Statistical Analysis Plan I4L-GH-ABES

A Prospective, Randomized, Open-Label Comparison of a Long-Acting Basal Insulin Analog, LY2963016, to Lantus® in Combination with Mealtime Insulin Lispro in Adult Chinese Patients with Type 1 Diabetes Mellitus

NCT03338023

Approved Date: 12Mar2018

1. Statistical Analysis Plan:

I4L-GH-ABES

A Prospective, Randomized, Open-Label Comparison of a Long-Acting Basal Insulin Analog, LY2963016, to Lantus® in Combination with Mealtime Insulin Lispro in Adult Chinese Patients with Type 1 Diabetes Mellitus

---

Confidential Information

The information contained in this document is confidential and the information contained within it may not be reproduced or otherwise disseminated without the approval of Eli Lilly and Company or its subsidiaries.

Note to Regulatory Authorities: this document may contain protected personal data and/or commercially confidential information exempt from public disclosure. Eli Lilly and Company requests consultation regarding release/redaction prior to any public release. In the United States, this document is subject to Freedom of Information Act (FOIA) Exemption 4 and may not be reproduced or otherwise disseminated without the written approval of Eli Lilly and Company or its subsidiaries..

LY2963016

Phase 3, randomized, multicenter, 2-arm, active-control, open label, parallel, 24-week treatment study to compare LY2963016 and Lantus® with mealtime insulin lispro in adult Chinese patients with type 1 diabetes mellitus, with 4-week post-treatment follow up.

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I4L-GH-ABES
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly on date provided below.

2. Table of Contents

Table of Contents

3.   Revision History

SAP Version 1 was approved prior to the First Randomized Patient.

4. Study Objectives

4.1. Primary Objective

The primary objective of this study is to test the hypothesis that LY2963016 QD is noninferior to Lantus® QD by a margin of 0.40%, as measured by change in hemoglobin A1c (HbA1c) from baseline to 24 weeks, when used in combination with premeal insulin lispro administered thrice a day (TID).

4.2. Secondary Objectives

The secondary objectives of the study are as follows:

- To test the hypothesis that Lantus® is noninferior to LY2963016, as measured by change in HbA1c from baseline to 24 weeks, when used in combination with premeal insulin lispro (TID) (this secondary objective is tested with a gated approach).

- To compare the safety of LY2963016 to Lantus® (proportion of patients with detectable of anti-glargine antibodies, hypoglycemia and injection site reaction) when used in combination with premeal insulin lispro.

- To compare change in HbA1c at 6, 12, and 18 weeks between LY2963016 and Lantus® when used in combination with premeal insulin lispro

- To compare 7-point self-monitored blood glucose (SMBG) profiles at baseline, 2, 4, 6, 12, 18, and 24 weeks between LY2963016 and Lantus® when used in combination with premeal insulin lispro

- To compare percentage of patients with HbA1c <7% and percentage of patients with HbA1c ≤6.5% at 6, 12, 18, and 24 weeks between LY2963016 and Lantus® when used in combination with premeal insulin lispro

- To compare LY2963016 to Lantus® when used in combination with premeal insulin lispro with regard to the following measures
    - intrapatient blood glucose (BG) variability
    - basal and prandial insulin dose
    - weight change

To compare patient-reported outcomes (PRO) between LY2963016 and Lantus® as measured by responses to the Insulin Treatment Satisfaction Questionnaire (ITSQ).

5.   A Priori Statistical Methods

5.1.   Determination of Sample Size

5.1.1.   Sample Size Section from the Protocol

Based on the primary objective, to show noninferiority of LY2963016 to Lantus® at the 0.40% noninferiority margin (NIM), 109 completers per arm (218 total) are needed at 24 weeks. This calculation assumes no treatment difference in HbA1c between LY2963016 and Lantus®, a common standard deviation (SD) of 1.05% for change from baseline in HbA1c, a two-sided significance level of 0.05, and over 80% power. Assuming a 15% dropout rate at 24 weeks, the required number of randomized patients is 129 per arm (258 total). This sample size is noted as the minimum sample size required in the study.

Blinded sample-size reestimation (bSSR) may be performed when approximately 40% of the subjects have been enrolled in the study and finished 24-weeks of treatment, or about 6 months prior to the completion of enrollment, depending on which occur first by team assessment. This reestimation will use a statistical model to estimate the variability in the change in HbA1c from baseline to 24 weeks using all available patient HbA1c values at the time of data cutoff. The estimate of variability will then be used to recalculate the sample size that would be needed to have 80% power for a NIM of 0.4%, assuming no difference between treatments. The sample size from the study is constrained between a predefined minimum sample size of 258, and a maximum sample size of 400 patients. If the recalculated sample size is smaller than minimum sample size planned, the study will enroll to the minimum sample size. If the recalculated sample size is larger than the planned minimum sample size, the team will make a decision whether to increase sample size to the reestimated sample size, bounded by the predefined maximum, or accept the consequent reduction in power.

5.1.2.   Blinded Sample Size Re-estimation

A bSSR procedure will be utilized to ensure that the trial has an appropriate sample size to achieve the pre-specified power. This procedure is blinded, in that it does not use information on treatment assignments. Only pooled data from a blinded snapshot will be used to estimate the variance of the primary endpoint. Based on the estimated variability, the final study size will be calculated. The allowable range of the final sample size is constrained in a pre-specified manner, subject to the planned and maximum sample size. The bSSR procedure described below was evaluated via simulation and is efficient in determining the appropriate sample size under a range of possible scenarios without inflating type 1 error. The methodology and implementation details are described below.

The bSSR procedure has four steps,

1) At a blinded snapshot, a Bayesian predictive model will use the available data to generate predictions (multiple imputations) of all future HbA_{1c} data yet to be observed in the trial using the available data at the snapshot (see Section 5.1.2.2).

2) These predicted datasets will each be analyzed using the primary analysis model. The residual errors from each model will be used to determine a predictive distribution of the total variability of change in HbA$_{1c}$ from baseline to the end of the study.
3) The predictive distribution of the variability will be used to determine the recommended sample size.
4) The study team will consider the recommended sample size and decide on the final sample size.

For practical reasons, the final sample size is constrained by a pre-specified minimum, N_{min}, and maximum, N_{max}. Therefore if the estimated final sample size is greater than N_{max}, 400 is recommended. If the estimated final sample size is less than N_{min}, 258 is recommended. The resulting estimated sample size is a recommendation, is non-binding, but represents the model estimate of the appropriate sample size. The final decision on the sample size would be subject to constraints from the maximum and minimum sample sizes. Factors influencing the final sample size are described in Section 5.1.2.4.

5.1.2.1. Traditional Power Analysis versus Blinded SSR for a Non-Inferiority Trial

When powering a non-inferiority trial, assumptions need to be made to determine the sample size for a pre-specified type I (α) and type II (β) error level. The non-inferiority margin is denoted as δ, the assumed difference is μ, the fraction dropping out of the study is f, and the assumed standard deviation of the outcome variable is σ. Putting this together the total number of patients to enroll (to ensure that the completers have the pre-specified power) for the trial can be estimated by,

$$N = 4 * \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{(\mu - \delta)^2/\sigma^2} \left(\frac{1}{1-f}\right).$$

When there is considerable uncertainty in the value of σ, a conservative assumption may be made, such that it is assumed to be sufficiently large to cover the range of the literature. This leads to larger sample sizes than are needed to achieve the desired level of confidence in the objective of interest.

Blinded SSR will be used to determine the sample size necessary to achieve the desired power using data from a blinded snapshot. These data will be used to estimate σ, denoted $\hat{\sigma}$. The sample size is estimated, $\widehat{N}$, using,

$$\widehat{N} = 4 * \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{(\mu - \delta)^2/\hat{\sigma}^2} \left(\frac{1}{1-f}\right)$$

with $\alpha = 0.025$ (one-sided), $\beta = 0.05$, $\mu = 0$, $\delta = 0.4\%$, $f = 0.15$. Blinded SSR is most effective in the cases where the therapies are expected to have similar effects, $\mu \approx 0$. Since this assumption is valid in this study, i.e. both treatments are expected to have similar mean responses, the data could be pooled to estimate variability.

5.1.2.2. Timing and Data used for the Blinded SSR

When approximately 40% of the pre-specified minimum subjects have been enrolled in the study and finished 24-weeks of treatment, or about 6 months prior to the completion of enrollment, depending on which occur first by team assessment, a blinded snapshot of the study data will be used to estimate the necessary final sample size. The snapshot will contain a scrambled patient identifier that is unique to each patient (but different than their unique subject identifier used in the study), the HbA$_{1c}$ values corresponding to each patient visit, the country of enrollment of the patient, and an indicator of whether the patient dropped out, completed, or is currently ongoing in the study, and other stratification factors. This information will be modeled to estimate the variability in HbA$_{1c}$ at the final visit and therefore inform the decision on final sample size. The treatment assignment will not be provided in the data snapshot and will not be used to estimate variability.

5.1.2.3. Blinded SSR Methodology

At the time of the blinded SSR, all available HbA$_{1c}$ data in the database will be used. The HbA$_{1c}$ value for subject i, and time t_j is denoted as y_{ij}, the total duration of the study is d, the subject random effect is, s_i, and the residual error is, ϵ_{ij}. The model for HbA$_{1c}$ a Bayesian Integrated Two-Component Prediction (ITP) (Fu, 2010) and is written as,

$$y_{ij} = \left(\theta + s_i + \epsilon_{ij}\right)\frac{\left(1 - e^{-kt_j}\right)}{\left(1 - e^{-kd}\right)}$$

with the between subject error,

$$s_i \sim N(0, \sigma_b^2),$$

and within subject error,

$$\epsilon_{ij} \sim N(0, \sigma_w^2).$$

This model is Bayesian and requires specification of prior distributions for each parameter. For each parameter, non-informative priors have been chosen,

$$k \sim \text{Uniform}(0.001, 10),$$

$$\theta \sim N(-1, 1000),$$

$$\sigma_b^2 \sim IG(0.01, 0.01),$$

and

$$\sigma_w^2 \sim IG(0.01, 0.01).$$

This Bayesian model will be fit using Markov Chain Monte Carlo (MCMC) with the OpenBUGS package in R. The MCMC algorithm will use one chain, with initial values for the parameters drawn from their corresponding prior distributions. An appropriate burn-in period will be used of at least 10,000 iterations to ensure that the MCMC chain has reached convergence. After the burn-in period, 10,000 MCMC samples will be obtained for all parameters in the model, as well as all missing y data. Therefore each MCMC sample contains a complete predicted dataset,

where all unobserved y data are imputed. These 10,000 predicted datasets will be used to estimate the variability in HbA_{1c}. To ensure independence of the MCMC samples, the chains will be thinned by at least 10 to ensure that the 10,000 selected samples are not highly correlated. Diagnostics will be assessed to ensure convergence of the chain.

For each of the 10,000 MCMC samples, the primary analysis described in Section **Error! Reference source not found.** (containing all terms other than treatment assignment), will be performed using the complete dataset corresponding to each imputation step. For each of these k analyses, $k=1$ to 10,000, the total variability will be estimated, as $\hat{\sigma}_{T,k}$ by evaluating the residual error of the model. For each value of $\hat{\sigma}_{T,k}$, the estimated sample size will be computed using the sample size formula introduced in the previous section for a non-inferiority trial,

$$\widehat{N}_k = 4 * \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{(\mu - \delta)^2/\hat{\sigma}_{T,k}{}^2} \left(\frac{1}{1-f}\right)$$

with $\alpha = 0.025$ (one-sided), $\beta = 0.05$, $\mu = 0$, $\delta = 0.4\%$, $f = 0.15$. The mean value of the 10,000 $\widehat{N}_k$ values, $\overline{N}$ as well as the distribution of $\widehat{N}_k$ will be reported. The recommended sample size, $\widehat{N}$, is then subject to constraints from the maximum and minimum sample sizes, using.

$$\widehat{N} = \min(\max(\overline{N}, N_{\min}), N_{\max})$$

5.1.2.4. Determination of the Final Sample Size
The recommended sample size, $\widehat{N}$, is non-binding, but the study team will use the recommendation when making the decision regarding the final sample size. The decision on the final sample size will consider the impact of the recommended sample size on two primary factors, 1) the feasibility of attaining the recommended sample size given the observed enrollment rate, and 2) the degree of certainty in the recommended sample size. Regardless of this decision process, the final target sample size will be no smaller than approximately $N_{\min}$ (258 for this study) and no larger than approximately $N_{\max}$ (400 for this study).

5.2. Statistical and Analytical Plans

5.2.1. General Considerations
All data will be entered, verified, and archived at a contract research organization (CRO) external to Lilly and/or at Lilly. Data listings, summaries, and analyses will be performed by a CRO and/or by Lilly under the guidance and approval of statisticians at Lilly. Statistical analysis of this study will be the responsibility of Lilly.

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 statistical analysis plan and/or in the clinical study report (CSR). Additional exploratory analyses will be conducted, as deemed appropriate.

The patient populations used in the study are described below:

1. All Patients Entered: all patients who entered this study and completed Visit 1.

2. All Randomized: all patients who were randomized to a treatment arm.

3. Full Analysis Set (FAS): based on the intent to treat (ITT) principle, all patients who were randomized and who have taken at least one dose of study medication. Patients are assigned to the treatment arm to which they were randomized.

4. Per-protocol (PP): patients in the FAS/ITT population who also meet the following criteria:

    a. violate no inclusion or exclusion criteria

    b. have not discontinued from the study prior to 24 weeks

    c. have not been off study medication for more than 10 consecutive days during the treatment period

    d. have not received chronic (lasting longer than 14 consecutive days) systemic glucocorticoid therapy (excluding topical, intra-articular, intraocular, and inhaled preparations)

Unless otherwise specified, listings will be prepared using all randomized patients. Efficacy and safety analyses will be conducted using the FAS population. Selected analyses will be conducted using the All Randomized population and the PP population.

Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and confidence intervals (CIs) will be calculated as 2-sided 95% CIs. All tests of interactions between treatment groups and other factors will be conducted at a 2-sided alpha level of 0.05. No adjustments for multiplicity will be performed. See Section 5.2.7 for the gate-keeping strategies used for primary/secondary endpoints.

The baseline is Visit 2. If baseline data are missing, the last measurement taken prior to this visit will be used for the baseline measurement.

The last visit for the 24-week treatment study is Visit 13. If the Visit 13 measurement is missing, the last post-baseline value will be carried forward to create the 24-week endpoint value using the last-observation-carried-forward (LOCF) methodology. If there are no measurements after Visit 2 (date of randomization), the 24-week endpoint values will be considered missing. If the baseline value is missing or a patient only has a baseline value with no follow-up values, then the patient will not be included in the efficacy analyses. The LOCF methodology will be utilized in the endpoint analyses using analysis of covariance (ANCOVA) of HbA_{1c}, laboratory chemistry and hematology, as well as analysis of hypoglycemia (rate and incidence) and insulin antibodies.

Unless otherwise noted, the analysis of the continuous secondary efficacy and safety variables (SMBG at each time point and summaries of SMBG, weight, dose, and vital signs) will be performed using the same mixed model repeated measures (MMRM) methodology for the primary efficacy analysis with the baseline value of the response variable added as a covariate with the FAS population. Continuous laboratory measures will be analyzed using the ANCOVA model.

Unless otherwise noted, the analysis of categorical variables to compare the 2 treatments will be performed using Fisher`s exact test or Pearson's chi-square test.

Values for the 7-point self-monitored blood glucose (SMBG) blood glucose profiles will be averaged over the two 7-point SMBG profiles obtained during 2-week period prior to each office visit (as specified in Protocol Section 7.2.3). For the average blood glucose calculation for a specific time point, if only 1 of 2 days of data is collected, then the value of the 1 day will be used. If 2 days of data are collected, then the average of the 2 days will be used. If more than 2 days of data are collected then choose the 2 according to the ranking first by the number of non-missing blood glucose measurements and then by day closest to the visit day. The handling of missing data for the ITSQ questionnaire is specified in Section 5.2.10.

Visits for laboratory data will be handled by conventions in the Lilly diabetes white paper (Standardization of Derivations for Efficacy Lab Variables).

All analyses will be implemented using SAS Version 8.2® or higher.

5.2.2. Patient Disposition

A listing of the primary reason for patient discontinuation will be presented for All Randomized patients. Summary analyses will be conducted for the All Randomized and FAS populations. Frequency counts and percentages will be presented for each treatment group and compared across treatment groups using Fisher's Exact Test or Pearson's Chi-square test.

A listing and summary analyses will also be provided for the discontinuation of study drug for the FAS population.

5.2.3. Patient Characteristics

Demographic and baseline characteristics will be summarized by treatment group for all randomized patients, the FAS and PP populations. For continuous measures, summary statistics will include sample size, mean, median, maximum, minimum, and standard deviations (SDs). The treatment groups will be compared using a 2-sample t-test. For categorical measures, summary statistics will include sample size, frequency, and percentage. Analysis will use Fisher's Exact test or Pearson's Chi-square test.

5.2.4. Concomitant Therapy

Concomitant medications, including previous therapies for diabetes, will be summarized by different categories and treatment group using the FAS population. All concomitant therapies that originally mapped using the WHODRUG dictionary in the Clintrial database will be further classified using Anatomical Therapeutic Chemical (ATC) level 2 codes for reporting purpose. Analyses will use test Fisher's Exact Test or Pearson's Chi-square test.

5.2.5. Treatment Compliance

No specific study data will be collected for analysis of treatment compliance.

5.2.6. Protocol Violations

A comprehensive listing of patients with important protocol deviations (IPD) or protocol violations that could potentially impact data interpretation, data integrity and patient safety across the I4L-GH-ABES study will be provided. A summary of IPD by treatment group and overall will also be provided.

Important protocol deviations will be identified from the clinical database and from site monitoring. Categories of IPDs will be documented in the "ABES trial issue management plan" that will contain detailed criteria used to identify IPDs. Detailed programming specifications for IPDs (ADaM PD dataset specification) will be stored in CLUWE prior to database lock.

5.2.7. Primary Efficacy Outcome and Methodology

The primary efficacy outcome will be the change in HbA1c level from baseline to 24 weeks. The primary analysis will be a likelihood-based, MMRM approach, treating the data as missing at random (MAR) for the FAS population. The MMRM model will evaluate the change from baseline to each postbaseline visit in HbA1c level as the dependent variable with treatment (LY2963016, Lantus®), pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither), visit, and interaction between visit and treatment as fixed effects; the baseline value of HbA1c as a covariate; and a random effect for patient. The MMRM model will be carried out using the observed margins (OM) option in SAS MIXED procedure. The standard least square means (LSMeans) have equal coefficients across classification effects; however, the OM option changes these coefficients to be proportional to the number of patients used to calculate the classification means. This will provide LSMeans estimates that are more representative of the patient population recruited in this study.

The primary treatment comparison is to compare LY2963016 versus Lantus® at the NIM of +0.4%. If the upper limit of the 95% CI on the change from baseline to 24-week endpoint HbA1c for LY2963016 versus Lantus® is below +0.4%, then LY2963016 will be declared noninferior to Lantus®. The LSMean and standard error derived from the MMRM model for each treatment will be used to test noninferiority. Type III sums of squares will be used to make the treatment comparisons.

- If the +0.4% NIM is met, a key secondary treatment comparison will be conducted to show noninferiority of Lantus® to LY2963016. This will be done by comparing Lantus® versus LY2963016 at the NIM of -0.4%. If the lower limit of the 95% CI on the change in HbA1c from baseline to the 24-week endpoint for LY2963016 versus Lantus® is above -0.4%, then Lantus® will be declared noninferior to LY2963016. The LSMean and standard error derived from the MMRM model for each treatment will be used to test noninferiority. This gate-keeping procedure controls the family-wise Type 1 error rate at a 1-sided 0.025 level.

- If LY2963016 is declared noninferior to Lantus® in the primary treatment comparison, and Lantus® is declared noninferior to LY2963016 in this secondary treatment comparison, then LY2963016 will be considered to have equivalent efficacy as Lantus®.

A first secondary analysis of the primary efficacy outcome will use the same MMRM model described above with the PP patient population. Significance tests will be based on LSMeans using the Type III sum of squares, and testing for noninferiority will occur as described above.

A second secondary analysis of the primary efficacy outcome will use an ANCOVA model with FAS population. The ANCOVA model will include treatment, pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) as fixed effects and the baseline value of HbA1c as a covariate. The ANCOVA model will be carried out using the OM option in SAS GLM procedure. If the 24-week HbA1c value is missing, the last postbaseline value will be carried forward and used in the analysis. This creates the 24-week endpoint value for HbA1c using the LOCF methodology. If there are no HbA1c data after the date of randomization, the endpoint will be considered missing and the patient will not be included in the analysis.

The analyses of the primary efficacy outcome will only be conducted for patients with both nonmissing baseline value and at least 1 nonmissing postbaseline value.

5.2.8. Secondary Efficacy Outcomes and Methodology

The continuous secondary efficacy outcomes include:

- actual and change in HbA1c from baseline to 6, 12, and 18 weeks or LOCF

- 7-point SMBG measurements and change from baseline to each visit (expressed as plasma-equivalent glucose values), as listed in the Study Schedule (see ABES Study Protocol Attachment 1)

  - premeal for each meal

  - postmeal for breakfast and lunch

  - bedtime

  - 3 am

- summaries of 7-point SMBG (actual and change from baseline)

  - daily mean BG level (average across all 7 time points)

  - daily mean premeal BG level (before breakfast, lunch, dinner)

  - daily mean postprandial BG level (breakfast, lunch and bedtime)

  - bedtime to 3 AM excursion

- intrapatient variability as measured by the SD of the 7-point SMBG.

- actual and change from baseline total insulin, basal insulin, and mealtime insulin lispro doses in U/day and U/kg/day

- actual and change from baseline in weight and BMI

- patient-reported outcomes as reflected in responses to ITSQ.

The analysis of the continuous secondary efficacy variables will be performed using the same MMRM model with treatment (LY2963016, Lantus®), pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither), visit, and interaction between visit and treatment as fixed effects; the baseline value of the response variable and the baseline HbA1c level as covariates; and a random effect for patient, with the FAS population.

The proportions of subjects achieving HbA1c target values (HbA1c <7.0% and ≤6.5%) at any point during the study (Weeks 6, 12, 18, 24, and 24-week endpoint [LOCF]) will be analyzed using Fisher's Exact test or Pearson's Chi-square test.

A listing of 7-point SMBG and 4-point SMBG will be presented for all the randomized subjects.

5.2.9. Pharmacokinetic/Pharmacodynamic Analyses
Not applicable.

5.2.10. Health Outcome/Quality of Life Analyses
The Insulin Treatment Satisfaction Questionnair (ITSQ) will be completed at baseline (Visit 2), prior to randomization and at Week 24 (Visit 13 or ED).

The following will describe the analyses details for ITSQ.

5.2.10.1. Insulin Treatment Satisfaction Questionnaire
The ITSQ is a validated instrument containing 22 items that assesses treatment satisfaction completed at weeks 0 (visit 2), and 24 (visit 13) or early discontinuation for persons with diabetes who are taking insulin. Items are measured on a 7-point scale in which lower scores reflect better outcomes. In addition to an overall score, the items that make up the 5 domains of satisfaction are categorized as:

- Inconvenience of Regimen (IR - 5 items)

- Lifestyle Flexibility (LF - 3 items)

- Glycemic Control (GC - 3 items)

- Hypoglycemic Control (HC - 5 items)

- Insulin Delivery Device (DD - 6 items).

All individual patient domain scores and overall scores will be calculated as the sum of the individual item scores in the domain or overall. If an item score is missing for a patient and ≤20% of the items within the domain (or overall) are missing for that patient, then impute the missing item score as the mean of all other patients with a non-missing same item score.

Individual patient with more than 20% missing items in a domain (or overall), the domain (or overall) score will be set to missing.

The individual raw and transformed domain scores are then calculated as follows: first, calculate the raw domain score as the mean of the items in the domain for that patient. Then, use the following formula to transform the raw domain scores on a scale from 0-100 where higher scores indicate better treatment satisfaction:

Transformed domain score= 100*[(7- raw mean domain score)/6]

The individual raw and transformed overall scores are then calculated as follows: first, calculate the raw overall score as the mean of the items for that patient. Then, use the following formula to transform the raw overall score on a scale from 0-100 where a higher score indicates better treatment satisfaction:

Transformed overall score= 100*[(7- raw mean overall score)/6]

The change from baseline to Week 24 or end of study transformed domain and overall scores will be analyzed using the ANCOVA model for the FAS population. The ANCOVA model will include treatment (LY2963016, Lantus®), pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) as fixed effects and the baseline value of HbA1c as a covariate.

The transformed domain and overall scores will be analyzed using the ANCOVA model for the FAS population. The ANCOVA model will have treatment (LY2963016, Lantus®), pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) as fixed effects and the baseline value of HbA1c as a covariate.

5.2.11. Safety Analyses

5.2.11.1. Adverse Events

Adverse events will be listed by patient, system organ class, Medical Dictionary for Regulatory Activities® (MedDRA) preferred term, severity, and relationship to the study disease, drug, device, or procedure for all patients. Adverse events will be summarized as treatment-emergent adverse events (TEAEs) for the FAS population. Treatment-emergent adverse events are defined as events that are newly reported after first study treatment following randomization, or reported to worsen in severity from baseline. The proportion of patients experiencing each TEAE will be presented by preferred term, system organ class (SOC), and treatment group. The proportion of patients experiencing each TEAE that are assessed as possibly related to the study disease, drug, device, or procedures will also be summarized. The number and proportion of patients will be presented and compared by treatment using Fisher's exact test or Pearson's chi-square test for the FAS population. TEAEs will be summarized by preferred term within SOC and by preferred term by decreasing frequency.

All SAEs will be listed by patient and summarized by treatment as counts and percentages. If a sufficient number of SAEs are reported, the proportion of patients with SAEs between treatment groups will be compared using Fisher's exact test. Similar analyses will be performed for discontinuations due to AEs.

5.2.11.1.1. Special Topic Assessment of Allergic Events

The special topic assessment of allergic events will be performed by an initial blinded review of preferred terms (PTs) by SOC in order to identify all possible cases of allergic events, followed by a comparison between treatment arms. The goal of the initial blinded review, which will be carried out by the safety physician, is to identify all reported allergic reactions. As an initial reference for this blinded review, Lilly will consider the list of PTs shown in Appendix 1. Justification in including or excluding events based on medical judgment and other supportive information will be provided when applicable. All allergic events will be listed by patient and summarized by treatment as counts and percentages. The proportion of patients with allergic reactions between treatment groups will be compared using Fisher's exact test or Pearson's chi-square test for the FAS.

5.2.11.1.2. Assessment of Injection Site Adverse Events

Whenever an injection site adverse event occurs, there is an evaluation of the pain, pruritus, and rash associated with the injection as well as of the characteristics of the injection site (abscess, nodule, lipoatrophy, lipohypertrophy, or induration). The proportion of patients experiencing injection site adverse events will be summarized and analyzed by treatment group. Additional analyses will be done as deemed appropriate.

5.2.11.2. Hypoglycemic Events

A **hypoglycemic episode** is that at any time a patient feels as he/she is experiencing a sign or symptom that is associated with hypoglycemia or has a BG level of ≤70 mg/dL (≤3.9 mmol/L), even if it was not associated with signs, symptoms, or treatment consistent with current guidelines (Seaquist et al. 2013).

Severe hypoglycemia: an event requiring assistance of another person to actively administer carbohydrates or glucagon or take other corrective actions. Plasma glucose concentrations may not be available during an event, but neurologic recovery following the return of plasma glucose to normal levels is considered sufficient evidence that the event was induced by low plasma glucose concentration. Sub-categories of severe hypoglycemia of primary interest are:

- **Severe hypoglycemia (BG level ≤ 70 mg/dL[3.9 mmol/L]):** severe hypoglycemia with accompanying BG level ≤70 mg/dL (3.9 mmol/L)

- **Severe hypoglycemia (not biochemically confirmed—BG missing):** severe hypoglycemia without an accompanying BG level (no BG level obtained prior to administration of treatment or at the time of neuroglycopenic symptoms)

- **Severe hypoglycemia (not biochemically confirmed—BG not aligned with severe symptoms):** severe hypoglycemia in which the event was reported/confirmed by the investigator as a severe hypoglycemic event associated with neuroglycopenic symptoms or cognitive impairment that required assistance for active administration of glucose or glucagon for recovery, but a BG level >70 mg/dL was reported on the case report form that cannot be resolved through queries.

Nocturnal hypoglycemia: any total hypoglycemic event that occurs after bedtime and prior to the first meal upon waking (eg, breakfast).

Nonnocturnal hypoglycemia: any event that occurs between waking and bedtime.

Documented symptomatic hypoglycemia: an event during which typical symptoms of hypoglycemia are accompanied by a measured plasma glucose concentration ≤70 mg/dL (≤3.9 mmol/L).

Asymptomatic hypoglycemia: an event not accompanied by typical symptoms of hypoglycemia but with a measured plasma glucose concentration ≤70 mg/dL (≤3.9 mmol/L).

Probable symptomatic hypoglycemia: an event during which symptoms typical of hypoglycemia are not accompanied by a plasma glucose determination but that was presumably caused by a plasma glucose concentration ≤70 mg/dL (≤3.9 mmol/L).

Pseudo-hypoglycemia: an event during which the person with diabetes reports any of the typical symptoms of hypoglycemia with a measured plasma glucose concentration >70 mg/dL (>3.9 mmol/L) but approaching that level.

Besides the hypoglycemic categories listed above, total hypoglycemic event will be defined as follows:

Total hypoglycemic event: any event which is either documented symptomatic hypoglycemia, asymptomatic hypoglycemia, probable symptomatic hypoglycemia or severe hypoglycemia.

For each category of hypoglycemia event, the incidence, the number of hypoglycemic events per patient, the rate of hypoglycemic events per year (that is, the number of hypoglycemic episodes per patient per patient year [365.25 days]) will be calculated. These measures will be summarized at baseline, titration, maintenance, and overall study periods and at endpoint.

The proportion of patients with at least 1 hypoglycemic event (total, severe, nocturnal, and others) or incidence during the study will be summarized (counts and percentages) and analyzed using Fisher's exact test or the Pearson's chi-square test for the FAS population.

Logistic regression will be used as a sensitivity analysis of the incidence of hypoglycemia. The model will have presence or absence of hypoglycemia as the dependent or response variable and treatment, baseline HbA$_{1c}$, pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) as the independent terms in the model.

The rate of hypoglycemic episodes per year (total, severe, nocturnal, and others) will be analyzed at baseline, titration, maintenance, and overall study periods and at endpoint using the Wilcoxon test. In addition, the hypoglycemia rates will also be analyzed using a negative binomial model for the FAS population with terms for treatment, baseline HbA$_{1c}$, pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither).

In addition, the total number of patients with at least 1 hypoglycemic episode divided by the total extent of exposure in patient-years will be calculated for the overall study period and summarized descriptively for each treatment group for total, severe, nocturnal, documented symptomatic, and asymptomatic hypoglycemia definitions only. Individual patient listing of hypoglycemic events by visit will be presented for the FAS population.

All the above analysis will be repeated with the threshold of **Clinically significant hypoglycemia**: <54 mg/dL (3.0 mmol/L).

5.2.11.3. Laboratory Measures – Chemistry/Hematology Panel

Continuous measures and their change from baseline to 24-week endpoint in the chemistry and hematology panels will be summarized using descriptive statistics at baseline and at 24-week endpoint for the FAS population.

In addition, continuous measures and their change from baseline to 24-week endpoint will be analyzed using the ANCOVA model with treatment, pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) as fixed effects and the baseline value of HbA$_{1c}$ level and the baseline of the response variable as covariates. The LSMeans will be estimated using the observed margins (OM) option in SAS GLM procedure.

For each chemistry and hematology analyte, the number and percent of patients with treatment emergent high (within normal range at baseline, greater than the upper limit at Week 24/Early Discontinuation) and treatment emergent low (within normal range at baseline, less than the lower limit at Week 24/Early Discontinuation) will be presented and analyzed using Fisher`s exact test or Pearson's chi-square test.

5.2.11.4. Laboratory Measures – Insulin Antibodies

At each visit (baseline, visits 4, 8, 11, and 24-week (last observation carried forward [LOCF])), total insulin antibody status (detected, not detected (ND), no test (NT)) will be collected. The lower limit of detection of total insulin antibodies is 0.26%. If antibodies are detected then the level of total percent binding will be recorded and cross reactive antibody status (detected, not detected (ND), no test (NT)) will be recorded. The lower limit of detection of cross-reactive antibodies is 1.06%. If cross-reactive antibodies are detected then the level of cross-reactive percent binding will be recorded.

Descriptive and inferential analyses will be performed for the set of FAS patients with a valid antibody testing (detected or ND) at baseline and at least one post baseline visit.

The proportion of patients with detected insulin antibodies will be summarized as counts and percentages at baseline, at each visit, at the 24-week endpoint (LOCF), and overall for the 24-week treatment period. At each of these time points, the proportion of patients with detected antibodies will be compared between treatment groups using Fisher's exact test. These analyses will be performed for patients with antibody levels detectable at baseline and for patients without detectable levels at baseline) in the FAS population.

In addition, the following listing will be provided for patients with detectable insulin antibodies at any time during the study: a listing of level of total insulin antibody percent binding across visits sorted by treatment and maximum postbaseline percent binding in descending order. Similar listings sorted by baseline percent binding as well as by endpoint percent binding will be provided. These listings will be provided for the FAS population.

The level of total insulin antibodies (expressed as percent binding) will be summarized by descriptive statistics (mean, median, SD, Q1, inter-quartile range (IQR), Q3, standard error, minimum, and maximum) at baseline, each visit, and endpoint (LOCF). At each of these time points, the median level of percent binding will be compared between treatment groups using the Wilcoxon rank sum test. Graphical displays of median percent binding and IQR at each visit and endpoint (LOCF) by treatment will be presented. These analyses will be performed for the FAS population, for patients with antibody levels detectable at baseline and without detectable levels at baseline.

All of the analyses above will be performed for the level of cross-reactive antibodies as well.

Categorical shift tables for total insulin antibodies with 4 categories (nondetectable, [0.26, 1.26), [1.26, 5), ≥5.0) of the level of insulin antibodies from baseline to endpoint (LOCF) and from baseline to maximum postbaseline value will also be presented.

Categorical shift tables for cross-reactive antibodies with 4 categories (nondetectable, [1.06, 2.06), [2.06, 5), ≥5.0) of the level of cross-reactive antibodies from baseline to endpoint (LOCF) and from baseline to maximum postbaseline value will also be presented.

The number and proportion of patients who have a treatment-emergent antibody response (TEAR) will be summarized by treatment at each post-baseline visit, at the 24-week endpoint (LOCF), and overall for the 24-week treatment period, then analyzed using Fisher`s exact test or Pearson's chi-square test. Treatment-emergent AR is defined as an absolute increase of at least 1% in insulin antibody levels (measured in % binding) AND at least 30% relative increase from baseline [for patients who are insulin antibody-positive at baseline], or turning from insulin antibody-negative status at baseline to insulin antibody-positive during the course of the study with at least 1.26% binding (1% above the lower limit of detection) following treatment with study drug. The TEAR definition for patients who were antibody-negative at baseline takes into account the assay variability near the cut point of the assay (detectable antibody if ≥0.26%).

For patients with TEAR, a visit-wise listing of total and cross-reactive insulin antibody percent binding and clinical outcomes (HbA$_{1c}$, total hypoglycemia rate, basal insulin dose (U/Day, U/kg/Day), TEAE preferred terms) will be presented. The listing will be sorted by treatment and endpoint total insulin percent binding.

Incidence tables showing the number and percent of patients with initial and continuing TEAR at each visit will be presented. Similar analyses will be performed for the detection of total insulin antibodies and for the detection of cross reactive insulin antibodies.

Box plots of clinical outcomes (HbA$_{1c}$, total hypoglycemia rate, basal dose) will be generated by treatment and TEAR status.

In addition, the following patient listings of insulin antibodies (percent binding) will be provided for patients with TEAR: a listing of level of total insulin antibody percent binding across visit sorted by treatment and maximum postbaseline percent binding; a listing of level of total insulin antibody percent binding across visit sorted by treatment and endpoint percent binding. Similar listings will be presented for cross reactive percent binding.

In addition, the potential impact of total insulin antibody level and TEAR status on clinical response will be evaluated in several ways:

1. The relationship between the natural logarithm of the last observed insulin antibody levels (Ln[antibody level]) and selected clinical response variables (efficacy and safety measures (eg HbA$_{1c}$, total hypoglycemia rate, basal dose [U/day, U/kg/day], total dose [U/day, U/kg/day], premeal dose [U/day, U/kg/day]) will be evaluated using scatterplots and analyzed using ANCOVA on efficacy/safety measure as the dependent variable with treatment, pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) as fixed effects; baseline HbA$_{1c}$ level, Ln(antibody level) and treatment-by-Ln(antibody level) interaction as covariates for the FAS patients with detectable antibodies postbaseline and with valid antibody testing at baseline (detected, ND). A significant treatment-by-insulin antibody interaction may be indicative of a differential treatment effect necessitating further exploration to determine the nature of the interaction. Only patients with non-missing baseline value and at least one non-missing post-baseline value of the clinical response variable will be included in the analysis.

2. The relationship between TEAR and selected clinical response variables (efficacy and safety measures (e.g. HbA$_{1c}$, total hypoglycemia rate, basal dose [U/day, U/kg/day], total dose [U/day, U/kg/day], premeal dose [U/day, U/kg/day]) will be analyzed using ANCOVA on efficacy/safety measure as the dependent variable with treatment, TEAR (yes/no), pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither) , treatment-by-TEAR interaction as fixed effects and baseline HbA$_{1c}$ level as a covariate for the FAS patients with detectable antibodies postbaseline and with valid antibody testing at baseline (detected, ND). A significant treatment-by-TEAR interaction may be indicative of a differential treatment effect necessitating further exploration to determine the nature of the interaction. Only patients with non-missing

baseline value and at least one non-missing post-baseline value of the clinical response variable will be included in the analysis.

3. The relationship between TEAR and overall incidence of categories of adverse events (TEAE, SAE, TEAE related to study drug, special topic allergic reactions, injection site reactions) will be assessed by showing the proportion of patients with an event for TEAR and non-TEAR patients by treatments for the FAS. For patients with and without TEAR, treatments will be compared using the Mantel-Haenszel test and the odds ratio and p-value from this test will be reported. The homogeneity of the odds ratios for TEAR and non-TEAR patients will be assessed using the Breslow-Day test.

The potential impact of cross-reactive antibody formation on clinical response similar to the analyses above will be done if there is sufficient number of patients with cross-reactive antibodies.

5.2.11.5. Vital Signs

Vital signs measures (systolic blood pressure, diastolic blood pressure, and heart rate) and their change from baseline will be summarized by descriptive statistics (mean, median, SD, standard error, minimum, and maximum) by visit for the FAS population.

In addition, the MMRM model will evaluate the change from baseline in vital sign measure as the dependent variable with treatment (LY2963016, Lantus®), pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), and pre-study metformin or acarbose usage (metformin only, acarbose only, or neither), visit, and interaction between visit and treatment as fixed effects; the baseline value of the vital sign measure and the baseline HbA_{1c} level as covariates; and a random effect for patient. The LSMeans will be estimated using the observed margins (OM) option in SAS MIXED procedure.

The number and percentage of patients with treatment-emergent outlier vital sign measures at each post baseline visit will be presented and analyzed using Fisher`s exact test or Pearson's chi-square test for systolic blood pressure, diastolic blood pressure, and pulse rate according to the categories as listed below:

- low systolic blood pressure ($\leq$90 mmHg and a decrease from baseline $\geq$20 mmHg)

- high systolic blood pressure ($\geq$140 mmHg and an increase from baseline $\geq$20 mmHg)

- low diastolic blood pressure ($\leq$50 mmHg and a decrease from baseline $\geq$10 mmHg)

- high diastolic blood pressure ($\geq$90 mmHg and an increase from baseline $\geq$10 mmHg)

- low heart rate (<50 bpm and a decrease from baseline $\geq$15 bpm)

- high heart rate (>100 bpm and an increase from baseline $\geq$15 bpm)

5.2.11.6. Other Safety Measures

Exposure:

Exposure to each treatment during the treatment period of the study will be calculated for each patient and summarized by treatment group. Exposure will be calculated as the number of days

from the date of first dose of the study drug (or if this information is missing, from the date of randomization) to the date of last treatment dose. Exposure will be expressed in days, months (30 days) and years (365.25 days).

Listings of Hepatic Disorders and Abnormal Liver Enzymes:
A listing will be provided for patients with pre-existing conditions or treatment emergent adverse events in the following MedDRA SMQs for Hepatic Disorders (Broad and Narrow SMQ):

Narrow	Biliary disorders	20000118
Narrow	Drug-related hepatic disorders – comprehensive	20000006

Additionally listings will be provided for all patients who meet at least one of the following liver enzyme outlier criteria for at least one visit: ALT >3 ULN, AST >3 ULN, total bilirubin >2 ULN, or alkaline phosphatase >2 ULN at any visit. These listings will include all liver enzymes at all visits.

Listings Based on Renal Function:
Listings will be provided for patients with severe reduction in GFR or kidney failure at any visit according to the following criteria:

Severe reduction in eGFR (15-29 mL/min/1.73 m^2)
Kidney Failure (eGFR< 15 mL/min/1.73 m^2).

These listings will include all estimated GFR (eGFR) at all visits.

5.2.12. Subgroup Analyses
The consistency of the treatment effect for the change in HbA1c level will be assessed in the following subgroups in the FAS population if there are sufficient numbers of patients in each treatment by subgroup (e.g. 20 patients per cell):

- Entry HbA_{1c} levels (<7%, ≥7%)
- Entry HbA_{1c} levels (<8.5%, ≥8.5%)
- Pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin)
- Entry BMI (<30, ≥30)
- Entry BMI (<25, ≥25)
- Entry age (<65, ≥65)
- Entry age (<75, ≥75)
- Gender
- Prestudy metformin or acarbose use (metformin only, acarbose only or neither)
- Time of basal insulin injection (daytime, evening/bedtime)
- Renal function, as estimated by estimated glomerular filtration rate (EGFR) using the MDRD formula (see Appendix 2). The following EGFR categories will be used:

- o normal or increased GFR: EGFR ($\geq$90 mL/min/1.73 m^2)
- o mild reduction in GFR: EGFR (60 to 89 mL/min/1.73 m^2)
- o Moderate reduction in GFR: EGFR (30 to 59 mL/min/1.73 m^2)
- o Severe reduction in GFR: EGFR (15 to 29 mL/min/1.73 m^2)
- o Kidney failure: EGFR (<15 mL/min/1.73 m^2).

The categories severe reduction in GFR and kidney failure will be combined if there are fewer than 20 patients in either of these categories.

The change in HbA1c from baseline to 24-week endpoint will be analyzed using MMRM with treatment, visit, prestudy treatment, prestudy metformin or acarbose usage, subgroup, subgroup-by-treatment interaction, subgroup-by-visit interaction, treatment-by-visit interaction, and treatment-by-visit-by-subgroup interaction as fixed-effects, the baseline value of HbA1c as a covariate, and a random effect for patient for the FAS population. If the subgroup is one of the stratification variables, then the subgroup will only be included once in the model. A significant treatment-by-subgroup interaction ($p<.05$) may be indicative of a differential treatment effect across levels of the subgroup, necessitating further exploration of the nature of the interaction.

The contrast from the MMRM model with the OM option using PROC Mixed in SAS above will be used to obtain the LSmeans for change in HbA$_{1c}$ levels from baseline to Week 24 within each treatment at each level of the subgroup. Additionally, the difference of these LSmeans between treatment groups with 95% confidence interval and p-values will be presented.

For each subgroup listed above, the change in weight from baseline to 24-week will be analyzed using the same methodology as the subgroup analysis of HbA$_{1c}$.

Additional subgroup analyses will be carried out on selected safety outcomes:

a) total, severe or nocturnal hypoglycemia at 24-week endpoint and for the overall 24-week period (rates and incidence)
b) treatment-emergent antibody response (TEAR) at 24-week endpoint and for the overall 24-week period
c) detectable insulin antibodies at 24-week endpoint and for the overall 24-week period
d) categories of adverse events:
   - o treatment-emergent adverse events (TEAEs) (overall incidence and preferred terms with >5% incidence)
   - o TEAEs related to study drug (overall incidence)
   - o special topic assessment of allergic events (overall incidence)
   - o injection site reactions (overall incidence)
   - o serious adverse events (overall incidence)

For rates and incidence of hypoglycemia (total, nocturnal, severe) and categories of adverse events, the same set of 11 subgroups used to analyze change in HbA1c and change in weight will be analyzed.

For each categorical safety outcome and subgroup listed above, the proportion of patients with an event will be compared between treatments for the FAS. Within each subgroup, treatments will be compared using the Mantel-Haenszel test and the odds ratio and p-value from this test will be reported. The homogeneity of the odds ratios across subgroups will be assessed using the Breslow-Day test.

For the 11 subgroups listed above, the rate of total, severe or nocturnal hypoglycemia per year at the 24-week endpoint and for the overall 24-week period will be analyzed. Within each subgroup, treatments will be compared using the negative binomial (NB) model with treatment as the only factor. The NB mean for each treatment and their ratio with corresponding 95% confidence interval will be presented. Additionally, within each subgroup the rates will be compared using the Wilcoxon test. The significance of the subgroup-by-treatment interaction will be evaluated using a negative binomial model with factors treatment, subgroup and treatment-by-subgroup interaction. The negative binomial model will be implemented using the SAS PROC GENMOD with natural logarithm of exposure time (in days) as an offset variable. The above subgroup analysis will be performed and p-values will be provided if there are at least a total of 20 events in the combined treatment groups.

For TEAR at 24 weeks endpoint and overall, subgroup analyses will be performed only for pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin), entry age (>65, ≤65), entry age (<75, ≥75) and renal function. For detectable antibodies at 24 weeks endpoint and overall, subgroup analyses will be performed only for pre-study treatment (Lantus®-bolus, other basal-bolus, premixed insulin).

Other subgroup analyses may be performed if deemed appropriate as exploratory analyses.

5.2.13. Interim Analyses

A single interim analysis on both efficacy and safety data may be performed to allow interaction with regulatory authorities. Analyses on both interim and final data will be performed using the same statistical methods, as described in this protocol and the SAP. The study will not stop for early efficacy at the interim analysis therefore no adjustment of Type I error is needed. If there are significant safety concerns arising from the interim analyses, the study team may decide to stop the trial early for safety concerns. The study team will decide, based on the trial operation and blinded safety information, whether and when the interim analyses will be performed.

If any unplanned interim analysis is deemed necessary, the appropriate Lilly regulatory scientist will be consulted to determine whether it is necessary to amend the protocol.

5.2.14. Analysis of Visit 801 Data

Data from the 4 week post-treatment follow up period (from Visit 13 to Visit 801) will be

appended in the listings. Treatment-emergent adverse events (TEAEs) at post-treatment followup

visit (V801) will be summarized for the FAS. Treatment-emergent adverse events at V801 are

defined as events that are newly reported or have worsened in severity after Visit 16 (or from the

last visit noted).

TEAEs of special topic assessment of adverse (allergic) events at V801 will also be summarized in a similar manner.

5.2.15. Exploratory Analysis
Additional exploratory analyses will be conducted when deemed appropriate.

5.2.16. Required Analyses for the Clinical Trial Registry (CTR)
The study team will create the CTR adverse event dataset based on the CTRAESUMM ADaM standard. A member of the CTR team will be responsible for generating the standard CTR reports from the CTR adverse event dataset and for uploading the reports to the CTR site. The CTR adverse event dataset will include the following requirements:

- Both Serious Adverse Events and 'Other' Adverse Events will be summarized and analyzed:

    o An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE)
    o An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious

- Serious Adverse Events and 'Other' Adverse Events will be summarized: by treatment group and by MedDRA preferred term.

- For each Serious AE and 'Other' AE term, the following will be provided for each treatment group:

    o the number of participants at risk of an event
    o the number of participants who experienced each event team
    o the number of events experienced

- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients in every treatment group may not be included if a 5% threshold is chosen (5% is the maximum threshold).

- AE reporting will be consistent with other document disclosures for example, the CSR, manuscripts, and so forth. A member of the CTR team will perform the quality checks to ensure that the AE reporting for the CTR is consistent.

6. Unblinding Plan

This is an open-label study in which investigators, patients, study site personnel, and study monitors will be aware of the treatment assignment. To minimize bias, review of summary data by the Lilly study team (i.e. CRP/CRS overseeing the global conduct of the study, statisticians, and statistical analysts) prior to the final database lock of the study (at the end of 24 weeks of treatment) will remain blinded to treatment assignment. Unblinding of an individual patient's study drug treatment assignment may occur in the course of consultation between the investigator and the study team (principally between the investigator and the CRP/CRS) or during review of SAEs. No systematic unblinding of study drug treatment assignments will be performed by the Lilly study team before the final database lock. Similar to a double-blind study, a minimum number of Lilly personnel will have access to the randomization table and treatment assignments before the final database lock.

7. References

Fu H, Manner D. Bayesian Adaptive Dose-Finding Studies with Delayed Responses. *Journal of Biopharmaceutical Statistics.* 2010: 20:5, 1055-1070.

[White Paper] Standardization of Derivations for Efficacy Lab Variables. Lilly Unpublished Manuscript.

[White Paper] White Paper on the Classification and Data Capture of Hypoglycemia Events. Lilly Unpublished Manuscript.

8. Appendices

Appendix 1. Listing of Allergic Reaction Terms

Allergic bronchitis	Hypersensitivity	Pruritus generalised
Allergic colitis	Idiopathic urticaria	Rash
Allergic cough	Immediate post-injection reaction	Rash erythematous
Allergic cystitis	Injection site dermatitis	Rash follicular
Allergic keratitis	Injection site eczema	Rash generalised
Allergic oedema	Injection site erythema	Rash macular
Allergic otitis media	Injection site hypersensitivity	Rash maculo-papular
Allergic pharyngitis	Injection site induration	Rash maculovesicular
Allergic respiratory symptom	Injection site inflammation	Rash papular
Alveolitis allergic	Injection site macule	Rash pruritic
Anaphylactic reaction	Injection site nodule	Rash pustular
Anaphylactic shock	Injection site oedema	Rash vesicular
Anaphylactoid reaction	Injection site papule	Reaction to drug excipients
Anaphylactoid shock	Injection site photosensitivity reaction	Reaction to preservatives
Angioedema	Injection site pruritus	Reversible airways obstruction
Arthralgia	Injection site pustule	Scleral oedema
Arthritis	Injection site rash	Scleritis allergic
Arthritis allergic	Injection site reaction	Skin oedema
Asthma	Injection site recall reaction	Small bowel angioedema
Auricular swelling	Injection site streaking	Stevens-Johnson syndrome
Bronchial hyperreactivity	Injection site swelling	Stridor
Bronchial oedema	Injection site urticaria	Suffocation feeling
Bronchospasm	Injection site vesicles	Swelling face
Circumoral oedema	Joint effusion	Swollen tongue
Conjunctival oedema	Joint swelling	Throat tightness
Corneal oedema	Laryngeal obstruction	Tongue oedema
Dermatitis	Laryngeal oedema	Toxic epidermal necrolysis
Dermatitis allergic	Laryngitis allergic	Toxic skin eruption
Dermographism	Laryngotracheal oedema	Tracheal obstruction
Diffuse cutaneous mastocytosis	Lip oedema	Tracheal oedema
Drug eruption	Lip swelling	Type I hypersensitivity
Drug hypersensitivity	Local swelling	Type II hypersensitivity
Drug rash with eosinophilia and systemic symptoms	Localised oedema	Type III immune complex mediated reaction

Encephalopathy allergic

Eosinophilic oesophagitis

Epiglottic oedema

Erythema multiforme

Erythema nodosum

Eye oedema

Eye swelling

Eyelid oedema

Face oedema

Gastrointestinal oedema

Gingival oedema

Gingival swelling

Haemorrhagic urticaria

Nasal oedema

Nephritis allergic

Oculorespiratory syndrome

Oedema mouth

Oedema mucosal

Oesophageal oedema

Orbital oedema

Oropharyngeal swelling

Palatal oedema

Periarthritis

Periorbital oedema

Pharyngeal oedema

Photosensitivity allergic reaction

Photosensitivity reaction

Pruritus

Pruritus allergic

Type IV hypersensitivity reaction

Urticaria

Urticaria cholinergic

Urticaria chronic

Urticaria contact

Urticaria popular

Urticaria physical

Urticaria pigmentosa

Urticaria pressure

Urticaria thermal

Urticaria vesiculosa

Urticaria vibratory

Visceral oedema

Wheezing

Appendix 2.   Estimation of Creatinine Clearance Using the Cockroft-Gault and MDRD for the Approximation of Glomerular Filtration Rate (GFR)

The most common equations used in the United States are the Cockcroft-Gault and Modification of Diet in Renal Disease (MDRD) study equations. The IDMS MDRD study equation is increasingly utilized in the United States (Levey et al. 2006)

Cockcroft-Gault equation — The Cockcroft-Gault equation allows the creatinine clearance to be estimated from the serum creatinine in a patient with a stable serum creatinine:

$$\text{Male eCrCl (mL/min)} = (140 - \text{age[years]}) \times (\text{weight[kg]}) / (\text{sCr} \times 72)$$

$$\text{Female eCrCl (mL/min)} = (140 - \text{age[years]}) \times (\text{weight[kg]}) \times 0.85 / (\text{sCr} \times 72)$$

where sCr is the serum Creatinine in mg/dl.

This formula takes into account the increase in creatinine production with increasing weight, and the decline in creatinine production with age. For women, the formula requires multiplication by 0.85 to account for smaller muscle mass compared to men.

Modification of Diet in Renal Disease (MDRD) Equation

Estimate the patient's creatinine clearance using Modification of Diet in Renal Disease (MDRD) equation as an approximation of glomerular filtration rate (GFR). The MDRD equation is:

$$\text{eGFR} = 175 \times \text{standardized sCr}^{-1.154} \times \text{age}^{-0.203} \times 1.212 \text{ [if black]} \times 0.742 \text{ [if female]}$$

Renal function, as estimated by estimated glomerular filtration rate (eGFR) using the MDRD formula will be used to calculate the following eGFR categories (Stage 1 to Stage 5):

Stage 1: Normal or increased eGFR (>90 mL/min/1.73 m^2)
Stage 2: Mild reduction in eGFR (60-89 mL/min/1.73 m^2)
Stage 3: Moderate reduction in eGFR (30-59 mL/min/1.73 m^2)
Stage 4: Severe reduction in eGFR (15-29 mL/min/1.73 m^2)
Stage 5: Kidney Failure (eGFR <15 mL/min/1.73 m^2)

References:

Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron*. 1976;16(1):31-41.

Levey AS, Coresh J, Greene T, Stevens LA, Zhang Y (Lucy), Hendriksen S, Kusek JW, Van Lente F, for the Chronic Kidney Disease Epidemiology Collaboration. Using standardized serum creatinine values in the Modification of Diet in Renal Disease Study equation for estimating glomerular filtration rate. *Ann Intern Med*. 2006;145:247-254.

Leo Document ID = e3a3e138-da12-4a4e-8152-31c386b0a56f

Approver: PPD
Approval Date & Time: 12-Mar-2018 08:22:03 GMT
Signature meaning: Approved