STATISTICAL ANALYSIS PLAN

Title:	A Phase 2b, double-blind, randomized, placebo-controlled study of RVT-101 in subjects with dementia with Lewy bodies (DLB)
Sponsor	Axovant Sciences Ltd.
Compound Name:	RVT-101
Protocol Number	RVT-101-2001
Indication	Dementia with Lewy bodies (DLB)
Development Phase	2b
IND #	127,379
Version/ Effective Date:	Version 1.0 14 November 2017
Axovant Sciences Study Director	Ilise Lombardo, MD Telephone: 646-787-5388 Email: ilise.lombardo@axovant.com

SIGNATURE PAGE

Eric Floyd, MS, MBA, PhD
Senior Vice President, Global
Regulatory Affairs

eric floyd (Nov 14, 2017)

_____
Signature

Nov 14, 2017

_____
Date

Ilise Lombardo, MD
Senior Vice President, Clinical
Research

_____
Signature

Nov 14, 2017

_____
Date

Divakar Sharma, PhD
Vice President, Biometrics

divakar sharma (Nov 14, 2017)

_____
Signature

Nov 14, 2017

_____
Date

Mark Harnett, MS
Lead Project Biostatistician

Mark D Harnett
Mark D Harnett (Nov 14, 2017)

_____
Signature

Nov 14, 2017

_____
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. LIST OF ABBREVIATIONS

Table 1:   List of Abbreviations

Abbreviation	Term
ADAS-Cog-11	Alzheimer's Disease Assessment Scale (11-item)
ADCS-ADL	Alzheimer's Disease Cooperative Study – Activities of Daily Living
AE	Adverse Event
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
CAF	Clinician Assessment of Fluctuation
CDR	Cognitive Drug Research
CIBIC+	Clinician's Interview-Based Impression of Change – plus caregiver interview
CMH	Cochran Mantel-Haenszel
COWAT	Controlled Oral Word Association Test
CRF	Case Report Form
CSI	Circadian Sleep Inventory
CSR	Clinical Study Report
DOB	Date of Birth
DLB	Dementia with Lewy Bodies
DS	Dependence Scale
EQ-5D-5L	EuroQol Five Dimensions Questionnaire, Five Level Version
FAS	Full Analysis
LBCRS	Lewy Body Composite Risk Score
LLN	Lower Limit of Normal
LOCF	Last observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities Terminology
MMRM	Mixed model for repeated measures
MMSE	Mini Mental State Examination
N	Total Sample Size
NEVHI	North-East Visual Hallucinations Interview
NPI	Neuropsychiatric Inventory
OC	Observed Cases
PCS	Potential Clinical Significance
POA	Power of Attention

Abbreviation	Term
PK	Pharmacokinetic
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-101-2001, a Phase 2b, double-blind, randomized, placebo-controlled study of intepirdine in subjects with dementia with Lewy bodies (DLB).

Study measurements and assessments, planned statistical methods, and derived variables are summarized in this plan. Planned tables, figures, and listings are specified. All decisions regarding final analyses, as defined in this SAP document, have been made prior to un-blinding the study database.

Note that study drug, Intepirdine, is also referred to (eg, in the study protocol) as RVT-101. For purposes of this SAP, as well as in the analysis tables, figures, and listings, the study drug is referred to as "Intepirdine". Doses will be designated as:

- Intepirdine 70 mg

- Intepirdine 35 mg

- Placebo.

A Pooled Intepirdine group is defined as the combination (pooling) of the Intepirdine 70 mg arm and the Intepirdine 35 mg arm.

3. STUDY DESIGN

This is a multi-center, double-blind, randomized, placebo-controlled, parallel-group study in patients with probable DLB who are either not receiving therapy for DLB or are receiving a stable regimen of standard therapy for DLB. The efficacy and safety of Intepirdine at doses of 70 mg and 35 mg daily will be evaluated over a 24-week treatment period. All subjects who are on stable doses of other therapies to treat cognitive deficits and/or hallucinations associated with DLB will continue to remain on them for the duration of the study.

Approximately 240 subjects will be randomized. The randomization ratio will be 1:1:1 (70 mg Intepirdine: 35 mg Intepirdine: placebo). Double-blind treatment will be administered as follows:

- 70 mg Intepirdine (Test therapy): Intepirdine 70 mg given as two tablets of intepirdine 35 mg

- 35 mg Intepirdine (Test therapy): Intepirdine 35 mg given as one tablet of intepirdine 35 mg and one tablet of placebo

- Placebo (Reference): Matching placebo, given as two tablets of placebo

The primary endpoint (UPDRS Part III) and the secondary endpoints (ADAS Cog 11 and CIBIC+) will be measured after 24 weeks of treatment. Study participation will last approximately 32 weeks: 0 to 4 weeks for Screening, a 2-week Single-Blind Run-In Period to evaluate baseline status, a 24-week randomized, double-blind, placebo-controlled Treatment Period and a 2-week Safety Follow-up Period for subjects who do not enter the extension study RVT-101-2002.

There will be weekly clinical assessments for the first two weeks of double-blind treatment, bi-weekly assessments thereafter until 12 weeks post-randomization and every six weeks thereafter. For certain visits (Visits 5, 7 and 9), subjects may have the option of whether to have assessments performed at the clinical study site or by a trained, visiting nurse in their own home.

An independent Safety Monitoring Committee (SMC) will review interim safety data accumulated after approximately 30 subjects have completed 4 weeks of double-blind treatment and throughout the study at points specified in the SMC Charter. The SMC will provide their recommendation regarding the acceptability of reducing the visit frequency by omitting certain visits (Visits 5, 7 and 9) for both newly enrolled subjects and subjects active in the study at the time of the recommendation. Study enrollment will not be stopped or slowed to wait for SMC's recommendation and will proceed as planned with all visits until the SMC recommendation is made. Site will be formally notified of the SMC's recommendation regarding whether Visits 5, 7 and 9 can be omitted. [The meeting did occur and the SMC recommended Visits 5, 7, and 9 be omitted.]

After completion of Visit 12, the subject may be considered for enrollment in the extension study RVT-101-2002. If the subject does not enter the extension study, he/she will return for the Safety Follow-up Visit (Visit 13). Subjects who enter the extension study will not be required to complete Visit 13.

A study schema is shown in Figure 1 and study scheduled assessments are presented in Table 2.

Figure 1:   Study Schematic (Study RVT-101-2001)

Abbreviations: qd = daily; V = visit; WK = week

[1] Subjects will be given the option of whether to have assessments performed at the clinical study site or by a trained, visiting nurse in their own home. Also, if and when the DSMB approves a reduced visit frequency, Visits 5, 7 and 9 will be skipped.

[2] Safety Follow-up Visit (V13) is conducted only for subjects who do not enter the extension study.

Table 2:   Time and Events Schedule (Study RVT-101-2001)

Study Period	Screening	Run-in	Baseline	Treatment									Follow-up[3]	
Study Visit Number	V1	V2	V3[1]	V4	V5[2]	V6	V7[2]	V8	V9[2]	V10	V11	V12	V13	ET
Study Week	W(-6)	W(-2)	W0	W1	W2	W4	W6	W8	W10	W12	W18	W24	W26	ET
Study Day relative to Baseline unless specified	Up to 28 days before V2	14 +5 days before V3	0	7 ± 5	14 ± 5	28 ± 5	42 ± 5	56 ± 5	70 ± 5	84 ± 5	126 ± 5	168 ± 5	14 to 19 days after V12 or ET	
Informed consent	X													
Inclusion/Exclusion criteria	X													
Demography	X													
Medical History	X													
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Blood alcohol and urine drug screen	X	X	X											
C-SSRS	X		X	X		X		X		X	X	X	X	X
Randomization			X											
Dispense IP[5]		X	X			X		X	X	X	X			
Assess IP compliance				X	X	X	X	X	X	X	X	X		X
Physical exam[6]	X	X	X	X		X		X			X	X	X	X
Complete neurological exam	X		X									X	X	X
MRI or CT[7]	X													
12-Lead ECG	X	X	X	X		X		X		X	X	X	X	X
Vital signs[8]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Review Adverse Events	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Questionnaire for symptoms of potential orthostasis	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Serum OR urine pregnancy test[9]	X		X									X	X	X
Hep B and Hep C screen[10]	X													
TSH, vitamin B_{12}, syphilis serology[11]	X													
Serum chemistry	X	X	X			X		X		X	X	X	X	X
Hematology	X	X	X			X		X		X	X	X	X	X

Study Period	Screening	Run-in	Baseline	Treatment										Follow-up[3]	
Study Visit Number	V1	V2	V3[1]	V4	V5[2]	V6	V7[2]	V8	V9[2]	V10	V11	V12	V13	ET
Study Week	W(-6)	W(-2)	W0	W1	W2	W4	W6	W8	W10	W12	W18	W24	W26	
Study Day relative to Baseline unless specified	Up to 28 days before V2	14 +5 days before V3	0	7 ± 5	14 ± 5	28 ± 5	42 ± 5	56 ± 5	70 ± 5	84 ± 5	126 ± 5	168 ± 5	14 to 19 days after V12 or ET	
Urinalysis	X	X	X					X		X	X	X	X	X
Intepirdine level						X		X						
MMSE	X	X	X											
LBCRS	X													
ADAS-Cog 13			X							X		X		
CIBIS			X											
CIBIC+						X		X		X	X	X		
CDR[4,12]	X	X	X			X		X		X	X	X		
COWAT	X	X	X			X		X		X	X	X		
ADCS-ADL			*X*							*X*		*X*		
NPI *Parts A, B, D & E*			X							X	X	X		
NEVHI			X							X		X		
EQ-5D-*5L*			X									X		
DS			X									X		
CAF	X		X					X		X		X		
UPDRS-III			X	X		X						X	X	
CSI			X							X		X	X	

Abbreviations: ADAS-Cog = Alzheimer's Disease Assessment Scale; ADCS-ADL = Alzheimer's Disease Cooperative Study – Activities of Daily Living; CAF = Clinician's Assessment of Fluctuation; CDR = Cognitive Drug Research computerized assessment; CIBIS = Clinician's Interview-Based Impression of Severity; CIBIC+ = Clinician's Interview-Based Impression of Change Plus caregiver Interview; COWAT = Controlled Oral Word Association Test; CSI = Circadian Sleep Inventory; C-SSRS = Columbia-Suicide Severity Rating Scale; CT = computed tomography; DS = Dependence Scale; ECG = electrocardiogram; EQ-5D-5L = EuroQual five dimensions questionnaire, five level version; Hep = hepatitis; IP = investigational product; LBCRS = Lewy Body Composite Risk Score; MMSE = mini-mental state examination; MRI = Magnetic Resonance Imaging; NEVHI = North East Visual Hallucinations Inventory; NPI = Neuropsychiatric Inventory; TSH = thyroid stimulating hormone; UPDRS-III = Unified Parkinson's Disease Rating Scale, Part III

1. The Single-Blind Run-In Period may be extended for 1 to 10 days) for subjects who do not meet MMSE stability criterion for randomization. If, after the first extension to the Single-Blind Run-In Period, the subject still does not meet the MMSE stability criterion for randomization, the run-in period may again be extended for an additional 1 to 10 days . No more than 2 extensions to the Single-Blind Run-In Period will be allowed. Subjects who require one extension to the run-in period will have a new Baseline visit occurring no more than 10 days after the originally intended Baseline at Visit 3. Subjects who require two extensions to the run-in period will have a new Baseline visit occurring no more than 17 days after the originally intended Baseline at Visit 3. No new IP will be dispensed for the extended run-in periods; subject should be instructed to continue taking blinded IP dispensed at Visit 2.

2. After the first approximately 30 subjects have completed 4 weeks of treatment and throughout the study at points specified in the SMC Charter, the SMC will meet and determine if a reduced visit frequency is acceptable. If and when the SMC approves a reduced visit frequency, Visits 5, 7 and 9 will be skipped. Subjects may have the option of whether to have assessments for Visits 5, 7 and 9 performed at the clinical study site or by a trained, visiting nurse in their own home.
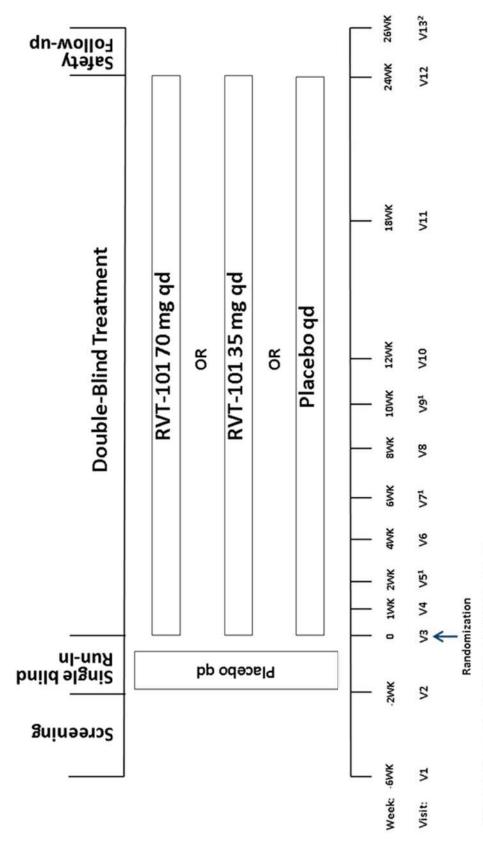
3. The Follow-Up Visit (V13) is required only for subjects who do not enter the extension study.

4. If, during the visit, the subject is unable or, in the judgment of the investigator, unlikely to be able to complete the cognitive assessments, the testing may be rescheduled within the visit windows.

5. The bottles and pill count should be checked for compliance at V4, V5, V7, and V9, but no new medication will be dispensed at those visits.

6. Full physical examination at Screening (V1), V3, V10, V12, V13and ET; abbreviated physical examination at V2, V4, V6, V8 and V11.

7. MRI or CT scan will be performed between V1 and V2 if no scan has been performed within the previous 12 months. These scan findings must be consistent with the diagnosis of DLB without any other clinically significant pathologies.

8. Vital signs will include systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, temperature, and body weight at each visit and height at screening. Postural changes in blood pressure and heart rate will also be assessed at every visit.

9. Required only for women of child bearing potential.

10. If these tests were performed within 3 months prior to the planned first dose of investigational product, testing is not required. Records must be present in the subject's source documents.

11. If these tests were performed within 12 months prior to the planned first dose of investigational product, testing is not required. Records must be present in the subject's source documents.

12. The CDR computerized assessment performed at Visits 1 and 2 will be training sessions.

4.   STUDY OBJECTIVES AND ENDPOINTS

Study objectives and endpoints are described in Table 3.

Table 3:   Objectives and Endpoints (Study RVT-101-2001)

Objective	Endpoint
Efficacy Primary	
• To assess the effects of Intepirdine versus placebo on the Unified Parkinson's Disease Rating Scale – Part III (UPDRS-III) after 24 weeks of treatment	Change from baseline in the UPDRS- III total score at Week 24
Efficacy Secondary	
• To assess the effects of Intepirdine versus placebo on cognition, including amnestic aspects of cognition, as measured by the ADAS-Cog-11 after 24 weeks of treatment	Change from Baseline in ADAS-Cog-11 at Week 24
• To assess the effects of Intepirdine versus placebo on global function as measured by the CIBIC+ after 24 weeks of treatment	Total score of the CIBIC+ at Week 24
Tertiary Endpoints	
• To assess the effects of Intepirdine versus placebo on activities of daily living as measured by the ADCS-ADL scale after 24 weeks of treatment	Change from Baseline in the ADCS-ADL at Week 24
• To assess the effects of Intepirdine versus placebo on the Basic and Instrumental subscores of the ADCS-ADL scale after 24 weeks of treatment	Change from Baseline in the Basic and Instrumental subscores of the ADCS-ADL at Week 24
• To assess the effects of Intepirdine versus placebo on the Unified Parkinson's Disease Rating Scale – Part III, 5-item subscale (UPDRS-5) after 24 weeks of treatment	Change from baseline in the UPDRS-5 total score at Week 24
• To assess the effects of Intepirdine versus placebo on cognition as measured by the composite z-score of the 7-domains of the CDR computerized assessment system after 24 weeks of treatment (CDR System domains include Power of Attention, Continuity of Attention, Quality of Working Memory, Quality of Episodic Memory, Speed of Memory, Cognitive	Change from Baseline in the composite z-score combining the 7 domains of the CDR computerized assessment system at Week 24

Objective	Endpoint
Reaction Time and Reaction Time Variability)	
• To assess the effects of Intepirdine versus placebo on working memory as a measure of executive function as assessed by the COWAT after 24 weeks of treatment	Change from Baseline in COWAT score at Week 24
• To assess the effects of Intepirdine versus placebo on cognitive function as measured by a composite z-score combining the 7 domains for the CDR computerized assessment system and the COWAT after 24 weeks of treatment	Change from Baseline in composite z-score combining the 7 CDR System domains plus the COWAT at Week 24
• To assess the effects of Intepirdine versus placebo on hallucinations and delusions as measured by a 2-item subscore on the NPI which is the sum of the scores for the hallucinations and delusions domains (Parts A and B) after 24 weeks of treatment	Change from Baseline in the sum of Parts A and B of the NPI at Week 24
• To assess the effects of Intepirdine versus placebo on visual hallucinations as measured by the NEVHI after 24 weeks of treatment	Change from Baseline on the total severity score and distress score of the NEVHI score at Week 24 for subjects with visual hallucinations
• To assess the effects of Intepirdine versus placebo on fluctuations in cognition and on confusion using the CAF after 24 weeks of treatment	Change from Baseline in CAF severity scores at Week 24 for subjects with cognitive fluctuations
• To assess the effects of Intepirdine versus placebo on subject dependence using the DS after 24 weeks of treatment	Change from Baseline on DS at Week 24
• To assess the safety and tolerability of Intepirdine after 24 weeks of treatment	AEs, physical examinations, vital signs (including orthostatic changes in BP and HR), ECGs, routine clinical laboratory assessments, questionnaire for symptoms potentially associated with orthostasis, and C-SSRS.
Exploratory	
• To assess the effects of Intepirdine versus placebo on quality of life as measured by the EQ-5D-5L after 24 weeks of treatment	Change from Baseline on EQ-5D-5L at Week 24
• To assess the effects of Intepirdine versus placebo on sleep-related behaviors as measured by the modified CSI	Change from Baseline on CSI score at Week 24

Objective	Endpoint
• To assess the effects of Intepirdine versus placebo on each domain of the CDR Computerized Assessment System after 24 weeks of treatment	Change from Baseline on each domain of the CDR computerized assessment system at Week 24
• To assess the effects of Intepirdine versus placebo on depression and anxiety as measured by a 2-item subscore on the NPI, which is the sum of the scores for the depression/dysphoria and anxiety domains (Parts D and E) after 24 weeks of treatment	Change from Baseline in sum of Parts D and E of the NPI at Week 24
• To estimate the pharmacokinetic (PK) parameters of Intepirdine and explore relationships to efficacy or safety endpoints, as appropriate	Plasma concentrations and PK parameters ($AUC_{\tau ss}$, C_{max-ss} and C_{min-ss}) for each subject
Safety	
• To assess the safety and tolerability of Intepirdine after 24 weeks of treatment	Incidence of AEs, and changes in physical examinations, vital signs (including orthostatic changes in BP and HR), ECGs, routine clinical laboratory assessments, questionnaire for symptoms potentially associated with orthostasis, and C-SSRS.

5.   DESCRIPTIONS AND SCORING OF EFFICACY ENDPOINTS

5.1.   Unified Parkinson's Disease Rating Scale – Part III (UPDRS-III)

The UPDRS-III is a widely used scale designed to follow the longitudinal course of Parkinson's disease. It is completed by a trained and qualified clinician, based on clinical examination, observation and questioning of the subject. The UPDRS-III is the primary endpoint in this study.

- UPDRS-III is a motor evaluation, consisting of 27 items each scored on a scale of 0 to 4, with higher scores indicating greater severity of the motor symptom. The Part III total is the sum of the 27 items, with scores ranging from 0 to 108.

- The UPDRS-5 is a subscale of UPDRS-III, and combines scores from 5 questions:

  – rest tremor (5 items)

  – action tremor (2 items)

  – rigidity (5 items)

  – facial expression

  – bradykinesia.

  Thus, UPDRS-5 is the sum of these 14 items, with scores ranging from 0 to 56, with higher scores indicating greater severity of the motor symptoms.

Missing items will be handled in the following manner:

For UPDRS-III, if ≥ 8 items are missing (of the 27 questions), the total score will be set to missing. If 7 or fewer items are missing, individual item scores that are missing will be imputed to be the AVERAGE of the non-missing item scores for that patient.

For instance, if a patient is missing one item, the AVERAGE of the 26 non-missing item scores will be obtained and imputed as the score for that one missing item.

For the subscale UPDRS-5, if ANY individual item is missing, the total will be set to missing.

5.2.   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 is a secondary endpoint in this study.

- 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.

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:

3. Word Recall

4. Commands

5. Constructional Praxis

6. Naming Objects and Fingers

7. Ideational Praxis

8. Orientation

9. Word Recognition

10. Remembering Test Instructions

11. Comprehension

12. Word Finding Difficulty

13. 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 4):

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

Table 4: **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)

For both the ADAS-Cog 11 and ADAS-Cog 13, 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 5.

Table 5:   **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)

5.3.   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 CIBIC+ is a secondary endpoint in this study. CIBIS is the baseline version.

The CIBIC+ 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. There is no 'change from baseline' calculated (eg, using the CIBIS). The CIBIC+ is analyzed by-visit using the observed values.

5.4.   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 is the sum of questions 7-23 and 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 be 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 6):

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

Table 6: **Calculation of Imputed ADCS-ADL Instrumental Score (Study RVT-101-2001)**

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.

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

5.5.   Controlled Oral Word Association Test (COWAT)

The COWAT, which is subtest of the Multilingual Aphasia Examination (Benton, 1994), is a verbal fluency test that measures spontaneous production of words beginning with some designated letter. The COWAT uses the three-letter set (F, A, S) to assess phonemic fluency. Individuals are given 1 minute to name as many words as possible beginning with one of these three letters. The procedure is then repeated for the remaining two letters.

The COWAT takes approximately 5 to 10 minutes to administer.

The COWAT total score is calculated as the sum of the number of correct (acceptable) words, as follows:

- The number of correct (acceptable) words starting with the letter **F** PLUS The number the sum of correct (acceptable) words starting with the letter **A** PLUS The number the sum of correct (acceptable) words starting with the letter **S**.

If any one (or more) of the number of correct (acceptable) words is missing, the COWAT will be set to missing.

5.6.   Neuropsychiatric Inventory (NPI)

The full NPI (Cummings 1994) is a 12-item behavior rating scale composed of a structured interview of the caregiver, which assess psychiatric disturbance. In this study, only Parts A, B, D & E of the NPI will be performed, which assess delusions, hallucinations, depression/dysphoria and anxiety, respectively. Both the frequency and the severity and caregiver distress of each behavior are determined. This scale will be administered by a rater who is trained for this study.

Scores are assigned to questions as demonstrated in Table 7. 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 7:   Scores for NPI Individual Questions (Study RVT-101-2001)

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.

For each behavioral domain (Parts A, B, C and E), the total score for that domain is calculated as the frequency score times the severity score.

There are two sub-scores that will be analyzed from the NPI:

- The 2-item sub-score, which is the sum of the total scores for Parts A and B (delusions and hallucinations domains)

- The 2-item sub-score, which is the sum of the total scores for Parts D & E (depression and anxiety domains)

- The Distress Scores will be presented as shifts from baseline to each time point, where values are rated as 'worsened', 'no change', or 'improved' from baseline distress score.

5.7.   North East Visual Hallucinations Inventory (NEVHI)

The NEVHI is a semi-structured interview that assesses visual hallucinations in elderly patients with visual and/or cognitive impairments (Mosimann 2008). The NEVHI has 3 sections. It will take approximately 3 minutes for subjects without visual symptoms and up to 12 minutes for subjects with hallucinations. Subjects must have a score of ≥1 on Section 1 (Phenomenology) in order to have been assessed for Section 2 (Severity) and Section 3 (Distress).

If there are no hallucinations identified (and thus the Phenomenology section is scored as a 0), then both severity and distress scores will also be 0.

- Section 1: Phenomenology. Screens for the presence of visual hallucinations and assesses hallucination phenomenology. Scoring ranges from 0 (no hallucinations) through 6. The following are the 6 types of phenomenology that can be reported (it is possible that an individual subject may have up to all 6 of these). Within Section 1, the score can range from 0 (all 'NO' response) up to 6 (all 'YES' questions).

    - 1.1 - Complex visual hallucinations

    - 1.2 - Illusions

    - 1.3 - Feeling of presence

    - 1.4 - Passage of Shadow

    - 1.5 - Other visual symptoms

    - 1.6. - Simple hallucinations

- Section 2: Severity. Assesses the severity of hallucinations by rating temporal aspects (frequency and duration) of hallucinations.

- Section 3: Distress. Assesses emotion, cognition, and behavior associated with the most prominent recurrent visual hallucination during the month prior to the assessment.

Missing Phenomenology data will be handled as follows:

- A Phenomenology section MUST have a non-missing total score in order for severity and distress scores to be calculated.

- If the Phenomenology section is completely missing, then the entire scale (including both severity and distress scores) is set to missing for all phenomenology types.

5.7.1.   Section 2 Scoring (Severity of Hallucinations)

Table 8 provides the scoring methodology for Section 2 (Severity) for the NEVHI. For each phenomenology identified in Section 1 (except for 1.5, Other visual symptoms), a frequency (never, less than once a week, 1-6 time a week, daily) will be assigned a score. Further, the duration (less than 5 minutes, 5 min to 2 hours, longer than 2 hours) will also be assigned a score.

Table 8:   Frequency and Duration Scoring for Section 2 (Severity) of the NEVHI (Study RVT-101-200)

	Answers in Section 1*:					
	1.1 CVH	**1.2 Illusion**	**1.3 Presence**	**1.4 Passage**	**1.5 Other**	**1.6 SVH**
2.1 How often have you had hallucinations** in the last month? (Frequency)					Not Calculated for this Phenomenology	
0 Never						
1 Less than once a week						
2 1-6 times a week						
3 Daily						
2.2 How long do your hallucinations** usually last? (Duration)						
1 Less than 5 minutes						
2 5 min to 2 hours						
3 Longer than 2 hours						
2.3 (for patients with multiple phenomenology): what hallucination has the highest impact on you?) (indicate with 1)						

* Section 1.1 refers to Complex visual hallucinations (CVH), 1.2: Illusions, 1.3: Passage of shadow, 1.4: Feeling of Presence, 1.5: Other, 1.6: Simple hallucinations (SVH).
** refers to simple/complex hallucination, feeling of presence, passage of shadow etc.

A Severity score is quantified for each hallucination experience as a product of its frequency and duration. The Severity score can range from 0 to 9, with 9 being the worst score.

- Severity score = frequency score x duration score

Missing frequency or duration values will be handled as follows:

- If the frequency is missing for a given phenomenology, the Severity Score for that phenomenology will be set to 0.

- If the duration is missing BUT the frequency is non-missing for a given phenomenology, the duration score should be set to "1" (less than 5 min).

The Severity score is only calculated for hallucinations (CVH, Illusion, Passage, Presence, SVH). Hence severity score is NOT calculated for Other Visual Symptoms (1.5 Other in Table 8).

Recurrent CVH are known to have the strongest impact on patients followed by recurrent Illusions (Urwyler 2014). Each **Severity score** is thus adjusted using a multiplication factor as shown in Table 9.

- Adjusted Severity Score for each Phenomenology = Severity Score (Table 8) x Multiplication factor (Table 9)

Table 9:   Multiplication Factor for each Severity Score for Each Phenomenology (Study RVT-101-2001)

	1.1 CVH	1.2 Illusion	1.3 Presence	1.4 Passage	1.6 SVH
Multiplication factor	4	3	1	1	1

The Total Severity Score is the sum of the 5 phenomenology Adjusted Total Severity scores (and thus the Total Severity Score ranges from 0 to 90, with higher scores indicating worse overall severity of hallucination phenomena).

5.7.2.   Section 3 Scoring (Distress Caused by the Hallucinations)

Section 3 applies to the phenomenology having the highest impact on the patient, where impact is determined in response to Question 2.3 in Table 8). If there is only one phenomenology present and the response to Question 2.3 is missing, the one phenomenology that is present will be used for the distress scoring. There are 6 questions relating to the distress, broken up into 3 areas:

- Emotions
  - Do you find the content of your hallucinations irritating?
  - Do you find the content of your hallucinations distressing?
- Cognitions
  - Does the content of your hallucinations make you worry that you are losing your mind?
  - While you are having a hallucination do you ever believe it is real?
- Behaviors
  - Do you ever act out on the content of your hallucinations?
  - Are you able to ignore your hallucinations, i.e. make them stop?

The Distress Score is the sum of all 6 questions. The minimal distress score is 0 and the highest (worst) distress score is 18.

Missing responses for the Distress Score will be handled as follows:

- If only 1 response is missing, the value for the response will be imputed as the average of the remaining 5 non-missing questions.
- If >1 response is missing, the Distress Score will be set to missing.

5.7.3. Analysis of NEVHI Outcomes

Data will be presented in two ways:

1. Number (%) of patients:

    a. worsening from baseline
    b. remaining unchanged from baseline
    c. improving from baseline.

This would be performed for each score separately (Total Severity Score and Distress Score). Patients with NO hallucinations (a Phenomenology score of 0) at a given visit would have a score of 0 for Severity and for Distress at that visit.

2. Mean changes over time in the Section 1 Total Score, the Total Severity Score and Distress Score.

5.8. Clinician Assessment of Fluctuation (CAF)

The CAF assesses fluctuating cognition and impaired consciousness over the previous month (Walker 2000). The CAF is one of the measures of cognitive fluctuation recommended by the DLB Consortium when applying the DLB diagnostic criteria (McKeith 2005).

The CAF is a 2-item questionnaire that asks about fluctuating cognition and fluctuation confusion, with each question reported as Yes, No, and Not Known.

- Fluctuating Cognition: Does the patient ever have spontaneous impaired alertness and concentration, ie, appear drowsy but awake, look dazed, not be aware of what is going on around? (Clear examples demonstrating impaired consciousness with variations in performance/cognition are required to receive a positive rating.) Have these episodes occurred within the last month?

- Fluctuation Confusion: Has the level of confusion experienced by the patient tended to vary a lot recently from day to day or week to week? Has it become worse, then improved for a while, ie, been up and down? (Significant fluctuation is regarded as present if distinct examples of differences in performance/cognition can be given on at least two occasions over the past month.)

If fluctuation is reported as 'yes', frequency (Table 10) and duration (Table 11) scores are assessed. These scores (frequency and duration) are multiplied to achieve a severity score (Van Dyk 2015).

- If the clinician assessment of fluctuation is that the patient does not have fluctuation, the patient is assigned a severity score of 0 for that question.

Table 10: Frequency of Fluctuating Cognition or Confusion Scores (Study RVT-101-2001)

Frequency of fluctuating Cognition or Confusion	
1	1 per month
2	Monthly – Weekly

| 3 | Weekly – Daily |
| 4 | $\geq$ Daily |

Table 11: **Duration of Fluctuating Cognition or Confusion Scores (Study RVT-101-2001)**

Duration of fluctuating Cognition or Confusion	
0	seconds
1	< 5 minutes
2	5 minutes – 1 hour
3	> 1 hour
4	> 1 Day

If either frequency or duration is missing for either question, the severity score is set to missing.

These data will be analyzed as follows for each question separately, by visit:

- The number (%) of patients with cognition (or confusion) will be tabulated..

- For each question, changes over time in severity scores.

- The number (%) of patients with EITHER a 'yes' on cognition OR a 'yes' on confusion (or a YES on both) vs those with a 'no' on BOTH.

5.9.   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 via summing the scores, with a range from 0 to 15 with higher scores indicating greater dependency.

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

5.10.   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).

5.11.   Modified Circadian Sleep Inventory – (CSI)

The Circadian Sleep Inventory is a two-part questionnaire that assesses sleep-related problems and behaviors. In this study, modifications have been made to remove questions assessing sleep apnea and restless leg syndrome.

The CSI has a Likert scale (How would you rate the overall quality of your subject's sleep in previous 4 weeks?).

The modified CSI also contains 12 questions that are answered by the caregiver about the subject's sleep-related behaviors, and an overall sleep quality rating. For each question, both the frequency and the severity of each behavior are determined and are multiplied. The CSI total score will be calculated as the sum of the scores for all questions. Note that if an answer to an item is 'no', the item score is zero. If the answer to an item is 'Don't know', the item score is set to missing.

A single missing value is allowed, with the mean of the 11 non-missing items used as the imputed value for that missing value.

5.12.   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.

5.13.   Cognitive Drug Research (CDR) Computerized Assessment System

The CDR computerized assessment system is a computerized battery of cognitive tests (Simpson 1991; Keith 1998). The CDR computerized assessment includes tests of attention and information processing speed (Simple Reaction Time, Choice Reaction Time, and Digit

Vigilance), verbal and visuospatial working memory (Numeric and Spatial Working Memory) and verbal and visual episodic memory (Immediate and Delayed Word Recall, Word Recognition and Picture Recognition).

The CDR computerized assessment will be administered at the study sites by trained administrators. During the Screening visit (Visit 1) and Visit 2, the CDR computerized assessment will be administered as training sessions to overcome test anxiety, ensure the patient is able and willing to undergo testing, and overcome initial training effects in order that a stable score is subsequently obtained at Baseline (Visit 3). All tasks are computer-controlled. The information is presented on the screen of a notebook computer and the responses recorded via a response module containing two buttons, one marked 'NO' and the other 'YES'. The test administrator sits with the patient throughout testing, explains each test using pre-set instructions, initiates each test and ensures the patient performs it appropriately.

When possible, the CDR computerized assessment should always be the first assessment conducted, and should be administered, within ±1 hour of the time of Baseline administration at all subsequent visits to control for potential impact of circadian variation in performance.

The tests are administered in the following order:

- Immediate Word Recall: A list of 12 words is presented on the screen at the rate of 1 every 2 seconds for the patient to remember. The patient is then given 1 minute to recall as many of the words as possible.

- Picture Presentation: A series of 14 pictures is presented on the screen at the rate of 1 every 3 seconds for the patient to remember.

- Simple Reaction Time: The patient is instructed to press the 'YES' response button as quickly as possible every time the word 'YES' is presented on the screen. Thirty stimuli are presented with a varying inter-stimulus interval.

- Digit Vigilance: A target digit is randomly selected and constantly displayed to the right of the screen. A series of digits is then presented in the center of the screen at the rate of 80 per minute and the patient is required to press the 'YES' button as quickly as possible every time the digit in the series matches the target digit. There are 15 targets in the series. The task lasts for 3 minutes.

- Choice Reaction Time: Either the word 'NO' or the word 'YES' is presented on the screen and the patient is instructed to press the corresponding button as quickly as possible. There are 30 trials for which each stimulus word is chosen randomly with equal probability and there is a varying inter-stimulus interval.

- Spatial Working Memory: A picture of a house is presented on the screen with 4 of its 9 windows lit. The patient has to memorize the position of the lit windows. For each of the 36 subsequent presentations of the house, the patient is required to decide whether or not the 1 window that was lit was also lit in the original presentation. The patient responds by pressing the 'YES' or 'NO' buttons as appropriate, as quickly as possible.

- Numeric Working Memory: A series of 3 digits is presented for the patient to hold in memory. This is followed by a series of 18 probe digits for each of which the patient

has to decide whether or not it is in the original series and press the 'YES' or 'NO' response button as appropriate, as quickly as possible.

- Delayed Word Recall: The patient is again given 1 minute to recall as many of the words as possible.

- Word Recognition: The original words plus 12 distracter words are presented one at a time in a randomized order. For each word the patient is required to indicate whether or not the patient recognizes it as being from the original list of words by pressing the 'YES' or 'NO' button as appropriate, as quickly as possible.

- Picture Recognition: The original pictures plus 14 distracter pictures are presented one at a time in a randomized order. For each picture the patient has to indicate whether or not the patient recognizes it as being from the original series by pressing the 'YES' or 'NO' button as appropriate, as quickly as possible.

Seven domains will be calculated from the CDR computerized assessment system tests (Power of Attention, Continuity of Attention, Cognitive Reaction Time, Response Variability, Quality of Working Memory, Quality of Episodic Secondary Memory, and Speed of Memory). Table 12 and Table 13 provide details on how these 7 domains are calculated.

Table 12:   CDR Computerized Assessments System Variables (Study RVT-101-2001)

Task Name	Variable	Abbreviation	Unit
Immediate Word Recall	Words Recalled Correctly	IRCL	#
	Errors	IRCLERR	#
Simple Reaction Time	Mean Reaction Time	SRT	ms
	Median Reaction Time	SRTM	ms
	Standard Deviation	SRTSD	ms
Digit Vigilance	Mean Reaction Time	VIGRT	ms
	Accuracy	VIGACC	%
	False Alarms	VIGFA	#
	Standard Deviation	VIGSD	ms
Choice Reaction Time	Mean Reaction Time	CRT	ms
	Median Reaction Time	CRTM	ms
	Accuracy	CRTACC	%
	Standard Deviation	CRTSD	ms
Spatial Working Memory	Median Reaction Time	SPMRTM	ms
	Sensitivity Index	SPMSI	SI
Numeric Working Memory	Median Reaction Time	NWMRTM	ms
	Sensitivity Index	NWMSI	SI
Delayed Word Recall	Words Recalled Correctly	DRCL	#
	Errors	DRCLERR	#
Word Recognition	Median Reaction Time	DRECRTM	ms
	Original Stimuli Accuracy	DRECOACC	%
	New Stimuli Accuracy	DRECNACC	%
Picture Recognition	Median Reaction Time	DPICRTM	ms
	Original Stimuli Accuracy	DPICOACC	SI
	New Stimuli Accuracy	DPICNACC	

Table 13: **Calculation of CDR Computerized Assessment System Domain Scores (Study RVT-101-2001)**

Name	Abbreviation	Unit	Calculation	Interpretation
Power of Attention	POW-ATT	ms	SRTM+VIGRT+CRTM	Lower scores are better
Continuity of Attention	CONT-ATT	#	(VIGACC*0.45)+(CRTACC*0.5)-VIGFA	Higher scores are better
Cognitive Reaction Time	COGRT	ms	CRTM-SRTM	Lower scores are better
Response Variability	POWATTCV	%	(SRTSD/SRT*100)+(CRTSD/CRT*100)+(VIGSD/VIGRT*100)	Lower scores are better
Quality of Working Memory	QL-WORK	SI	SPMSI+NWMSI	Higher scores are better
Quality of Episodic Secondary Memory	QL-EPIS	#	(DRECOACC+DRECNACC-100)+(DPICOACC+DPICNACC-100)+((IRCL-IRCLERR)*100/12)+ ((DRCL-DRCLERR*100/12)	Higher scores are better
Speed of Memory	SPEEDMEM	ms	SPMRTM+NWMRTM+DRECRTM+DPICRTM	Lower scores are better

Based on the derived domain scores, a standardized z-score will be derived for each CDR system domain using the formula below:

$$\text{Domain z-score} = (X2 - X1)/SD1$$

Where:

X2 is the domain score for each individual at any study assessment time post Baseline
X1 is the Baseline domain score for that individual
SD1 is the domain standard deviation of the population at Baseline

A composite z-score for the 7 CDR domains will be calculated based on the individual domain scores as:

(Continuity of Attention z-score + Quality of Episodic Secondary Memory z-score - Response Variability z-score + Quality of Working Memory z-score - Power of Attention z-score - Cognitive Reaction Time - Speed of Memory z-score)/7

Missing data will be handled as follows:

- If any individual item (variable) is missing for a Domain Score, the Domain Score will be set to missing. Thus, no imputation of individual Domain Scores will be performed; all variables are needed in order to calculate each Domain Score.

- If a single Domain Score is missing for the derivation of the Composite score (ie, one of the 7 Domain Scores is missing), the average of the 6 non-missing Domain z-scores will be used to replace the missing Domain z-score.

- If >1 Domain Scores are missing, the Composite score will be set to missing.

6.   SAMPLE SIZE CONSIDERATIONS

The sample size is based on assumptions of treatment benefit for one primary endpoint, the UPDRS- III total score. A sample size of 70 subjects per treatment group will allow a treatment difference of 4 points between placebo and active treatment in the change from baseline in UPDRS- III score to be detected with 88% power and a 0.05 significance level assuming an underlying standard deviation (SD) of 7.5

There are two secondary endpoints:

- ADAS-Cog 11 total score. A sample size of 80 subjects per treatment group will allow a treatment difference of 3 points between placebo and active treatment in the change from baseline in ADAS-Cog-11 score to be detected with 88% power and a 0.05 significance level assuming an underlying standard deviation (SD) of 6. Under the assumptions of a 2.5, 2.0 and 1.5 point treatment effect, the power is 74%, 55% and 35%, respectively.

- CIBIC+. A sample size of 80 subjects per treatment group will allow a treatment difference of 0.5 points between placebo and active treatment in the observed values in the CIBIC+ to be detected with 91% power and a 0.05 significance level assuming an underlying standard deviation (SD) of 0.95. Under the assumptions of a 0.4 and 0.3 point treatment effect, the power is 75% and 51%, respectively.

Randomization will be stratified according to Baseline MMSE score in the groupings of 14-17 points, 18-21 points and 22-26 points and according to whether subjects are or are not taking a cholinesterase inhibitor as a concomitant medication.

7. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

The primary statistical framework will be to test the superiority of Intepirdine over placebo. The null hypothesis is as follows, with a significance level of ≤0.05 needed for the null hypothesis to be rejected.

- There is **NO** statistically significant difference between Intepirdine and placebo in the mean change from baseline to Week 24 in the UPDRS- III.

- There **IS** a statistically significant difference between Intepirdine and placebo in in the mean change from baseline to Week 24 in the UPDRS- III.

The individual Intepirdine arms will be tested, as follows:

- Intepirdine 70 mg vs placebo

- Intepirdine 35 mg vs placebo

All hypothesis tests will be 2-sided, performed at the 5% level of significance.

The primary endpoint needs to achieve a significance level of 0.05 within an Intepirdine dose to allow for testing of the secondary endpoints within that dose.

7.1. Interim Analyses

No formal efficacy interim analyses are planned. An external SMC will review the safety data on an ongoing basis. Internal reviews of interim blinded safety data (by the Sponsor medical monitor) will occur on a regular basis.

7.2. Final Analysis

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.

- The study database is un-blinded.

7.3. 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 (and in this order, left to right, in the tables):

- Placebo

- Intepirdine 35 mg

- Intepirdine 70 mg

- Total (all 3 treatment groups combined)

For the primary efficacy and secondary efficacy variables ONLY, a table will be presented that will compare the POOLED data (pooled across the 35 mg and 70 mg Intepirdine arms) vs placebo.

Data analyses will be performed by, or under the direct supervision of, Axovant Sciences. Listings will be sorted by subject, period and time, or as appropriate. Tabulations will be presented by treatment and time. Unless stated otherwise, descriptive summaries will include n, mean, standard deviation, median, minimum, and maximum for continuous variables, and n and percent for categorical variables.

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:

- $\%CVb = 100 * \sqrt{(e^{SD^2} - 1)}$, 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 external vendor data, will be provided in data listings, sorted (in general) by patient and date. For date fields, Study Day will be calculated and presented in these listings.

7.4.   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.

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

- Placebo Run-In Population: All patients who took at least 1 dose of IP in 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.

- Randomized Population: The Randomized Population will include all patients who are randomized.

- Safety Population: All subjects who were randomized and took at least one dose of double-blind investigational product.

- Intent-to-Treat (ITT) Population: All subjects randomized to treatment who have taken at least one dose of double-blind investigational product and who have a baseline and at least one post baseline primary efficacy assessment for the UPDRS-III _or_ for the ADAS-Cog 11. This will be the primary population used for the efficacy analysis of all variables EXCEPT the UPDRS Part III and Part 5.

- UPDRS Primary Analysis Population: This will be the primary population used for the UPDRS Part III efficacy analysis, and is comprised of all ITT patients EXCLUDING those with no change OR a worsening in UPDRS-III total score prior to a dose increase in anti-parkinsonian medication (including but not limited to L-dopa, dopamine agonists, and monoamine oxidase inhibitors) and where there is _subsequent_ improvement or stabilization of UPDRS-III total score.

  – Dose increase in anti-parkinsonian medication is defined as an increase in dose, dose frequency or commencement of anti-parkinsonian medication where the total daily dose of anti-parkinsonian medication at Week 24 is higher than at Baseline.

- UPDRS Sensitivity Analysis Population: All ITT patients EXCLUDING those who increase or start anti-parkinsonian medication treatment during the double-blind treatment period.

  – Dose increase in anti-parkinsonian medication is defined as an increase in dose or frequency or commencement of anti-parkinsonian medication where the total daily dose of anti-parkinsonian medication at Week 24 is higher than at Baseline.

- Completers (CS) Population: ITT subjects who have completed the Week 24 visit with non-missing endpoints for all three variables, the UPDRS-III, ADAS-Cog 11, and CIBIC+. This population will be used for supportive analysis of all efficacy variables except the UPDRS variables.

- UPDRS Completers Primary Analysis Population: UPDRS Primary Analysis Population Patients who have a Week 24 UPDRS Part III value. This will be run only on the UPDRS Part III and Part 5 variables.

- Pharmacokinetics (PK) Population: All subjects in the Safety Population who undergo plasma PK sampling and have at least one post-baseline evaluable PK concentration result.

- Per-Protocol (PP) Population: Subjects in the ITT Population who have no major protocol violations. The PP Population will not be analyzed if this population comprises more than 95% or less than 50% of the ITT Population. This population will be used for confirmatory analysis all efficacy variables except the UPDRS variables. The PP Population will be identified prior to breaking the study blind. Major protocol violations will be defined as follows:

  – Subject had >7 consecutive days interruption in double-blind study medication

- Treatment with excluded concomitant medications during the double-blind treatment period.

- If taking an acetylcholinesterase inhibitor, subject had a change in AChEI dose or frequency during the double-blind treatment period.

- Significant study drug administration errors.

- Invalid baseline primary or secondary efficacy assessment.

- Subject does not have MMSE score 14-26 inclusive at Screening and Baseline

- Subject's baseline MMSE declined by 4 or more points from Visit 2 (Run-in) to Baseline

- Evidence of other causes of dementia other than DLB as the primary cause

- History or presence of significant medical or psychiatric condition at Screening.

- Subject does not live with (or have substantial Periods of contact with) a regular caregiver [at least 10 hours total and at least 3 days per week]).

- Subject is <80 % compliant or > 120% compliant between Visit 11 (Week 18) and Visit 12 (Week 24) with study drug. 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.

- UPDRS Per-Protocol (PP) Population: Subjects in the UPDRS Sensitivity Population who have no major protocol violations. The PP Population will not be analyzed if this population comprises more than 95% or less than 50% of the ITT Population. This population will be used for confirmatory analysis of the primary efficacy variable, UPDRS Part III (and Part 5). The PP Population will be identified prior to breaking the study blind. Major protocol violations will be defined as follows:

  - Subject had >7 consecutive days interruption in double-blind study medication

  - Treatment with excluded concomitant medications during the double-blind treatment period.

  - If taking an acetylcholinesterase inhibitor, subject had a change in AChEI dose or frequency during the double-blind treatment period.

  - Significant study drug administration errors.

  - Invalid baseline primary or secondary efficacy assessment.

  - Subject does not have MMSE score 14-26 inclusive at Screening and Baseline

  - Subject's baseline MMSE declined by 4 or more points from Visit 2 (Run-in) to Baseline

  - Evidence of other causes of dementia other than DLB as the primary cause

  - History or presence of significant medical or psychiatric condition at Screening.

- – Subject does not live with (or have substantial Periods of contact with) a regular caregiver [at least 10 hours total and at least 3 days per week]).

- – Subject is <80 % compliant or > 120% compliant between Visit 11 (Week 18) and Visit 12 (Week 24) with study drug. 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.

7.5. Multiplicity Considerations

This is a Phase 2b study in which there are two dose levels being compared to placebo, and there are two efficacy endpoints being assessed.

- All hypothesis tests will be 2-sided, performed at the 5% level of significance.

- The primary endpoint needs to achieve a significance level of 0.05 within a Intepirdine dose to proceed to testing the secondary endpoints and therefore maintain an overall 5% significance level for that dose. The two secondary endpoints are tested simultaneously at the 5% significance level.

The primary comparisons of interest are the individual Intepirdine arms vs placebo pairwise tests. The doses will be assessed in a hierarchy, with the 70 mg dose (vs placebo) assessed first. If significant on the primary endpoint (p-value ≤ 0.05), then the 35 mg dose (vs placebo) will be assessed. Otherwise, the 35 mg dose will not be assessed.

7.6. 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 there are multiple assessments collected on the same date, the average of these assessments will be used.

7.7. Derived and Transformed Data

The following definitions will be used.

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.

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

7.7.1.   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:

- Start and stop dates of adverse events

- Start and stop dates of concomitant medication

- Start and stop dates of double-blinded medication

Imputation rules for start and stop dates for the above 3 categories are described in Section 7.7.3.

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

Note that for missing DLB diagnosis date, if the month is missing, January will be imputed. If the day is missing, the 1st of the month will be imputed. Thus, if a patient has only a year of diagnosis, then the imputed diagnosis date will be January 1 of that year.

7.7.2.   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 double-blind 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 double-blind 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**. For efficacy and health outcomes data, assessments are regarded as Follow-Up if the assessment occurred >7 days after the last dose date of randomized investigational product.

Data collected at assessment visits provide information of the status at that point in time (eg, 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.

Data will be phased first. All data that slots into the screening 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.

Post-baseline assessments from the Double-Blind Treatment Period data will be visit slotted according to the time intervals shown (Table 14) 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 assigned to analysis windows based on Table 14 . This table defines the analysis windows associated with the screening, double-blind treatment period and follow-up periods of the study for assessments collected at each visit.

Table 14: **Windows of Relative Time Used to Classify Subject Assessments (Study RVT-101-2001)**

			Actual Study Day[1]	
Period	Visit Number	Nominal Day	Lowest	Highest
Screening (Week -6)[2]	1	-29		-1
Run-in (Week -2)[3]	2	-14	-39	-1
Baseline (Week 0)	3	1	1	1
On Treatment				
Week 1	4	7	2	10
Week 2	5	14	11	20
Week 4	6	28	21	35
Week 6	7	42	36	49
Week 8	8	56	50	63
Week 10	9	70	64	77
Week 12	10	84	78	105
Week 18	11	126	106	147
Week 24	12	168	148	last dose+7
Follow-up (Week 26)/ET	13	182	+7	

[1] Actual study day is derived by (visit date – run-in date) + 1 for Screening visit; (visit date – first dose date) + 1 for other visits except for Screening visit.
[2] Visit window for Screening visit is relative to run-in visit instead of Baseline visit.
[3] Visit windows may be adjusted to account for 2 10-day extensions to the run-in period allowed per protocol.

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 analysis by visit (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 (or is missing), the other assessment would be used. For example, for ADAS-Cog, if the total score could not be calculated for the closest assessment, another assessment would be used. For "anytime" analysis of the safety data, all data points will be used.

- If on one of the dates only one blood pressure measurement is available (ie, only one of either the supine, sitting or standing measurements), then the assessments recorded on the other date, which slots to the same visit interval, will be used in tabulations and analyses. A similar rule applies for pulse rate assessments.

- Orthostatic blood pressure and heart rate will be calculated using vital sign measurements from whatever 2 body positions they are collected. If vital signs are collected in all 3 body positions (supine, sitting and standing), the supine and standing measurements will be used to calculate orthostatic blood pressure and heart rate.

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 by visit. For the "anytime" analysis, all data points will be used.

7.7.3.   Imputation Rules for Missing Start and Stop Dates of Prior and Concomitant Medications, Adverse Events and Birth Dates

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.

7.8.   Subgroups

While this study is not powered for subgroup analyses, the primary and secondary efficacy endpoints will be analyzed by defined subgroups. Additional subgroups may be added as the data dictate, and subgroup analysis may be performed on other variables. 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.

Assessment of safety by subgroups will be performed for OTAEs.

- UPDRS based on baseline score
    - UPDRS <= 10 and >10
    - UPDRS < Baseline Overall Median and >= Baseline Overall Median (using overall SAFETY Population)
- Cognitive Severity based on baseline MMSE
    - 14-19 and 20-26
- Geographic Region
    - US
    - EU and Canada (English speaking)
    - All other Europe (non-English speaking)
- Age (years)
    - < Median and ≥Median (using overall SAFETY Population)
    - <65 and ≥65
- Gender
    - Female and Male
- Race
    - White and Non-White. Not that a few patients may be excluded from this analysis as their country privacy laws prohibit collection of race.
- Ethnicity

- – Hispanic and Not Hispanic

- Use of Acetylcholinesterase medications

    - – Using AChEIs vs Not Using AChEIs

    - – Using Donepezil

    - – Using Rivastigmine

- Use of Anti-parkinsonian medications (including but not limited to L-dopa, dopamine agonists, and monoamine oxidase inhibitors)

    - – Using

    - – Not using

## 8.	STUDY POPULATION

## 8.1.	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 and screen-failed prior to the Placebo Run-in period. The denominator for those outcomes for this analysis will be the Screened Population.

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 Placebo Run-in Population 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. The following will also be presented for the Randomized Population:

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

- An overall summary of the number and percentage of subjects who completed or discontinued from the Treatment Period will be displayed by treatment group. Reasons for discontinuation during the Treatment Period will be presented in the order they are displayed in the CRF. Subjects who discontinued from the Treatment Period will be listed by treatment group and subject.

- A cumulative distribution plot of the time to discontinuation during the Treatment Period will be provided and will include all treatment groups to allow for a visual inspection of the time to discontinuation. A similar plot will be provided for those patients who discontinue due to an on-treatment adverse event. The denominator for this plot will be the underlying population (and thus 'censoring' will not be applied).

## 8.2.	Screen Failures

Screen failure data will be tabulated by reason patients were found to not be eligible for the study. Patients may have screen failed due to any of the inclusion/exclusion criteria reasons as well as:

- Loss to follow-up

- Deviation/noncompliance with the protocol

- Adverse event

- Sponsor termination of the patient from the study

- Investigator decision

- Caregiver consent withdrawn

- Subject consent withdrawn

A listing of reasons of screen failures will be provided.

8.3.   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 stratum per the randomization, AChEI use (yes/no)} will be tabulated, along with the defined subgroups (see Section 7.8).

- Dementia with Lewy Bodies diagnosis information will be summarized in a similar fashion.

- Other characteristics collected on the CRF will be summarized by the number and percentage of subjects, or as appropriate.

- Baseline values for the efficacy endpoints will be provided as part of the baseline characteristics assessment.

- Details of the primary caregiver (relationship to subject, cohabitating with subject, # hours spent with subject, # days spent with subject) will be included in a tabulation.

- A CT scan or MRI may be performed at screening; outcomes will be listed.

8.3.1.   Lewy Body Composite Risk Score (LBCRS)

The LBCRS is a scale that has been shown to help discriminate DLB from other causes of dementia, such as Alzheimer's disease, using a cut-off score of 3 (Galvin 2015). LBCRS will be administered during screening to collect information regarding the clinical presentation of DLB symptoms.

Ten questions are completed by a clinician after interview with patient and caregiver and a complete physical and neurological exam. The clinician rates the presence or absence of physical signs and can elicit information from either the patient or caregiver regarding the presence or absence of symptoms. 1 point is awarded for each "Yes" response.

The LBCRS total score is the sum of the items 1 through 10, with scores ranging from 0 to 10. Interpretation as to whether the patient is a Lewy Body 'case' is as follows:

- 0 to 2: Non-Lewy Body Case
- >2: Probably Lewy Body Case

The number (%) of patients meeting the threshold for >2 on the LBCRS will be tabulated, as well as the distribution of scores (ranging from 0 to 10). These data will be presented in the baseline characteristics table and will also be listed.

8.4.   Subject Inclusion and Exclusion Criteria

Patients not meeting eligibility criteria will be listed.

8.5.   Medical History

Medical history 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.

8.6.   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
  - Post-treatment period
- New-onset concomitant medications (those starting only after the date of first dose of double-blind study medication)

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.

A concomitant medication can be counted in more than one period.

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

A listing of non-medication therapy will be provided.

8.7.   Extent of Exposure

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

- Duration of exposure in days = (Double-Blind Treatment stop date – Double-Blind Treatment 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

Exposure will also be categorized as <23 weeks and ≥23 weeks.

8.8.   Treatment Compliance

Compliance with study drug will be assessed over the Double-Blind Treatment Period and Run-In Period using drug dispensing records. Treatment compliance will be computed by determining the number of tablets taken relative to the number of tablets expected. Patients are expected to take two tablets per day, one from each of two bottles.

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.

8.9.   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. Any dose of intepirdine greater than 105 mg within a 24-hour time period will be considered an overdose. Cases of overdose were to be documented in the CRF. Overdose in the absence of other AEs were not be reported as an AE in its own right.

9.   EFFICACY

9.1.   General Considerations

Tabulation of the primary and secondary endpoints will generally be presented for the ITT, Per-Protocol, and Completers Population. The primary and secondary endpoints will be tabulated for the subgroups, while other variables may be added for subgroup analysis. Efficacy assessments will be performed on the observed cases dataset unless otherwise specified. The UPDRS Primary Analysis Population will form the basis for the primary conclusions regarding the primary efficacy variable (for the UPDRS Part III), while the ITT Population will form the basis for conclusions regarding all other endpoints. The UPDRS Part 5 will also be assessed using the UPDRS Primary Analysis Population.

Baseline or pre-dose assessment is the last available non-missing assessment prior to time of first dose of double-blind treatment, unless otherwise specified. If there are multiple efficacy assessments for a given parameter collected on the same date, the later of these assessments will be used.

9.2.   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 UPDRS-III score in patients with DLB disease.

- H_a: There __IS__ a difference between treatment with intepirdine and treatment with placebo with respect to changes over 24 weeks in the UPDRS-III score in patients with DLB disease.

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

9.3.   Analysis of the Primary Efficacy Endpoint

The primary efficacy endpoint is the change from Baseline in the UPDRS-III total score at Week 24. The primary method of analysis will be using a mixed model for repeated measures (MMRM), with supportive method being the Wilcoxon rank statistic (a nonparametric test).

For the primary efficacy and secondary efficacy variables ONLY, in addition to the primary pairwise analyses, a table will be presented that will compare the POOLED data (pooled across the 35 mg and 70 mg Intepirdine arms) vs placebo. This will be done ONLY for the MMRM analysis.

9.3.1.   MMRM

Primary treatment comparisons between Intepirdine and placebo will be performed on the change from Baseline to Week 24 using an MMRM with restricted maximum likelihood estimation, an unstructured covariance matrix, and the Kenward-Roger approximation for denominator degrees of freedom. The model will include terms for treatment, visit, treatment by

visit interaction, pooled geographic region, baseline MMSE score, the baseline value, and the baseline value by visit interaction.

- 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.

  – For 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 primary efficacy endpoint, as appropriate.

As additional supportive information, treatment differences at *each post baseline visit* will also be derived using the MMRM model; however, primary inferences will be drawn from treatment differences derived from the MMRM models at Week 24.

The estimated treatment difference ("Intepirdine 35mg – Placebo" and "Intepirdine 70mg – Placebo") at each visit will be displayed in the summary of statistical analysis together with the 95% confidence interval and the associated p-value.

No imputation of the missing values will be performed for the primary analyses, ie, the data used in the analysis will be the actual observed responses at each visit.

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.

9.3.1.1.   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.

  – If convergence issues arise, the base by visit interaction may be dropped.

Sample SAS code is as follows. The actual SAS code may be modified as appropriate. Note that the treat*region interaction will be assessed for significance (at alpha 0.10), and if not significant, will be dropped from the model.

```
ods trace on;
```

```
Proc mixed data=MMRM1 method=reml PLOTS(MAXPOINTS=200000000000);
  class usubjid trt avisitn region;
  model chg=trt avisitn trt*avisitn trt*region base basemmsescore
base*avisitn region /ddfm=kr;
  repeated avisitn/subject=usubjid type=un r;
  lsmeans trt*avisitn/pdiff cl ;
  lsmeans trt*region/pdiff alpha =0.10;
  ods output lsmeans=_LS diffs=_lsdiff  tests3 =tst;
run;
ods trace off;
```

9.3.2.   Description of Missing Data for the Primary Efficacy Endpoint

The patterns of missing data will be presented, with the number (%) of patients in each treatment group with each pattern of missingness, as demonstrated in Table 15. These data will be examined 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. Note that for the primary efficacy endpoint, data were only collected at baseline, Week 4, and Week 24, and thus the patterns would be presented accordingly.

Table 15:   Examples of Patterns of Missingness (Study RVT-101-2001)

Pattern	Treatment Group	Number (%) of Patients	Baseline	Week 1	Week 4	Week 12	Week 18	Week 24
1	Intepirdine 35mg	xx (xx.x%)	X	X	X	X	X	.
	Intepirdine 70mg	xx (xx.x%)	X	X	X	X	X	.
	Placebo	xx (xx.x%)	X	X	X	X	X	.
2	Intepirdine 35 mg	xx (xx.x%)	X	X	X	X	.	.
	Intepirdine 70mg	xx (xx.x%)	X	X	X	X	.	.
	Placebo	xx (xx.x%)	X	X	X	X	.	.
Etc.								

9.3.3.   Wilcoxon Rank Test

Analysis using a non-parametric (distribution-free) Wilcoxon Rank Sum Test will be performed. This analysis will be performed for each protocol-scheduled visit using the observed cases change from baseline data. *Sample* SAS code is as follows. The actual SAS code may be modified as appropriate.

```
Proc npar1way data=dat;
  By visit;
  Class treat;
  Var chg_from_baseline;
  ods select WilcoxonScores
```

9.3.4.   Sensitivity Analysis from Imputation of Missing Efficacy Data for Primary Efficacy Endpoint

9.3.4.1.   Cumulative Distribution Function

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.

9.3.4.2.   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. If there are no on-treatment values for a patient, the baseline (last pre-treatment) observation will be carried forward.

Sample SAS code is as follows (where basescore is the baseline score for the endpoint and center is the pooled geographic region variable.

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

9.3.4.3.   Multiple Imputation

Missing data for the primary endpoint 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 nature of missing data will be explored and the extent of missing data pattern (over time) will be summarized.

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 score, geographic region, and previous score (ie, the non-missing score from previous visit). The first visit to be imputed (as necessary) would be the Week 4 time point, then the Week 24 time point. Thus, the process is repeated sequentially, by visit, until the monotone missing data pattern is completely filled in (imputed). In the case of the secondary endpoint, in which values are collected more frequently, there will be more '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 4) and second visit (Week 24) for UPDRS- III 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 4);
  proc mi data=DATAIN_MONO out=DATAREG1 seed=465 nimpute=1;
      by _Imputation_;
      var visitnum1 basescore MMSEbase region previous;
      monotone regression(visitnum1);
      run;
  ** Second procedure imputes missing values for the second visit (Week 24)
  proc mi data=DATAREG1 out=DATAREG2 seed=465 nimpute=1;
      by _Imputation_;
      var visitnum1 visitnum2 basescore MMSEbase region 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 score, region, 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.

9.4.   Analysis of the Secondary, Tertiary, and Exploratory 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 as well as a Wilcoxon rank (nonparametric) test. For those endpoints that are collected only once on-treatment, 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.

When required, the ANCOVA will have effects for treatment, baseline variable score, baseline MMSE score, and geographic region. The observed case analysis will be considered the primary assessment of this variable.

No methods for imputation of missing values are planned for the exploratory endpoints. Exploratory endpoints will only be assessed using the MMRM analysis method.

9.4.1.   Secondary Efficacy Endpoint Analyses

The secondary efficacy analyses endpoints are the change from baseline in the ADAS-Cog 11 score at Week 24, and the observed scores of the CIBIC+ at Week 24.

This will be analyzed using similar MMRM methods as the primary endpoint (however, for the CIBIC+, no baseline value or interaction with visit will be included). A Wilcoxon Rank test will also be performed, as will the sensitivity analysis for LOCF to Week 24 and multiple imputation. The sensitivity analyses as being performed on the primary endpoint (CDF, LOCF, and multiple imputation) will be performed for the secondary endpoints.

For the CIBIC+, The number and percentage of subjects in each category of CIBIC+ will also be summarized by visit for each treatment group and tested using a CMH RMS test.

9.4.2.   Tertiary Endpoints

Tertiary efficacy endpoints include the following. Note that analysis will be performed for each protocol-scheduled time point for each analysis variable and method, but the focus will be on the Week 24 time point.

- Change from Baseline on UPDRS-5 at Week 24

    – MMRM on observed data (also on UPDRS subgroups)

    – Wilcoxon Rank test on observed data

- Change from Baseline on ADAS-Cog 13 at Week 24

    – MMRM on observed data

    – Wilcoxon Rank test on observed data

- Change from Baseline in the Total score, and the Basic, Instrumental, Independence subscores of the ADCS-ADL scale at Week 24

    – MMRM on observed data (also on Concomitant Parkinsonian Meds subgroups)

    – Wilcoxon Rank test on observed data

- Change from Baseline in the Total Composite z-score and in the actual score (NOT the z-scores) of each individual domain of the CDR computerized assessment system at Week 24.   Thus, in addition to the Total Composite, this includes the following 7 domains:

Power of Attention
Continuity of Attention
Cognitive Reaction Time
Response Variability
Quality of Working Memory
Quality of Episodic Secondary Memory
Speed of Memory

    – MMRM on observed data

    – Wilcoxon Rank test on observed data

- Change from Baseline in COWAT total score at Week 24

- – MMRM on observed data

- – Wilcoxon Rank test on observed data

- Change from Baseline in a composite z-score of the 7 CDR System domains plus the COWAT total score at Week 24

  - – MMRM on observed data

  - – Wilcoxon Rank test on observed data

- Change from Baseline in the sum of Parts A and B (Delusions and Hallucinations), and the sum of Parts of D and E (Depression and Anxiety) of the NPI at Week 24

  - – MMRM on observed data

  - – Wilcoxon Rank test on observed data

- Change from Baseline on the total severity score and distress score of the NEVHI score at Week 24 for subjects with visual hallucinations

  - – MMRM on observed data

  - – Wilcoxon Rank test on observed data

- Number (%) of patients with shifts from baseline in the NEVHI Total Severity Score and in the Distress Score (ie, 2 separate analyses) at Week 24, using a CMH row means score (CHM RMS) test. The categories for shifts are ordered as:

  - – worsening from baseline

  - – remaining unchanged from baseline

  - – improving from baseline.

  Sample SAS code for the CMH RMS is as follows (where MMSEbase is the baseline randomization stratum level for the patient, NOT the actual MMSE value):

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

- Change from Baseline in CAF severity score for each question (there are 2 questions) at Week 24. Note that patients with no cognition or confusion scores are assigned a 0 and thus are included in this analysis.

  - – MMRM on observed data

  - – Wilcoxon Rank test on observed data

- The number (%) of patients with CAF cognition or confusion (ie, a yes on one or the other question) at Week 24. This will be generated separately for each of the 2 questions as well as 'overall', ie, having a YES on either question.

  - – CMH test for general association (similar SAS code as for the CMH RMS but the RMS option is not needed).

- Change from Baseline on DS Total Score at Week 24

  − MMRM on observed data

  − Wilcoxon Rank test on observed data

9.4.3.   Analysis of Exploratory Efficacy Endpoints

Exploratory endpoints include:

- For each of the five EQ-5D dimensions (Mobility, self-care, usual activities, pain/discomfort, anxiety/depression):

  − CMH RMS at Week 24

  − Using the Dichotomized values (No Problems vs Problems), a CMH general association analysis will be run at Week 24.

- Change from Baseline on EQ-5D Visual analogue score at Week 24

  − MMRM on observed data

- Change from Baseline on CSI score at Week 24

  − MMRM on observed data

- Change from Baseline on Likert scale (overall quality of sleep) at Week 24

  − MMRM on observed data

- Change from Baseline in sum of Parts D and E of the NPI at Week 24

  − MMRM on observed data

10.   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 changes from pre-treatment to 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 before dosing with the double-blind study medication will be used as the baseline value for that parameter. If a dose is taken on the same day as the assessments, the assessment will be assumed to be baseline.

Adverse events will be classified using MedDRA Version 18.1 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.

10.1.   Adverse Events (AEs)

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 on or after the first dose of single-blind study medication but 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.

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.

- Double-Blind Treatment Period (using the Safety Population as the denominator)

  - OTAEs. OTAEs will also be presented by the defined subgroups.

  - OTAEs occurring in $\geq$ 5% and 10% in EITHER Intepirdine treatment arm. These thresholds may be adjusted as the data dictate.

  - 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

  - An analysis of Adverse Reactions (defined as events occurring with higher incidence in either Intepirdine arm than in the Placebo arm) will be provided via a tabulation of these OTAEs.

- Post-Treatment Period (using the Safety Population as the denominator)

  - AEs

  - SAEs (Fatal and Non-Fatal)

A listing off deaths will be presented.

The time to first onset of OTAEs will be presented using the following bins of onset:

- 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 may be presented via Kaplan-Meier plots and a log-rank statistic will be calculated (stratified by randomization baseline MMSE score). These events may include:

- OTAE resulting in discontinuation from study medication

- SAE

- Death

- Falls

  - SOC=Injury, poisoning, and procedural complications, PT=Fall.

  - The proportion of patients falling at any time will also be compared between the treatment groups using a CMH controlling for baseline stratification of MMSE and AChEI use.

10.2.   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 by visits and at ANY TIME which includes 7 days after the date of the last dose of double-blind study medication. 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 16 and for hematology in Table 17.

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.

In addition, the incidence of OTAEs relating to haematology and to 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 16:   Serum Chemistry Potentially Clinically Significant Criteria (Study RVT-101-2001)

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 17:   Haematology and Differentials Potentially Clinically Significant Criteria (Study RVT-101-2001)

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
Absolute Lymphocyte Count (10^9/L)	Male/Female	0.7	4.1

10.3.   Vital Signs

Vital sign data (blood pressure [BP], pulse rate, 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 by vist and at ANY TIME. Orthostatic BP measurements will also be tabulated using the PCS criteria.

PCS ranges for these parameters are provided in Table 18.

Table 18:   List of Potentially Clinically Significant Ranges for Vital Sign Parameters (Study RVT-101-2001)

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
Orthostatic Hypotension
≥ 30 mmHg decrease (reduction) in SBP upon positional change
≥ 20 mmHg decrease (reduction) in DBP upon positional change

10.4.   Physical and Neurological Examinations

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

10.5.   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 and 7 days after the last dose (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 QTcF values (Table 19).

Table 19:   List of Potentially Clinically Significant Ranges for ECG Parameters (Study RVT-101-2001)

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

10.6.   Pregnancy

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

10.7.   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 outcomes from the past year), while baseline 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 20 presents the general manner in which these data will be tabulated.

Table 20: **Sample Table Demonstrating How C-SSRS Outcomes will be Presented (Study RVT-101-2001)**

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.

10.8. Questionnaire for Symptoms of Potential Orthostasis (QSO)

Data recorded from the questionnaire for symptoms of potential orthostasis will collect the occurrence of the following symptoms:

- Dizziness (yes/no)
- Fainting (yes/no)
- Fall (yes/no).

For each symptom, information on the following will be collected:

- Severity (mild/moderate/severe)
- Number of times (since last minute)
- Comparison to prior episodes (worse/same/better)

The number of occurrences of each symptom over time will be presented. Severity, for those with each symptom, will be tabulated by visit.

These will be summarized and listed by treatment and visit, as well as 'at any time' during double-blind treatment.

10.8.1.   Falls as Reported from QSO

The rate of falls per month will be analyzed using the data on fall occurrence that is collected at each visit in the Treatment Period via the Questionnaire for Signs of Potential Orthostasis.

Rate of falls per month = (total number of falls / duration (days)) x 30, where:

- Total number of falls = summation of falls over all visits in the Treatment Period (through Week 24)

- Duration of treatment (days) through Week 24 = (last dose date of double blind IP– first dose date of double blind IP) + 1

Incidence rate will be summarized by treatment. The incidence rate will be analyzed using the MMSE model methodology as described for the primary efficacy endpoint but without the 'visit' terms.

Additional analysis will be performed for each symptom by treatment, visit and overall

11. PHARMACOKINETIC ANALYSES

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

Individual plasma intepirdine PK parameters (AUCτss, Cmax-ss and Cmin-ss) will be estimated from a prior population PK model developed from previously reported Phase 1, 2, and 3 studies. Model parameters will not be re-estimated with data from the current study, rather, individual PK parameters will be estimated using a post-hoc procedure. A report describing this procedure and listings and summaries of PK parameters will be included as an appendix to the CSR.

11.1. Pharmacokinetic/Pharmacodynamic analyses

The relationship between intepirdine exposure and the primary endpoints of this study will be assessed graphically (scatter plot, linear or non-linear regression/smoothing lines). Namely, the relationship between individual intepirdine AUCτss, Cmax-ss and Cmin-ss and the following:

- Change from baseline total score of the UPDRS Part III at Week 24
- Change from baseline total score of the ADAS Cog 11 at Week 24
- Total score of the CIBIC+ at Week 24

Additional linear/non-linear modeling of these endpoints may be performed based on these initial analyses. Additional modeling may be described in a separate SAP and report.

12.   REFERENCES

References provided with Protocol.