Statistical Analysis Plan Version 3 I7S-MC-HBEH

Effect of LY3154207 on Cognition in Mild-to-Moderate Dementia Due to Lewy Body Dementia (LBD) Associated with Idiopathic Parkinson's Disease (PD) or Dementia with Lewy Bodies (DLB)

NCT03305809

Approval Date: 08-Feb-2021

1. Statistical Analysis Plan: I7S-MC-HBEH: Effect of LY3154207 on Cognition in Mild-to-Moderate Dementia Due to Lewy Body Dementia (LBD) Associated with Idiopathic Parkinson's Disease (PD) or Dementia with Lewy Bodies (DLB)

LY3154207 Lewy Body Dementia

Study I7S-MC-HBEH is a multicenter, randomized, double-blind, parallel-group, placebo-controlled, fixed-dosage, Phase 2a study comparing 3 dosages of LY3154207 (10, 30, or 75 mg administered orally once a day) with placebo over 12 weeks in subjects with mild-to-moderate Parkinson's disease dementia.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I7S-MC-HBEH
Phase 2

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly: 28 November 2017

Statistical Analysis Plan Version 2 electronically signed and approved by Lilly: 17 July 2020

Statistical Analysis Plan Version 3 electronically signed and approved by Lilly on date provided below.

Approval Date: 08-Feb-2021 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

SAP Version 1 was approved prior to unblinding.

SAP Version 2 was approved prior to unblinding and includes updates from Amendment (b) and information on coronavirus disease 2019 (COVID-19) impact.

SAP Version 3 was approved after unblinding. The overall changes and rationale for the changes incorporated in Version 3 are as follows:

- Edits were made to comply with ALCOA+++ principles.
- Administrative changes (not affecting content), such as correcting typographical and formatting errors, were made.

4. Study Objectives

Table HBEH.4.1 shows the objectives and endpoints of the study.

Table HBEH.4.1. Objectives and Endpoints

Objectives	Endpoints
Primary To test the hypothesis that LY3154207 administered at 10 mg, 30 mg, and/or 75 mg daily (or 50 mg based on interim analysis) (QD) oral dosing for 12 weeks will result in a significant improvement in cognition in subjects with mild-to-moderate LBD compared with placebo.	Change in the CoA composite score of the CDR-CCB from baseline to Week 12
Secondary **Efficacy:**	
To evaluate the global efficacy of LY3154207	ADCS-CGIC score from baseline to Week 12
To evaluate the efficacy of LY3154207 on cognitive outcomes	Change in the CDR-CCB PoA composite score from baseline to Week 12
To evaluate the efficacy of LY3154207 on cognitive outcomes	Change in the ADAS-Cog_{13} score from baseline to Week 12
To evaluate the efficacy of LY3154207 on cognitive outcomes	Change in the MoCA score from screening to Week 12
To evaluate the efficacy of LY3154207 on neuropsychiatric symptoms	Change in the NPI total and individual items scores from baseline to Week 12
To evaluate the effect of LY3154207 on daytime sleepiness	Change in the ESS score from baseline to Week 12
To evaluate the effect of LY3154207 on Parkinson's disease severity	Change in the MDS-UPDRS total score (sum of parts I-III) from baseline to Week 12
To evaluate the efficacy of LY3154207 on functional outcome	Change in the PDAQ-15 total score from baseline to Week 12
To evaluate the effect of LY3154207 on verbal fluency	Change in D-KEFS Verbal Fluency test score from baseline to Week 12
To evaluate the effect of LY3154207 on motor symptoms and signs	Change in MDS-UPDRS Parts II (motor experiences of daily living) and III (motor exam) from baseline to Week 12

Objectives and Endpoints

Objectives	Endpoints
Secondary **Safety:**	
To evaluate the effect of LY3154207 on acute changes of vital signs on the first day of dosing	Number of subjects who met the potentially clinically significant vital signs criteria at 3 consecutive time points at Visit 3 (Day 1 stopping rules)
To evaluate the effect of LY3154207 on SBP and DBP on the first day of dosing	Change in in-clinic BP from 0 up to 8 hours post dose on the first day of study drug dosing
To evaluate the effect of LY3154207 on pulse rate on the first day of dosing	Change in in-clinic pulse rate from 0 up to 8 hours post dose on the first day of study drug dosing
To evaluate the effect of LY3154207 on SBP and DBP from baseline to Week 12	Change in in-clinic mean BP at baseline to mean SBP at Week 12
To evaluate the effect of LY3154207 on pulse rate from baseline to Week 12	Change in in-clinic mean pulse rate at baseline to mean pulse rate at Week 12
To evaluate the effect of LY3154207 using HBPM	Change in HBPM for SBP, DBP, and pulse rate from baseline to Week 12
To evaluate the effect of LY3154207 on physical withdrawal symptoms	Change in the PWC-20 total score from Week 12 to in-clinic follow-up visit
Pharmacokinetics: To assess the PK of LY3154207 in a population of subjects with mild-to-moderate dementia due to Lewy Body Dementia	Steady-state trough plasma concentrations of LY3154207 at Week 12

Objectives and Endpoints

Objectives	Endpoints
Tertiary/Exploratory **Efficacy:**	
To evaluate the efficacy of LY3154207 on cognitive outcomes	Change in total score of the ADAS-Cog_{11} from baseline to Week 12
To evaluate the efficacy of LY3154207 on cognitive outcomes	Change in the CDR-CCB Cognitive reaction time composite score, Reaction Time Variability composite score, and Attention Ratio and Speed Score composite scores from baseline to Week 12 Change in the CDR-CCB individual task outcomes from baseline to Week 12 including simple reaction time, digit vigilance, choice reaction time, numeric working memory, spatial working memory, verbal memory, and visual memory.
To evaluate the effect of LY3154207 on non-motor symptoms	Change in MDS-UPDRS • Part I Non-motor experiences (Total Score)
To evaluate the effect of LY3154207 on sleep	Change in MDS-UPDRS items 1.7 sleep problems and 1.8 daytime sleepiness
To evaluate the effect on motor complications	Change in MDS-UPDRS Part IV (motor complications) from baseline to Week 12
To evaluate the effect of LY3154207 on depressive symptoms	Change in NPI score (Part D Depression/Dysphoria) from baseline to Week 12 Change in MDS-UPDRS item 1.3 Depressed Mood from baseline to Week 12 Change in GDS total score from baseline to Week 12
To evaluate the effect of LY3154207 on apathy	Change in NPI score (Part G Apathy and Difference) from baseline to Week 12 Change in item 1.5 Apathy in MDS-UPDRS from baseline to Week 12
To evaluate the effect of LY3154207 on fatigue	Change in MDS-UPDRS item 1.13 fatigue from baseline to Week 12

Objectives and Endpoints

Objectives	Endpoints
Tertiary/Exploratory **Safety:**	
To evaluate the effect of LY3154207 on hallucinations and psychosis including dose response	Change in item 1.2 score of the MDS-UPDRS Part IA from baseline to Week 12 Change in score for the delusion (Part A) and hallucinations (Part B) domains of the NPI from baseline to Week 12
To evaluate the effect of LY3154207 on impulse control symptoms	Change in QUIP total score from baseline to Week 12
To evaluate the effect of LY3154207 on dyskinesias	Change in MDS-UPDRS Items 4.1 and 4.2 from baseline to Week 12
Digital Biomarkers:	
To evaluate the feasibility of using the Lilly Trial application (Lilly Trial app) to evaluate the effect of LY3154207 on motor tasks or cognitive outcomes in a subset of subjects	To explore trends, utilizations, and changes in quantitative motor and cognitive outcomes as measured by the Lilly Trial App scores from predose assessments to 12 weeks
To explore the feasibility of using actigraphy to evaluate the effect of LY3154207 on various sleep parameters and daytime activity including but not limited to sleep latency, duration of sleep, sleep–wake cycles, sleep architecture, and sleep disruption	Change in actigraphy measured sleep parameters as measured by an actigraphy watch from predose assessments to 12 weeks
Pharmacogenomic:	
To evaluate genetic interactions with study treatment response or safety if appropriate/as warranted	Explore relationships between pharmacogenomics variations and relevant clinical outcomes.

Abbreviations: ADAS-Cog_{11} = 11-item Alzheimer's Disease Assessment Scale – Cognitive subscale; ADAS-Cog_{13} = 13-item Alzheimer's Disease Assessment Scale – Cognitive subscale; ADCS-CGIC = Alzheimer's Disease Cooperative Study – Clinician Global Impression of Change; CDR-CCB = Cognitive Drug Research Computerized Cognition Battery; CoA = Continuity of Attention; DBP = diastolic blood pressure; D-KEFS = Delis-Kaplan Executive Function System; DLB = dementia with Lewy bodies; ESS = Epworth Sleepiness Scale; GDS = Geriatric Depression Scale; HBPM = home blood pressure monitoring; LBD = Lewy Body Dementia; MDS-UPDRS = Movement Disorder Society's Unified Parkinson's Disease Rating Scale; MoCA = Montreal Cognitive Assessment; NPI = Neuropsychiatric Inventory; PDD = Parkinson's disease dementia; PDAQ-15 = Penn Parkinson's Daily Activities Questionnaire-15; PK = pharmacokinetics; PoA = Power of Attention; PWC-20 = physician withdrawal checklist-20; QD = once a day; QUIP = Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease; SBP = systolic blood pressure.

5. Study Design

5.1. Summary of Study Design

Study I7S-MC-HBEH (HBEH) is a multicenter, randomized, double-blind, parallel-group, placebo-controlled, fixed-dosage, Phase 2a study comparing 3 dosages of LY3154207 (10, 30, or 75 mg administered orally once a day [QD]) with placebo over 12 weeks in subjects with mild-to-moderate dementia associated with Lewy body dementia (LBD) (Parkinson's disease dementia [PDD] and dementia with Lewy bodies [DLB]). The study includes a Screening Period (Visits 1 to 2) of a minimum of 7 days and up to 14 days, a 14-day Pretreatment Period (Visits 2 to 3), a 12-week Treatment Period (Visits 3 to 11), and a 14-day Safety Follow-Up Period (Visits 11 to 801) (Figure HBEH.5.1). Subjects who meet entry criteria will be randomized in a 1:1:1:1 ratio to LY3154207 (10 or 30 or 75 mg QD) or placebo. The primary objective of this study is to test the hypothesis that LY3154207 administration for 12 weeks will result in a significant improvement in cognition as measured by the change from baseline to Week 12 in the Continuity of Attention (CoA) composite score of the Cognitive Drug Research Computerized Cognition Battery (CDR-CCB), in subjects with mild-to-moderate LBD, compared to placebo. The CoA has demonstrated a significant treatment effect in previous trials in subjects with PDD and is resistant to learning and placebo effects (Wesnes et al. 2005; Rowan et al. 2007).

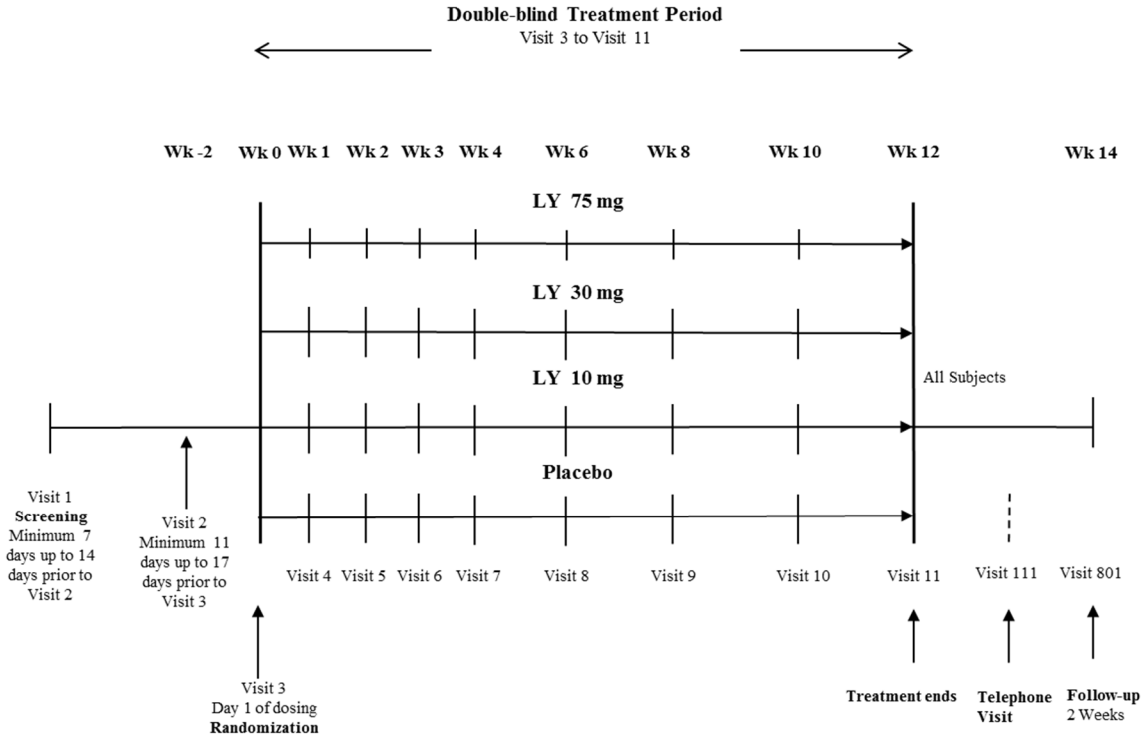

Figure HBEH.5.1. Illustration of study design for Clinical Protocol I7S-MC-HBEH.

5.2. Determination of Sample Size

Approximately 750 subjects will be screened to achieve 340 randomized subjects. The randomization will be 1:1:1:1 to LY3154207 (10, 30, or 75 mg QD) or placebo treatment arms. CCI [redacted] If the true effect size for LY3154207 relative to placebo on the CoA is 0.4 with the sample size of 340, there will be greater than 80% probability of passing this criterion. A 20% discontinuation rate by Week 12 has been accounted for in the sample size determination.

Sample size estimation was based on an article by Wesnes et al. (2005) where the difference in mean change from baseline for rivastigmine compared to placebo and the pooled standard deviation (SD) for CoA were estimated as 5.7 units and 18 units, respectively. The rivastigmine effect corresponded to a reduction in the deficit of subjects compared to age-matched healthy volunteers of approximately 45%. Based on this result, the effect size of 0.4 in this study corresponds to a mean difference of 7.2 units (approximately 57% reduction in the deficit).

This sample size will allow assessment of the in-clinic blood pressure (BP) and pulse rate (PR) values on the first day of dosing in order to evaluate any changes after the first dose of treatment that was seen in Phase 1. If the SD of the in-clinic systolic blood pressure (SBP) value is assumed to be 15 mm Hg, then the difference from the mean to the upper limit of a 95% confidence interval (CI) of the change from baseline in SBP relative to placebo is expected to be <4.7 mm Hg. If there is any initial rise in BP or PR, the in-clinic 6- and 12-week assessments along with the home blood pressure monitoring (HBPM) and ambulatory blood pressure monitoring (ABPM) measures will be used to assess whether the rise accommodates. The difference from the mean to the upper limit of a 95% CI in the change in average SBP measured at each of these visits relative to placebo is expected to be <3.5 mm Hg. This assumes an SD of 10 mm Hg and a 20% discontinuation rate.

Simulations were conducted in FACTs Version 5.1.1 using a Bayesian normal dynamic linear model (NDLM).

5.3. Method of Assignment to Treatment

Subjects who meet all criteria for enrollment will be randomized to double-blind treatment at Visit 3. Subjects will be randomized in a blinded fashion to the 4 treatment arms (LY3154207 10, 30, or 75 mg QD or placebo) in a 1:1:1:1 ratio. Randomization will use a minimization procedure in order to balance for investigator site and the current use of acetylcholinesterase inhibitors (AChEIs) (Yes, No) using an interactive web-response system (IWRS).

The IWRS will be used to assign blister packs containing double-blind investigational product to each subject. One blister pack will be dispensed for each week of dosing (for example, for 14-day study intervals, 2 blister packs will be dispensed at each visit). Site personnel will confirm that they have located the correct blister packs by entering a confirmation number found on the blister packs into the IWRS.

6. A Priori Statistical Methods

6.1. Populations for Analyses

For purposes of analyses, the following populations are defined:

Population	Description
Enrolled	All subjects who sign informed consent.
Evaluable Patient population (EPP)	All randomized subjects who received at least 1 dose of study medication, who have the baseline efficacy assessment, and have at least 1 post-dose efficacy assessment. Analyses will be according to the treatment the subject actually received.
Completers	All randomized subjects who have completed all study procedures including follow-up and/or early termination/discontinuation visit.
Safety population	All randomized subjects who take at least 1 dose of double-blind study treatment. In the event of a treatment error, subjects will be analyzed according to the treatment they actually received.
Per Protocol population	Subjects who complete the study with no major protocol deviations.
PK Analysis population	PK analyses will be conducted on subjects who receive at least 1 dose of the study drug and have 1 measurable concentration.

Abbreviations: PK = pharmacokinetic.

6.2. General Considerations

Statistical analysis for this study will be the responsibility of Eli Lilly Company (Lilly) or its designee. The statistical analyses performed as part of the interim analyses will be conducted by the designated statistical analysis center, which forms part of the internal assessment committee (IAC) (see Unblinding and Communication Plan for more details).

Analyses will be conducted in SAS Version 9.4 or greater, and Spotfire graphical package will be used to review safety data on an ongoing basis during the trial (for example, for trial level safety reviews). Packages to be used for the analysis of digital biomarker data will be specified in the Digital Biomarker Analysis Plan.

As HBEH is a Phase 2 study, the appropriate estimand is a de jure estimand, where efficacy of LY3154207 is assessed under the paradigm of all subjects taking study drug as intended. Intercurrent events for HBEH are defined to be when subjects discontinue the study prior to completing the 12 weeks of treatment. While subjects are allowed to be on steady doses of symptomatic treatments, the protocol does not allow for subjects to initiate any therapeutic agent that has been demonstrated to be effective for treatment of cognition in LBD during the trial. The primary analyses use a Bayesian mixed-model repeated measures (MMRM) analyses of the CDR-COA to evaluate cognition in subjects with mild-to-moderate LBD compared with placebo. This analysis assumes that the intercurrent events lead to data that is missing at random (MAR).

Unless otherwise specified, all efficacy analyses will be conducted on the Evaluable Patient population (EPP). Analyses will be conducted according to the treatment the subject actually received.

When change from baseline is assessed, subjects will be included in the analysis only if both the baseline and a postbaseline measure are available. Unless otherwise defined, a baseline measure is the last non-missing observation collected prior to the subject receiving study treatment.

Safety analyses will be conducted on the safety analysis set comprising all subjects who received at least 1 dose of study medication. Summary statistics, tabulations, and graphs will be provided as appropriate.

Longitudinal analyses from baseline to Week 12 will use the change from baseline score as the response variable. An unstructured covariance matrix will be considered for MMRM; however, if this fails to converge, other structures will be considered based on the trial design. Unless stated otherwise, for all MMRM analyses, the difference in least squares (LS) means between LY3154207 and placebo, corresponding 95% credible or CIs, and p-values will be presented.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the clinical study report (CSR).

Only key analyses and results relating to primary and secondary objectives will be presented in the CSR. Exploratory and sensitivity analyses are not expected to be included in the CSR, unless they are relevant to the trial conclusions. Further exploratory analyses, in addition to those specified in this SAP may be conducted, if deemed appropriate.

6.3. Adjustments for Covariates

Unless stated otherwise, the following covariates will be included in efficacy analyses: baseline, baseline by visit interaction, age, and concomitant use of AChEIs at baseline.

6.4. Handling of Dropouts or Missing Data

There will be no imputation for missing data unless specified otherwise.

General rule for the imputation of dates

If the adverse event (AE) onset date is missing or partial, the date will be compared as far as possible with the date of first dose of study medication when determining whether or not the AE is present at baseline. In this case, the AE will be assumed to be treatment emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the AE started prior to the first dose of study medication.

Efficacy analysis

The chosen method of analysis for the longitudinal data, MMRM, can handle missing data, as it uses all available data across subjects to estimate parameters. Under the assumption of data being MAR or the non-random missing data is ignorable, the estimates based on restricted maximum likelihood (REML) are unbiased. In the Bayesian analysis, under the assumption of MAR and distinct parameters determining the probability model and missingness mechanism, Bayesian inferences can be obtained for model parameters using a likelihood based only on the observed data (Gelman et al. 2004).

Repeated measures analyses will only use data from visits at which the data were scheduled to be collected (Andersen and Millen 2013). When subjects discontinue from the study early, there may be efficacy or safety data measurements at visits at which the variables were not scheduled to be collected. These data will be used in all other analyses.

Subjects with missing baseline data or with no postbaseline visit data will not be included. The only exception to this is for the primary analysis (see Section 6.5).

6.5. Handling Missing Items for Scales

Total and subscale scores for questionnaire-based safety and efficacy outcome measures will be derived from individual items. If any of the individual items are missing or unknown, every effort will be made to obtain the score for the missing item or items. If this is not possible, any total or sum involving that item will be considered missing. However, imputations will be adopted for the Alzheimer's Disease Assessment Scale – Cognitive subscale (ADAS-Cog) if complete data cannot be gathered and potentially for the baseline measurements of the primary outcome, CoA. For additional considerations due to the COVID-19 pandemic, see Section 6.12.7.

ADAS-Cog$_{13}$

For the 13-item Alzheimer's Disease Assessment Scale – Cognitive subscale (ADAS-Cog$_{13}$), if 4 or fewer of a total of 13 items (approximately 30%) are missing, the total score (maximum = 85) will be imputed as follows: The total from the remaining items will be multiplied by a factor that includes the maximum score for the missing items. For example, if the first item, "Word-Recall Task," which ranges from a score of 0 through 10 (maximum = 10), is missing, and the second item "Commands," which ranges from a score of 0 through 5 (maximum = 5), is missing then the multiplication factor = 85/(85 - [10 + 5]) = 85/70 = 1.21. Thus, the total score for this example will be the sum of the remaining 11 items multiplied by 1.21. The imputed number will be rounded up to the nearest integer. If more than 4 items are missing, the total score for ADAS-Cog$_{13}$ at that visit will be considered missing. The same imputation technique will be applied to the 11-item Alzheimer's Disease Assessment Scale – Cognitive subscale (ADAS-Cog$_{11}$) subscore.

CoA

Missing baseline data for CoA will be imputed using a regression approach. A simple linear regression model of CoA at baseline in terms of CoA scores at screening will be produced. Scores from the 2 screening assessments will be included as separate explanatory variables (first screening score and second screening score). This model will be based on data from subjects who have both screening and baseline CoA scores.

If missing, a subject's Visit 2 baseline score will be imputed with a predicted value from the model based on the subject's own CoA screening data. If a subject is missing both baseline and screening visit CoA data, the subject will be excluded from the analyses of CoA.

6.6. Multicenter Studies

Approximately 75 sites, predominantly in the United States (with some sites in Canada), will be participating in the trial. In the IWRS process, minimization will be adopted to maintain balance in treatment allocation at each site. Site is not an intended covariate for any statistical modeling, as the majority of sites are expected to only recruit a small number of subjects. Adjusting for such a large number of small sites would result in a high possibility of detecting a false site effect, due to multiplicity, and could also result in unreliable estimates of the treatment effect.

6.7. Multiple Comparisons/Multiplicity

No adjustments will be made for multiple comparisons.

6.8. Subject Disposition

All subjects who discontinue from the study will be identified, and the extent of their participation in the study will be reported. The reason for discontinuation and the time in days to study disposition will be summarized as frequencies and percentages.

A listing of the subject disposition will also be produced.

6.9. Subject Characteristics

Baseline characteristics will be summarized for the EPP and the per-protocol population by treatment group and overall. Summaries will include descriptive statistics for continuous measures and frequencies and percentages for categorical measures.

Subject characteristics that will be summarized include:

- age (continuous and categorized into <65, 65-84 and 85+)
- gender
- race (and ethnicity)
- height
- weight
- body mass index (weight (kg) / [height (m)]^2)
- country
- tobacco use at baseline
- alcohol use at baseline
- AChEI use at baseline
- memantine use at baseline
- dopaminergic therapy in levodopa equivalency dose (LED)
- time since Parkinson's disease (PD) or dementia diagnosis (whichever occurred first)
- diagnosis (PDD vs LBD)

- baseline severity of impairment as measured by CoA, Penn Parkinson's Daily Activities Questionnaire-15 (PDAQ-15), ADAS-Cog_{13}, Power of Attention (PoA), Delis-Kaplan Executive Function System (D-KEFS), Montreal Cognitive Assessment (MoCA), Neuropsychiatric Inventory (NPI), Epworth Sleepiness Scale (ESS), and Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS), and
- cardiovascular risk factors (diabetes, hypertension, hyperlipidemia, coronary artery disease, cardiovascular disease, and peripheral arterial disease).

6.10. Concomitant Therapy

Prior medications are defined as those that stop before randomization (Visit 3). Concomitant medications are defined as those being taken on or after randomization (Visit 3).

A summary of concomitant medications will be presented, with frequencies and percentages for each treatment group. If the start or stop dates of therapies are missing or partially missing, to the extent that it cannot be determined whether the therapy is prior or concomitant, the therapy will be deemed concomitant.

Summary tables will also be provided for PD medications, antidepressants, and cognitive enhancing medications; specifically, a separate summary for prior and concomitant use of AChEIs will be produced.

Investigators are permitted to adjust a subject's PD, cognitive, and antipsychotic medications in order to manage motor or other symptoms during the study; the frequency and reason for change in medication will be captured in a listing.

Medications will be coded using the World Health Organization (WHO) drug dictionary.

6.11. Treatment Compliance

Compliance is defined as having taken ≥80 % to <120 % of the prescribed amount of medication. Treatment compliance information for each treatment arm will be collected through counts of dispensed and returned capsules. The number of capsules taken relative to the number expected to be taken will be summarized.

6.12. Efficacy Analyses

6.12.1. Primary Outcome and Methodology

The primary objective is to evaluate whether LY3154207, administered at 10 mg, 30 mg, and/or 75 mg QD oral dosing for 12 weeks, will show an improvement in cognition compared to placebo in subjects with mild-to-moderate LBD. This objective will be assessed in the EPP using a Bayesian MMRM analysis of the change in the CDR-CCB CoA data from baseline to Week 12. Continuity of Attention (CoA) is defined as the sum of correct responses from the Choice Reaction Time and Digit Vigilance Task with the number of Digit Vigilance Errors subtracted.

The primary analysis on CoA will occur when all subjects have had the opportunity to complete 12 weeks of treatment.

The analysis of CoA will utilize a Bayesian MMRM model. The Bayesian analysis will use uninformative priors for all terms in the model. For the model coefficients, these will be diffuse Normal distributions centered on zero. Prior for covariance matrix will follow an inverse-Wishart distribution (the multivariate generalization of an inverse gamma distribution for a variance component). The MMRM model will account for longitudinal data assessed throughout the study after 1, 2, 4, 6, 8, 10, and 12 weeks of dosing. The change in CoA from baseline to Week 12 will be the dependent variable. The model will comprise fixed (baseline value, treatment, visit, age, and use of AChEIs) with the interaction terms (treatment by visit, baseline value by visit) with repeated measures of visit within subject. Visit will be considered a categorical variable with values equal to the visit numbers at which the CoA was assessed. Covariates may be centered prior to analysis to aid model convergence.

The primary comparison will be the contrast between treatments and placebo for the Week 12 change from baseline against a null hypothesis of no difference. This will be presented as both a difference in estimated means of the 12-week change and the effect sizes over the 12 weeks. The average effect size from the last 4 visit time points (Weeks 6, 8, 10, and 12) will be the primary comparison between treatments and placebo. These will be averaged after the model has been fit.

Example codes, for the purposes of aiding programming of the primary analysis, using Bayesian MMRM and equivalent frequentist MMRM, are given below.

Bayesian MMRM Model:

```
Proc MCMC data=coa_final nbi=2000 ntu=50 nmc=10000 seed=2019143 thin=10 diag=all
plots(smooth)=all dic outpost=coa_test_final_10000
        array CoA [9] x1 - x9;
        array mu [9] mu1 - mu9;
        array covar [9, 9];
        array s [9, 9];
        array sigma [9];
        array beta_v [8]   beta_v1 - beta_v8;     %* visit specific effect;
        array beta_t [3]    beta_t1 - beta_t3  ;     %* treatment specific effect;
        array beta_vt[8,3]  beta_vt1 - beta_vt24;    %* visit treatment interaction;
        array beta_bv[8]   beta_bv1 - beta_bv8;             %* baseline visit interaction;

        begincnst;
                call identity(s);
        endcnst;

        parms beta0 beta1 beta2 beta3 0;
        parms beta_v1 beta_v2 beta_v3 beta_v4 beta_v5 beta_v6 beta_v7 beta_v8  0 ;
        parms beta_t1  beta_t2  beta_t3 0;
        parms beta_vt1 beta_vt2 beta_vt3 beta_vt4 beta_vt5 beta_vt6 beta_vt7 beta_vt8  beta_vt9
        beta_vt10 beta_vt11 beta_vt12 beta_vt13 beta_vt14 beta_vt15 beta_vt16 beta_vt17 beta_vt18
        beta_vt19 beta_vt20 beta_vt21 beta_vt22 beta_vt23  beta_vt24  0;
        parms beta_bv1 beta_bv2 beta_bv3 beta_bv4 beta_bv5 beta_bv6 beta_bv7 beta_bv8  0;
        parms covar ;

        %macro tprior1;
        %do i = 1 %to 8;
                prior beta_v&i ~ N (0, prec=0.00001);
                prior beta_bv&i ~ N (0, prec=0.00001);
        %end;
        %do i = 1 %to 3;
                prior beta_t&i ~ N (0, prec=0.00001);
        %end;
```

```
        %do i = 1 %to 24;
               prior beta_vt&i ~ N (0, prec=0.00001);
        %end;
        %mend;
        %tprior1;

        prior beta0 ~ N (0, prec=0.00001);  %* overall intercept;
        prior beta1 ~ N (0, prec=0.00001);  %* baseline coefficient;
        prior beta2 ~ N (0, prec=0.00001);  %* age coefficient;
        prior beta3 ~ N (0, prec=0.00001);  %* AChEI coefficient;
        prior covar ~ iwish(9, S);

        do i=1 to 9;
        if trt > 1 then do;
               if i>1 then mu[i]=beta0+beta1*base+beta2*age+beta3*AChEI+beta_t[trt-1]+beta_v[i-
               1]+beta_vt[(i-1),(trt-1)] +beta_bv[i-1]*base ;
               else mu[i]=beta0+ beta1*base + beta2*age + beta3*AChEI+ beta_t[trt-1];
               end;
               else if trt eq 1 then do;
               if i>1 then mu[i]=beta0+ beta1*base + beta2*age + beta3*AChEI + beta_v[i-1] +
               beta_bv[i-1]*base;
               else  mu[i]=beta0 + beta1*base + beta2*age + beta3*AChEI;
               end;
        end;
        do i=1 to 9;
          Sigma[i] = sqrt(covar[i,i]);
        end;

        model coa ~ mvn(mu, covar);
run;
```

Corresponding Frequentist MMRM Model:

```
proc mixed data=efficacy_data order=data;
       class treatmentt subject visit AChEI;
       model CFB_CoA = base treatment visit visit*base treatment*visit age AChEI/
       ddfm=kr outp=pred solution residual;
       repeated visit/subject=subject type=un  ;
       lsmeans treament*visit  /pdiff cl alpha=0.05;
       ods output lsmeans=lsmean diffs=diff;
run;
```

6.12.2. Additional Analyses of the Primary Outcome

The primary efficacy analysis will be repeated with a per-protocol analysis set including only those subjects who completed the study with no major protocol deviations to verify robustness of the results.

6.12.3. Secondary Efficacy Analyses

The secondary efficacy outcomes (PDAQ-15, ADAS-Cog_{13}, CDR-CCB PoA, D-KEFS, MoCA, NPI, ESS, and MDS-UPDRS) will be assessed at the 12-week time point as the primary analysis. Each of these secondary endpoints will be analyzed using a frequentist MMRM model, which will include terms for: baseline value, treatment, visit, age, concomitant use of AChEIs, and the relevant interaction terms (treatment by visit, baseline by visit). The change in score from baseline will be the dependent variable. If subscores for a particular scale are deemed clinically relevant to analysis, it will be completed as an exploratory analysis.

For the Alzheimer's Disease Cooperative Study – Clinician Global Impression of Change (ADCS-CGIC), a logistic regression analysis will be performed with treatment, visit, age, concomitant medication of AChEIs, and relevant interaction terms. Two different endpoints will be assessed: "moderately or better," defined as subjects who had moderate/marked improvement and "minimally or better" defined as subjects who had minimal/moderate/marked improvement on the ADCS-CGIC scale.

Levodopa equivalency dose (LED) will be calculated using the conversion factors (Tomlinson et al. 2010) of the actual daily dose (mg) of the relevant PD drugs (Table HBEH.6.1). After the LED is calculated for each concomitant medication that a subject is taking, the total LED will be used as a covariate for the analyses of the MDS-UPDRS scales. The impact of LED may be explored for various efficacy and safety measures.

Table HBEH.6.1. LED Conversion Factors

Medication	Multiplication Factor
Immediate release L-dopa dose	1
Controlled release L-dopa dose	0.75
*Entacapone	LD × 0.33
*Tolcapone	LD × 0.5
Duodopa	1.11
Pramipexole (as salt)	100
Ropinirole	20
Rotigotine	30
Selegiline – oral	10
Selegiline – sublingual	80
Rasagiline	100
Amantadine	1
Apomorphine	10

Abbreviation: LED = levodopa equivalency dose.
Note: *LD represents immediate release L-dopa dose for particular patient.

6.12.4. Tertiary/Exploratory Efficacy Analyses

The following exploratory efficacy endpoints will be analyzed using an MMRM model:

- change in the total score for the ADAS-Cog_{11} from baseline to Week 12 (a subset of ADAS-Cog_{13})
- change in the CDR-CCB Cognitive reaction time composite score, Reaction Time Variability composite score, and Attention Ratio and Speed Score composite scores from baseline to Week 12
- change in the CDR-CCB individual task outcomes from baseline to Week 12 including simple reaction time, digit vigilance, choice reaction time, numeric working memory, spatial working memory, verbal memory, and visual memory
- change in total MDS-UPDRS (sum of Part I-III) and individual subscales non-motor experiences of daily living (Part I), motor experiences of daily living (Part II), motor exam (Part III), and motor complications (Part IV) subscales from baseline to Week 12

- change in items 4.3 through 4.6 (time off, impact of off, complexity of off, off dystonia) of the MDS-UPDRS from baseline to Week 12
- change in NPI score (Part D Depression/Dysphoria) from baseline to Week 12, and
- change in Geriatric Depression Scale (GDS) total score from baseline to Week 12.

Where appropriate, further exploratory analyses will be completed.

6.12.5. Sensitivity Analyses

A frequentist MMRM analysis of the primary analysis will be conducted.

An alternative MMRM analysis of change in CoA from baseline to Week 12, using the average change from baseline over Week 6 to Week 12 as the response variable, may also be considered. If there is a treatment difference prior to Week 12, which remains through to Week 12, this approach may be more sensitive than just looking at Week 12.

If baseline values are imputed from screening values in the primary analysis of CoA, a sensitivity analysis will be conducted to compare this approach with excluding subjects with missing baseline scores.

Differences in baseline characteristics of the treatment groups may be investigated using the Chi-squared test and 2 sample t-tests for categorical and continuous variables, respectively, if appropriate.

6.12.6. Dose Response Modeling

Dose response modeling may be utilized to understand the dose response relationship of select efficacy and safety measures including CoA, ADAS-Cog, MDS-UPDRS, SBP, diastolic blood pressure (DBP), and PR.

A trend test may be completed to assess evidence of a dose response by fitting a contrast statement in the MMRM model. In its simplest form, the analysis assigns equally spaced coefficients in the contrast, that sum to zero across the levels of the treatment factor (for example, -3, -1, 1, 3). This analysis assumes that the change in benefit from 75 mg compared to 30 mg is equivalent to the change from 30 mg to 10 mg and 10 mg to placebo. Alternative coefficients may be used if this assumption is inappropriate.

A Bayesian model averaging approach may be used to estimate the dose response relationship. This Bayesian model averaging approach is the Bayesian analog of the Multiple Comparison Procedure – Modeling (MCP-Mod) methodology (Bretz et al. 2005), and the Qualification of the MCP-Mod procedure (FDA 2015) is supportive in the use of MCP-Mod or Bayesian model averaging to assist in dose selection decisions.

Bayesian model averaging is a general mixture distribution, where each mixture component is a different parametric model. Prior weights are placed on each model and the posterior model weights are updated based on how well each model fits the data. Let $\mu(d)$ represent the mean of the dose response curve at dose d, $y = \{y_1, \ldots, y_n\}$ be the observed data, and $m \in \{1, \ldots, M\}$ be

an index on the M parametric models. Then the posterior of the dose response curve, $\mu(d)$, of the Bayesian model averaging model is

$$p(\mu(d) \mid y) = \sum_{m=1}^{M} p(\mu(d) \mid y, m)\, p(m \mid y)$$

$$p(m \mid y) = \frac{p(y \mid m)p(m)}{\sum_{m^*} p(y \mid m^*)p(m^*)}$$

where $p(\mu(d) \mid y, m)$ is the posterior mean dose response curve from model m, $p(m \mid y)$ is the posterior weight of model m, $p(y \mid m)$ is the marginal likelihood of the data under model m, and $p(m)$ is the prior weight assigned to model m. In cases where $p(y \mid m)$ is difficult to compute, Gould (2019) proposes using the observed data's fit to the posterior predictive distribution as a surrogate in calculating the posterior weights; this is the approach used in this analysis.

Potential candidate models that will be used are quadratic, log-quadratic, linear, log-linear, and EMAX models.

6.12.7. COVID-19

Due to the COVID-19 pandemic quarantine and travel restrictions, alternative methods of collecting key efficacy and safety data were implemented to allow subjects to safely remain in the trial.

Beginning in March 2020, enrolled subjects unable to attend in-person clinic visits due to COVID-19 restrictions were able to continue in the study by replacing in-person visits with telephone visits. All subjects had to have access to investigational product and an HBPM device to participate in the study remotely. During the telephone visits, investigators collected vital signs using the HBPM, AEs, concomitant medications, and the relevant information from the Columbia-Suicide Severity Rating Scale (C-SSRS). Additional efficacy outcomes were also collected, at appropriate visits, by phone and included the ADCS-CGIC, D-KEFS, PDAQ-15, MDS-UPDRS Parts I/II/IV, NPI, Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease (QUIP), Geriatric Depression Scale – Short Form (GDS-S), ESS, and physician withdrawal checklist-20 (PWC-20). In April 2020, the study implemented an option to allow web-based videoconference visits to assess additional measures that required visual inspection: the ADAS-Cog13, MDS-UPDRS Part III, and MoCA. The primary outcome, CDR-CCB CoA, was not collected remotely and is missing for all virtual visits. Visits conducted via telephone and/or video were documented within the study database. Given the variability in modality of collection among subjects, visits, and scales, sensitivity analyses will be conducted on the relevant scales excluding data collected via virtual visits.

Missingness due to COVID-19 does not depend on the outcome measurements and therefore will still be considered MAR. The MMRM analyses approach can handle missing values.

Since there is no anticipated interruption or discontinuation of treatment, the definition of an intercurrent event remains unchanged.

Additionally, a COVID-19 impacted population will include any subject after 01 March 2020 who had at least 1 virtual visit. This will allow for a variety of exploratory analyses regarding the impact of COVID-19 and mitigation strategies.

6.13. Pharmacokinetic/Pharmacodynamic Analyses

LY3154207 plasma concentrations will be illustrated graphically and summarized descriptively. Subject data, including but not limited to plasma LY3154207 concentrations, from other clinical studies evaluating LY3154207 may be combined with data from this study to support additional analyses.

If warranted and based on availability of data, the exposure-response relationship of plasma LY3154207 concentrations to efficacy endpoints and/or safety endpoints may be explored.

6.14. Safety Analyses

All subjects who receive at least 1 dose of double-blind study treatment will be evaluated for safety and toxicity.

Safety parameters (AEs, laboratory tests, vital signs, electrocardiograms [ECGs], and discontinuation rates) for meeting Day 1 (Visit 3) stopping rules will be summarized for the treatment period using descriptive statistics for continuous variables and frequencies and percentages for categorical variables.

Comparisons between the 3 LY3154207 doses and placebo will be made using a frequentist MMRM model for the in-clinic BP, PR, and HBPM measurements during the treatment period.

6.14.1. Extent of Exposure

Days of exposure will be calculated for all subjects as the (Date of last dose – date of first dose + 1). Summary statistics will be presented for the total number of days of exposure by treatment.

Cumulative exposure will be calculated for each subject and will be summarized with descriptive statistics by treatment arm.

A listing of drug accountability by subject will be created and the reason for any drug interruptions given.

6.14.2. Adverse Events and Serious Adverse Events

Treatment-emergent AEs will be defined as events that first occurred or worsened on or after randomization (Visit 3). Should there be insufficient data to determine whether an AE is treatment emergent, the AE will be considered treatment emergent.

Treatment-emergent AEs will be calculated based on adverse event identifier (AEID) and coded according to established Medical Dictionary for Regulatory Activities (MedDRA) terms and summarized by MedDRA System Organ Class (SOC) and Preferred Term. Adverse event terms and severity grades will be assigned by the investigator using Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

An overview of AEs, including the number and percentage of subjects who died, suffered serious AEs (SAEs), discontinued due to AEs, and who suffered treatment-emergent AEs (TEAEs), will be provided.

Summaries of AEs by decreasing frequency of Preferred Term within SOC will be provided for the following:

- preexisting conditions (medical history)
- TEAEs
- TEAEs by maximum severity
- TEAEs occurring in greater than 1% of subjects by Preferred Term
- SAEs, and
- AEs reported as reason for discontinuation.

These summaries will include number and percentages of subjects with TEAEs.

A summary of TEAEs by visit will also be provided.

SAEs and discontinuations due to AEs will be listed.

An SAE is any AE from this study that results in 1 of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (i.e. immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect
- important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the subject or may require intervention to prevent 1 of the other outcomes listed in the definition above. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse.

6.14.3. Deaths, Other Serious Adverse Events, and Other Notable Adverse Events

Any deaths will be listed and summarized.

Separate summaries by treatment group will be produced for any major cardiovascular TEAEs (including congestive heart failure, hypertension, myocardial infarction, stroke, and tachycardia) and major nervous system disorders TEAEs (including dyskinesias, impulse control disorders, delirium, confusion, and psychosis).

6.14.4. Clinical Laboratory Evaluation

Laboratory measures (clinical chemistry, hematology, and urinalysis) will be analyzed as continuous data (change from baseline) measured as International System of Units (SI) or as

categorical data (proportion of treatment-emergent abnormalities). If there are multiple records of laboratory measurements at baseline or postbaseline visits, the last record will be used.

Change from baseline to postbaseline visits at which laboratory measurements are taken will be summarized using descriptive statistics. A listing of abnormal laboratory data will be produced.

For all laboratory tests, frequencies of subjects with notable changes (that is, increases or decreases of a prespecified amount unique to each analyte) from baseline to each postbaseline visit will also be summarized for all subjects and stratified by low, normal, or high at baseline.

6.14.5. Vital Signs and Other Physical Findings

Vital Signs

Vital sign measurements and weight will be analyzed as continuous data (change from baseline) and the treatment-emergent abnormalities as categorical data.

If there are multiple records of vital sign or weight measurements at baseline or a postbaseline visit, the last record will be used. The only exception is for orthostatic BP and PR taken at Visit 1, Visit 4, Visit 5, Visit 6, Visit 7, Visit 9, Visit 10, unscheduled visit, and early termination/discontinuation visit and the initial measurement at time 0 of the in-clinic BP and PR monitoring. At these visits, 3 seated BP/PR measurements will be taken and the mean of these will be utilized for analysis purposes. Orthostatic changes will be calculated by comparing the mean of the 3 seated BP and PR values to the single standing BP and PR values.

The incidence of treatment-emergent abnormal high or low vital signs will be presented by treatment group and visit. Treatment-emergent vital sign evaluations are defined for evaluations collected after the initiation of study medication. Any vital sign meeting the criteria will be considered abnormal. For each vital sign at each postbaseline visit, only subjects who had a baseline result and had a non-missing result at that postbaseline visit will be included in the denominator when computing the proportion of subjects with treatment-emergent high, low, or abnormal values. Criteria for abnormal postbaseline vital signs are presented in Appendix 1.

The following will be summarized by treatment group and visit for all subjects in the Safety population: SBP, DBP, and PR (collected in sitting position); orthostatic SBP, orthostatic DBP, and orthostatic PR (measurement after 5 minutes in the seated position minus that after 3 minutes in the standing position); temperature; and weight. Body mass index at baseline will be summarized by treatment. Summary statistics will be presented for observed values at baseline and for change from baseline results at each scheduled postbaseline visit.

A listing of treatment-emergent abnormal vital signs by subject and visit will also be presented.

6.14.6. Electrocardiograms

The ECG measurements will be analyzed as continuous data (change from baseline) or as categorical data (proportion of treatment-emergent abnormalities).

Electrocardiogram is measured in triplicate during the double-blind period, and the average of triplicates will be used at baseline and each double-blind period visit. If there are multiple records after averaging ECG triplicates within a visit, the last record of averages will be used.

All ECG parameter changes from baseline to each postbaseline visit at which ECG measurements are taken will be summarized. All analyses of QTc will be carried out using the Fridericia's correction (QTcF) method.

Incidence of treatment-emergent abnormal ECGs will be also be summarized. For example, the frequency of subjects with abnormal treatment-emergent increases from baseline in QTcF intervals will be summarized into the following categories: >30 msec and >60 msec. Treatment-emergent elevated QTc intervals will be summarized into the following categories: >450 msec, >480 msec, and >500 msec. Baseline is the average of the last set of triplicates measured prior to taking study drug. Each category will show the number of subjects who have at least 1 treatment-emergent value meeting that criteria, and therefore those meeting the top criteria of >60 and >500 with be counted in lower categories as well.

6.14.7. Additional Safety Sections

6.14.7.1. Secondary Safety Analyses: In-Clinic Blood Pressure/Pulse Rate

- Day 1 (Visit 3) In-clinic BP and PR change post first dose

Change in in-clinic BP and PR from 0 up to 8 hours post dose measured on the first day of study drug dosing will be analyzed using an MMRM analysis with baseline value, treatment, time post dose (hours) as fixed effects, and treatment by time and baseline by time as interaction terms with repeated measures of time (hour) within subject. Two baselines will be considered: the Visit 3 pretreatment value (Time 0) and the time-matched baseline values from Visit 2 (hourly value 0-6 hours). For the second of these, there is no time-matched value for the Visit 3 7- and 8-hour time points, so these will use the Visit 2 6-hour time point as their baseline value. A separate change from baseline analysis will be completed for each baseline approach.

- Change in in-clinic BP and PR from baseline (Visit 2, predose) up to Week 12 (Visit 11)

To assess the impact of LY3154207 on BP and PR over time, change from baseline in in-clinic BP and PR to Week 12 will be analyzed using MMRM with response variable being the daily average of 0 to 6 hours, with fixed effects for baseline value, treatment, visit, and treatment by visit and baseline by visit as interaction terms with repeated measures of visit within subject. Change from baseline in BP and PR will be assessed at both 6 and 12 weeks. Analysis at 6 weeks will help to evaluate any significant changes (if any) in BP and PR earlier than 12 weeks (end of study) and if they accommodate by 6 weeks.

6.14.7.2. Tertiary/Exploratory Safety Analyses

The following exploratory safety endpoints will also be analyzed using an MMRM model adjusting for baseline reading:

- daily average change in HBPM measurements and PR from baseline to Week 12
- change in MDS-UPDRS items 4.1 and 4.2, time with dyskinesias, and impact of dyskinesias
- change in item 1.2 (hallucinations/psychosis) score of the MDS-UPDRS Part IA from baseline to Week 12
- change in score for the delusion (Part A) and hallucinations (Part B) domains of the NPI from baseline to Week 12
- change in QUIP score from baseline to Week 12, and
- change in PWC-20 total score will be analyzed using an analysis of covariance (ANCOVA) model with a covariate for baseline score. The primary analysis will be a change from baseline (Visit 11) to the in-person follow-up visit (Visit 801).

6.14.7.3. Home Blood Pressure Monitoring Analyses

Change from baseline in (i) average daily, (ii) morning, and (iii) evening home BP and PR to Week 3 will be analyzed using an MMRM with fixed effects for baseline value, treatment, day, and treatment by day and baseline by day interaction terms with repeated measures of day within subject. The baseline value for each analysis will be defined as the average of all relevant samples taken prior to first dose of LY3154207. A similar analysis will be completed for the 7 days before and after last dose. The MMRM analysis of the HBPM will provide daily estimates of change from baseline for the first 3 weeks of treatment. If there is a rise in BP and PR, the duration taken to accommodate can be assessed using the daily estimates of change from baseline. Visualizations will be used to characterize the temporal course of BP and PR changes, including accommodation. Additional exploratory analyses may be considered as necessary to further characterize the cardiovascular effects of LY3154207.

6.14.7.4. Addendum Safety Analyses

ABPM Addendum Objective

The objectives of the ABPM assessments are to explore any diurnal effects of LY3154207 on SBP and PR in subjects with PDD and whether these diurnal effects change over the study period—in particular, to evaluate the effect of LY3154207 on nocturnal dip in SBP.

ABPM Analysis

Only subjects participating in the ABPM addendum will be included in this analysis. The ABPM results will be summarized and represented graphically.

The ABPM will be used to assess the effect of LY3154207 on the dip in nocturnal SBP. A nocturnal dip is defined as a greater than 10% reduction in average SBP during the night compared to the average during the day. In order to reduce variability in the subject's sleep and awake period, the nighttime is defined as 0100 hours to 0600 hours. Similarly, the daytime, for purposes of analysis, is defined as 0900 hours to 2100 hours, when most subjects would be

expected to be awake, and the morning surge is defined as the mean of 4 readings over a period of 2 hours after subjects wake up from the lowest nocturnal SBP (mean of 3 readings centered around the lowest nighttime SBP). For each ABPM assessment period following dose administration (assessments will be between Visits 3 and 4, between Visits 7 and 8, and between Visits 10 and 11), the proportion of subjects who have a nocturnal dip will be summarized by treatment group, and a logistic regression will be performed to evaluate any treatment effect. The regression model will include the % reduction in SBP at baseline (measured between Visits 2 and 3) as a covariate. Odds ratios and associated 95% CIs will be presented.

Individual subject time courses will be reviewed graphically to investigate whether there is any morning surge effect.

In addition, the 24-hour average and the daytime and nighttime average SBP and PR changes from baseline at each ABPM assessment period will be analyzed using an MMRM analysis. Only the first 24 hours of data collected for each ABPM assessment period will be included. This will assess for any treatment effects and if there are any, whether they accommodate over the trial. The analyses will include fixed effects for baseline value, visit, treatment, treatment by time and baseline by time interaction with repeated measures of visit within subject. The time variable relates to the ABPM assessment period and will be assigned value of 1, 2, 3, or 4. An unstructured covariance matrix will be considered initially; however, if this fails to converge, other suitable structures will be considered. Baseline values will be defined as the 24-hour mean of the recordings taken following Visit 2; the mean of the recordings taken between 0900 hours to 2100 hours following Visit 2; and the mean of the recordings taken between 0100 hours to 0600 hours following Visit 2.

6.15. Subgroup Analyses

Change in CoA, ADAS-Cog_{13}, and CGIC from baseline to Week 12 may be assessed based on the following subgroups: concomitant PD therapy; use of concomitant AChEI therapy; use of memantine therapy; mild/moderate cognition (moderate: 10-16 total MoCA score; mild: 17-23 total MoCA score); and timing of dementia onset (less than or equal to 1 year versus greater than 1 year relative to onset of motor symptoms). These analyses will involve running the analysis outlined in Section 6.12.1 on subsets of the EPP according to each level of each variable listed above separately.

All subgroup analyses will be considered exploratory, and additional subgroup analyses may be performed as suggested by the data.

6.16. Protocol Deviations

Important protocol deviations that have the potential to compromise data integrity and subject safety will be listed and summarized by treatment group for the Safety population.

A list of the categories, subcategories, and study-specific terms of important protocol deviations can be found in the Trial Issue Management Plan v1.0 (available in the electronic trial master file [eTMF]).

Discontinuation due to meeting Day 1 cardiovascular safety criteria does not qualify as a protocol deviation.

6.17. Interim Analyses and Data Monitoring

A limited number of pre-identified individuals may gain access to the limited unblinded data, as specified in the unblinding plan. Information that may unblind the study during the analyses will not be reported to study sites or the blinded study team until the study has been unblinded.

The efficacy and safety interim analyses will be completed and reviewed by an IAC. Only the IAC is authorized to evaluate unblinded interim efficacy and safety analyses. Study sites will receive information about interim results ONLY if they need this information to ensure the safety of their subjects.

Safety interim analyses will be conducted on the number of subjects on each treatment who met the potentially clinically significant vital signs criteria at 3 consecutive time points at Visit 3 (Day 1 stopping rules). This will be done after 50, 100, and 150 subjects have completed Visit 3. CCI [redacted] Those already on 75-mg dose who passed the Day 1 stopping rules will remain on 75 mg. The specific analysis will be defined in the Interim Statistical Analysis Plan. In the event of an unacceptable rate of subjects meeting Day 1 stopping rules at other doses, adjustments to doses may be made for subsequently randomized subjects at the discretion of the IAC. A safety and efficacy interim analysis will be conducted when approximately 170 randomized subjects have had the opportunity to complete Visit 11 (Week 12) assessments. The statistical analysis will include all 170 subjects in the interim analysis, which will evaluate both efficacy and safety. In addition, the pharmacokinetic data may be reviewed. The efficacy analysis may be used for internal decision making, but is not planned to stop the study for efficacy or futility because this is an exploratory study and there are other key endpoints being assessed, not just the primary endpoint.

For more details, see the Unblinding and Communication Plan.

6.18. Planned Exploratory Analyses of Digital Biomarkers

Digital biomarkers (PRESENCE app and Actigraphy watch) data from baseline to Week 12 will be explored using appropriate visual, univariate, and multivariate methods to understand the relationship of the biomarker endpoints to other clinical endpoints in terms of treatment effect. The data may also be used to explore the validity and reliability of these assays in subjects with PDD. If applicable, potential predictive biomarkers associating with interpretable clinical benefit may be further examined with biological evidence. Further details are outlined in the Digital Biomarker Exploratory Statistical Analysis Plan.

6.19. Development Safety Update Report

Annual listings of deaths from any cause and discontinuations due to AEs during the current development safety update report (DSUR) period will be created for the DSUR each year until trial completion. The format of the tables will be based on the most current guidance document.

An age flag will be created to meet requirements of DSUR summary table: <65, 65-84, and 85+.

6.20. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

- Summary of AEs, provided as a dataset that will be converted to an XML file. Both Serious Adverse Events and 'Other' Adverse Events are summarized: by treatment group, by MedDRA Preferred Term.
    - An adverse event is considered 'Serious' whether or not it is a treatment-emergent adverse event (TEAE).
    - An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
        - the number of participants at risk of an event
        - the number of participants who experienced each event term, and
        - the number of events experienced.
    - Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
    - Adverse event reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

7. References

Andersen SW, Millen BA. On the practical application of mixed effects models for repeated measures to clinical trial data *Pharm Stat.* 2013;12(1):7-16. https://doi.org/10.1002/pst.1548

Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*. 2005;61(3)738-748. https://doi.org/0.1111/j.1541-0420.2005.00344.x

[FDA] United States Food and Drug Administration Office of Clinical Pharmacology Division of Pharmacometrics. Janssen Pharmaceuticals and Novartis Pharmaceuticals request for qualification of MCP-Mod as an efficient statistical methodology for model-based design and analysis of Phase II dose finding studies under model uncertainty. 2015. Accessed February 03, 2020. https://www.fda.gov/media/99313/download

Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. 2nd ed. CRC Press; 2004.

Gould AL. BMA-Mod: A Bayesian model averaging strategy for determining dose-response relationships in the presence of model uncertainty. *Biom J*. 2019;61(5):1141-1159. https://doi.org/10.1002/bimj.201700211

Rowan E, McKeith IG, Saxby BK, et al. Effects of donepezil on central processing speed and attentional measures in Parkinson's disease with dementia and dementia with Lewy bodies. *Dement Geriatr Cogn Disord.* 2007;23(3):161-167. https://doi.org/10.1159/000098335

Tomlinson CL, Stowe R, Patel S, et al. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov Disord*. 2010;25(15):2649-2653. https://doi.org/ 10.1002/mds.23429

Wesnes KA, McKeith I, Edgar C, et al. Benefits of rivastigmine on attention in dementia associated with Parkinson disease. *Neurology.* 2005;65(10):1654-1656. https://doi.org/10.1212/01.wnl.0000184517.69816.e9

8. Appendices

Appendix 1. Potentially Clinically Significant Vital Signs

Parameter	Unit	Observed Value	Change Relative to Baseline	Change from Supine to Standing
SBP (supine or sitting)	mm Hg	≥180 ≤90	Increase of ≥20 Decrease of ≥20	– –
DBP (supine or sitting)	mm Hg	≥105 ≤50	Increase of ≥15 Decrease of ≥15	– –
Pulse (supine or sitting)	bpm	≥120 ≤50	Increase of ≥15 Decrease of ≥15	– –
Orthostatic hypotension (systolic)	Decrease in SBP when going from 5 minutes supine to 3 minutes standing of ≥20 mm Hg or the inability to stand quickly for the measurements due to symptoms of orthostasis			
Orthostatic hypotension (diastolic)	Decrease in DBP when going from 5 minutes supine to 3 minutes standing of ≥10 mm Hg or the inability to stand quickly for the measurements due to symptoms of orthostasis			
Orthostatic pulse (tachycardia)	Increase in heart rate when going from 5 minutes supine to 3 minutes standing of ≥30 bpm			

Abbreviations: DBP = diastolic blood pressure; SBP = systolic blood pressure.