.0	0.5	NA	0.6	NA	NA	NA	NA	Microscopic examination
Calculated Absolute Basophils	All	M/F	10^9/L	0.0	0.2	NA	0.3	NA	NA	NA	NA	Microscopic examination
Percent Unclassified Cells	All	M/F	%	0	0	NA	1	NA	NA	NA	NA	Microscopic examination
Reactive Lymphocytes	All	M/F	NA	NA	NA	NA	Few	NA	NA	NA	NA	Microscopic examination
Nucleated RBC's	All	M/F	%	0	0	NA	1	NA	NA	NA	NA	Microscopic examination
Atypical Lymphocytes	All	M/F	NA	NA	NA	NA	10%	NA	NA	NA	NA	Microscopic examination
Platelet Estimate	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination
Toxic Granulation	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination
Smudge Cells	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination
Hypersegmented Neutrophils	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination
Giant Platelets	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination
Bizarre Platelets	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination
Diff Comment	All	M/F	NA	NA	NA	NA	NA	NA	NA	NA	NA	Microscopic examination

1. Grade 2 CTCAE Absolute Lymphocyte Count value falls within ACM Normal Range.

LL/HH flagging values are consistent with CTCAE Version 4 Grade 2 values. Low and High Call alerts are consistent with CTCAE Version 4 Grade 3 values.