Statistical Analysis Plan I8H-MC-BDCV (Version 3)

A Phase 3, Parallel-Design, Open-Label, Randomized Controlled Study to Evaluate the Efficacy and Safety of LY3209590 as a Weekly Basal Insulin Compared to Insulin Glargine in Adults With Type 2 Diabetes on Multiple Daily Injections

NCT05462756

Approval Date: 18-MAR-2024

Title Page

Protocol Title: A Phase 3, Parallel-Design, Open-Label, Randomized Controlled Study to Evaluate the Efficacy and Safety of LY3209590 as a Weekly Basal Insulin Compared to Insulin Glargine in Adults with Type 2 Diabetes on Multiple Daily Injections

Protocol Number: I8H-MC-BDCV

Compound Number: LY3209590
Short Title: Efficacy and Safety of LY3209590 Compared with Daily Insulin Glargine in Adults with Type 2 Diabetes on Multiple Daily Injections

Acronym: QWINT-4

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

IND: 129390

EudraCT: 2021-005878-25

Document ID: VV-CLIN-056525

Table of Contents

Version History

This Statistical Analysis Plan (SAP) for study I8H-MC-BDCV is the first version and approved prior to the first participant visit. This version is based on the protocol and protocol amendments approved on 31 March 2022, 05 April 2022, and 10 May 2022, respectively.

Table 1 SAP Version History Summary

Change	Description	Rationale
Version 2, Jan 22, 2024	1. Edited the estimand for the composite measure of achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L], or severe) in Section 1.1. 2. Edited the analysis method for Time in glucose range between 70 and 180 mg/dL in Section 4.4.2. 3. Added sensitivity analyses in Section 4.3.3.	To address FDA's feedback
	1. Added DID-EQ to secondary objectives in Section 1.1 and statistical analysis method in Section 4.4.2. 2. Modified statistical method for PRO outcomes form MMRM to ANCOVA based on 1 post-baseline measure.	Per protocol amendment (b) to address regulatory feedback
	1. For laboratory analysis, removed shift analysis, treatment emergent high/low and add elevated or low values meeting specified levels. 2. Risk difference and 95% CI will be used for safety categorical data analysis.	Per Lilly PSAP update in Version3
	1. Added analysis for CGM-based hypoglycemia in Section 4.6.3.1.2. 2. Updated the definition for CGM-based hypoglycemia and other updates in Appendix 6.6	Per CGM international consensus statement 2023
	1. Added graphical testing figure and description in Section 2.1. 2. Added more subgroups in Section 4.7.2.	As planned.
	1. Updated data cutoff for glargine arm in EAS2 in Section 3.	According to PK profile
	1. Simplified immunogenicity section by referring to the PSAP in Section 4.7.1.	To simplify the SAP
	Minor changes and reorganizations in the entire document.	For clarity.
Version 3, Mar 15, 2024	1. Added treatment regimen analysis method for fasting serum glucose, fasting glucose from SBMG, and CGM parameters for supportive secondary endpoints and tertiary endpoints in Section 4.4.2 and Section 4.5.1 2. Modified missing data handling approach for nocturnal hypoglycemia for the composite endpoint of HbA1c<7% at Week 26 without nocturnal hypoglycemia in Section 4.4.1. 3. Added sensitivity analyses for CGM that includes all data in derived variables in Appendix 6.	According to FDA's feedback

	4. Added a sensitivity analysis that considers all hypoglycemic events as one hypoglycemic event until a succeeding glucose value is ≥ 70 mg/dL in Section 4.6.2.2.1. 5. Updated the subgroup for race by each category.	
	Added information for Device Product Complaints in Section 4.6.3.	According to the new template

1. Introduction

The study protocol contains a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints. The changes to the analyses described in the protocol and in the prior version of SAP are listed in the Version History section. This SAP includes the analyses details for efficacy, safety measures, patient-reported outcomes, and parameters based on continuous glucose-monitoring (CGM). Pharmacokinetic/pharmacodynamics (PK/PD) analyses will be conducted by the PK/PD group and will be described in the PK analysis plan. The specifications for ADaM data, tables, figures and listings will be described in separate documents.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
Demonstrate noninferiority of LY3209590 compared to insulin glargine on glycemic control.	Change in HbA1c from baseline at Week 26
Key Secondary (Multiplicity Adjusted)	
Demonstrate LY3209590 is superior to insulin glargine in the selected parameters of glycemic control.	• Change in HbA1c from baseline at Week 26 • The percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26 • The event rate of participant-reported clinically significant nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26
Other Secondary	
To investigate the effect of LY3209590 compared with insulin glargine in additional parameters of diabetes management.	• Change in fasting glucose measured by SMBG from baseline to Week 26 • Time in glucose range between 70 and 180 mg/dL (3.9 and 10.0 mmol/L), inclusive measured during the CGM session that occurs between Week 22 and Week 26 • Time in hypoglycemia range with glucose <54 mg/dL (3.0 mmol/L), measured by CGM that occurs between Week 22 and Week 26 • Time in hyperglycemia range with glucose >180 mg/dL (10.0 mmol/L), measured by CGM that occurs between Week 22 and Week 26 • Glucose variability measured during the CGM session that occurs between Week 22 and Week 26 • Insulin dose at Week 26 o basal o bolus o total, and o basal/total insulin dose ratio
To investigate the safety of LY3209590 compared with insulin glargine.	• Incidence and rate of composite of Level 2 and 3 hypoglycemia events during treatment period • Body weight change from baseline to Week 26
To investigate the treatment impact of LY3209590 compared with insulin glargine on patient-reported outcome measures.	• Treatment experience at Weeks 26 for DID-EQ
Tertiary	
• Investigate the treatment impact of LY3209590 compared with insulin glargine on other measures of efficacy, safety, and patient-reported outcomes.	Efficacy • Percentage of participants achieving HbA1c <7% at Week 26 • Percentage of participants achieving HbA1c ≤6.5% at Week 26 • Change from baseline to Week 26 in fasting serum glucose as measured by central laboratory

Objectives	Endpoints
	Safety • Rate and incidence of Level 2 hypoglycemia events during treatment period • Rate and incidence of Level 3 hypoglycemia events during treatment period • Incidence of positive treatment-emergent antibody of LY3209590
	Patient-reported outcomes • Frequency of responses to "Basal Insulin Experience: Likelihood of incorporating into routine" at Week 26 • Frequency of responses to "Basal Insulin Experience: Preference" at Week 26 • Change in EQ-5D-5L from baseline at Week 26
• To characterize the PK/PD of LY3209590.	• LY3209590 PK and concentration response relationships to key safety and efficacy measures

Abbreviations: CGM = continuous glucose monitoring; DID-EQ = Diabetes Injection Device Experience Questionarire; EQ-5D-5L = EuroQol 5 dimensions 5 levels; HbA1c = hemoglobin A1c; PK = pharmacokinetics; SMBG = self-monitoring of blood glucose.

Primary estimand (for primary objective)

United States registration

The *primary* clinical question of interest is:

> What is the treatment difference between LY3209590 and insulin glargine in change from baseline at Week 26 in hemoglobin A1c (HbA1c) in study eligible participants regardless of treatment discontinuation for any reason or initiation of rescue therapy?

The ***treatment regimen estimand*** will be used for the primary objective and is described by the attributes in this table.

Treatment Regimen Estimand Attribute	Description
Treatment condition	Randomized treatment regardless of treatment discontinuation or initiation of rescue therapy
Population	Targeted study population. See Section 3 for detail.
Endpoint	Change from baseline to Week 26 in HbA1c
Remaining intercurrent events	None. The 2 intercurrent events, treatment discontinuation for any reason and initiation of rescue therapy, are addressed by the treatment condition of interest attribute.
Population-level summary	Difference in mean changes between treatment conditions

Abbreviation: HbA1c = hemoglobin A1c.

Rationale for treatment regimen estimand:
The treatment regimen estimand estimates how participants with type 2 diabetes (T2D) are treated in clinical practice and takes into account both efficacy and safety.

Registration for countries outside the United States

The *primary* clinical question of interest is:

> What is the treatment difference between LY3209590 and insulin glargine in change from baseline at Week 26 in HbA1c in study eligible participants who adhere to the randomized treatment without an intercurrent event during the study treatment period?

The ***efficacy estimand*** will be used for the primary objective and is described by the attributes in this table.

Efficacy Estimand Attribute	Description
Treatment condition	Randomized treatment
Population	Targeted study population. See Section 3 for detail.
Endpoint	Change from baseline to Week 26 in HbA1c
Remaining intercurrent events	None. The 2 intercurrent events, treatment discontinuation for any reason and initiation of rescue therapy, are handled by the hypothetical strategy, for example, the potential outcome for those participants if the intercurrent events have not occurred will be estimated
Population-level summary	Difference in mean changes between treatment conditions

Abbreviation: HbA1c = hemoglobin A1c.

Rationale for efficacy estimand: The treatment efficacy estimand supports the interpretation of the treatment effect as participants adhere to study treatment.

Secondary estimands (for multiplicity-adjusted objectives)

The superiority test in change from baseline to Week 26 in HbA1c will also be based on the primary estimands described above.

The percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L], or severe) during treatment phase up to Week 26 will be based on a composite estimand. Section 4.4.1.2 contains additional details.

The participant-reported clinically significant nocturnal hypoglycemia (<54 mg/dL[3.0 mmol/L], or severe) is 1 of the safety measures for the study. The event rate will be based on all available data during the specific analysis period. Relative rate between randomized treatment groups will be used for treatment comparison.

1.2. Study Design

- Study BDCV is a Phase 3, multicenter, randomized, open-label, comparator-controlled study to evaluate the efficacy and safety of once-weekly basal insulin (LY3209590) compared to insulin glargine in participants with T2D who are on once- or twice-daily basal insulin and at least 2 injections per day of prandial insulin prior to entering the study.

- The study includes a 3-week screening/lead-in period, a 26-week treatment period, and a 5-week safety follow-up period.
- Participants will continue prior stable diabetes therapy with 0 to 3 allowed non-insulin antihyperglycemic medications during the study.
- Participants will discontinue their previous basal insulin doses at randomization and will be assigned to either LY3209590 or insulin glargine, and all participants will be treated with prandial insulin, insulin lispro.
- Assignment to treatment groups will be determined by an interactive web-response system (IWRS). Participants will be randomly assigned to 1 of the 2 treatment groups in 1:1 ratio (LY3209590: insulin glargine) at Week 0 (Visit 3). Stratification will be by country, HbA1c stratum at Week -3 (Visit 1) (<8.0%, ≥8.0%), and routine use of personal continuous glucose monitoring (CGM) or flash glucose monitoring (FGM) at randomization.
- This is a treat-to-target study, both treatment groups will undergo titration to achieve an FBG target within the range of 80-120 mg/dL,. The primary endpoint at Week 26 (Visit 23) is based on participants reaching a stable insulin dose in order to achieve glycemic stability approximately 12 weeks in advance of the 26-week primary endpoint. HbA1c is the primary endpoint and reflects glycemic control from the prior 8 to 12 weeks.
- Rescue therapy will be considered during the treatment period if the participants meet the protocol criteria of severe, persistent hyperglycemia.
- If study intervention is permanently discontinued, the participant will be encouraged to continue participation in the study for continued monitoring. Both efficacy (including HbA1c) and safety data will continue to be collected per the schedule of activities in the protocol.

2. Statistical Hypotheses

Primary Hypothesis

The primary objective of this study is to test the hypothesis that LY3209590 is noninferior to insulin glargine on glycemic control as measured by change from baseline to Week 26 (Visit 23) in HbA1c in participants with T2D currently on basal insulin and at least 2 injections per day of prandial insulin.

Thus, the null hypothesis (H_0) to be tested is the difference in the change from baseline to Week 26 (Visit 23) in HbA1c (LY3209590 – insulin glargine) is greater than the noninferiority margin (NIM). The upper bound of the 2-sided 95% confidence interval (CI) will be used for testing the noninferiority.

Secondary Hypotheses

The key secondary (multiplicity adjusted) objectives are to test the hypotheses that LY3209590 is superior to insulin glargine with respect to

- change from baseline in HbA1c at Week 26
  H_0: the difference (LY3209590 – insulin glargine) ≥ 0.0
- The percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26
  H_0: the odds ratio (LY3209590 vs insulin glargine) ≤ 1
- the event rate of nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment period up to Week 26
  H_0: the relative event rates (LY3209590 vs insulin glargine) ≥ 1

These hypotheses and the primary hypothesis will be tested using a strategy to control the overall type 1 error.

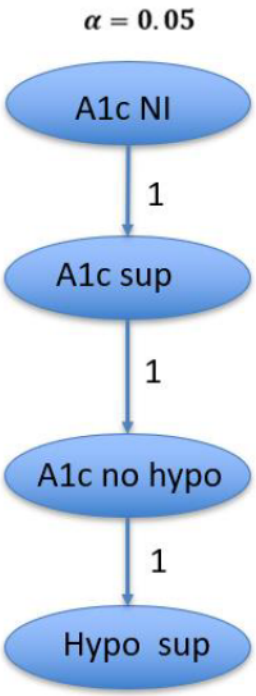
2.1. Multiplicity Adjustment

A graphical approach (Bretz et al. 2009; 2011) for multiple comparisons will be used to ensure the strong control of the overall type I error rate for testing the primary and key secondary (multiplicity adjusted) objectives.

As illustrated in Figure BDCV 2.1, the primary and key secondary endpoints will be sequentially tested using the primary statistical method described in Section 4.3 to compare LY3209590 and insulin glargine. The overall significance level (α) will be set to 0.05. The total α will be used for the primary objective first, and the noninferiority test for the primary objective will be based on the 2-sided 95% CI. Once the upper limit of the 2-sided 95% CI is below the NIM, the non-inferiority is achieved and the α of 0.05 will be transitioned to test superiority for the key secondary objectives sequentially. If a test in this sequence is not significant, all the subsequent tests will be considered non-significant (Hsu and Berger 1999). This is a predetermined fixed testing order; thus it is a gatekeeper method. The graphical approach is a closed testing procedure; hence, it strongly controls the family-wise error rate (Hsu and Berger 1999; Alosh et al. 2014).

The NIMs of 0.4% (for treatment regimen estimand) and 0.3% (for efficacy estimand) will both be tested to meet different regulatory requirements. No multiplicity adjustments will be made for

conducting separate analyses relative to the efficacy and treatment regimen estimands. Each estimand will have its own family-wise error rate of 0.05.



Primary: Change in HbA1c from baseline at Week 26 (LY3209590 noninferior to insulin glargine)

Key Secondary 1: Change in HbA1c from baseline at Week 26 (LY3209590)

Key Secondary 2: The percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26 (LY3209590 superior to insulin glargine)

Key Secondary 3: The event rate of participant-reported clinically significant nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26 (LY3209590 superior to insulin glargine)

Abbreviations: NI = noninferiority; HbA1c = hemoglobin A1c; hypo = hypoglycemia; sup = superiority.

Figure BDCV 2.1. Illustration of graphical multiple testing procedure with initial α allocation and weights.

3. Analysis Sets

This table defines the populations and data sets for the purpose of analysis.

Analysis Populations/ Data Sets	Description
Entered Population	All participants who sign the informed consent form
Randomized Population	All randomized participants. Participants will be analyzed according to the treatment they were assigned.
Modified Intent-to-Treat (mITT) Population	All randomized participants who took at least 1 dose of study treatment. Participants will be analyzed according to the treatment they were assigned.
Efficacy Analysis Set 1 (EAS1) for treatment regimen estimand on efficacy measures	The data will include • mITT Population, excluding participants discontinuing the study treatment due to inadvertent enrollment, and • all measurements regardless of the use of study treatment or the initiation of rescue therapy.
Efficacy Analysis Set 2 (EAS2) for efficacy estimand and for other secondary/tertiary objectives on efficacy measures	The data will include • mITT Population, excluding participants discontinuing the study treatment due to inadvertent enrollment, and • measurements up to the early discontinuation of study treatment or the initiation of rescue therapy. The data cutoff for participants who had intercurrent events is defined by the earliest date from below dates for individual participants other than the analysis for study dose: • the date of last study dose + 10 days for LY3209590, or +1 day for glargine • the start date of the first treatment of rescue therapy The data cutoff for the analysis on study dose is defined by the earliest date from below dates for individual participants: • the date of last study dose • the start date of the first treatment of rescue therapy
Safety Analysis Set (SS)	The data will include • mITT Population, and • all measurements regardless of the use of study treatment or the initiation of rescue therapy.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designees. 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 changes to the data analysis methods described in the protocol, and the justification for making the change, will be described in this SAP or the clinical study report. Additional exploratory analyses of data will be conducted as deemed appropriate.

Unless otherwise stated, the efficacy analyses will be based on either efficacy analysis set 1 (EAS1, see the definition in Section 3) or efficacy analysis set 2 (EAS2, see the definition in Section 3). For treatment regimen estimand, efficacy analysis set 1 (EAS1, see the definition in Section 3) will be used. For efficacy estimand, EAS2 will be used. The treatment comparison will be based on either 2-sided test or 95% CI.

Unless otherwise noted, the safety analyses will be conducted on Safety Analysis set (SS, see the definition in Section 3). Percentages will be calculated using the safety population as the denominator. For events that are sex-specific, the denominator and computation of the percentage will include only participants from the given sex.

For continuous measures, summary statistics will include sample size, mean, SD, median, minimum, and maximum for both the actual and the change from baseline measurements. For certain variables that are considered to be log-normally distributed, the geometric mean and coefficient of variation (CV) will be provided instead. Either the mixed model repeated measurement (MMRM) model or the analysis of covariance (ANCOVA) model will be used to analyze continuous outcomes. Least-squares (LS) means and standard errors derived from the analysis models will also be displayed. Treatment comparisons will be displayed showing the treatment difference LS means and the 95% CIs for the treatment differences, along with the p-values for the treatment comparisons. For certain safety laboratory measures, log-transformed values will be analyzed in the statistical model instead. The actual change from baseline and percentage change from baseline will be presented using the derivation based on the output from the statistical model and the assumption of log-normality.

For categorical measures, summary statistics will include sample size, frequency, and percentages. Fisher's exact test or logistic regression will be used for treatment comparisons, unless otherwise stated.

For laboratory values, both conventional (CN) and System International (SI) units will be presented. Therefore, both % and mmol/mol will be presented for HbA1c and both mg/dL and mmol/L will be presented for glucose measurements.

In this study, the negative binomial regression will be used to analyze the number of hypoglycemic episodes. Group Mean instead of LS mean will be estimated and delta method will be used to estimate the standard error of the Group Mean (Qu and Luo 2015). Group Mean is defined as the mean response in the treatment group for the studied population. The difference between LS mean and the Group Mean is that LS mean estimates the response by taking the inverse link function on mean covariates, while the Group Mean takes the inverse link function on individual patient covariates first and then averages over all patients. For severe

hypoglycemia, the empirical method based on exposure adjusted rate (calculated by total number of events divided by total exposure) may be used for the treatment comparison if the number of episodes is too small and leads to convergency issues in the negative binomial regression model.

The table below describes the definition of baseline and post-baseline observations for different analyses.

Analysis	Baseline Observations	Post-Baseline Observations
HbA1c (treatment regimen estimand)	The baseline is the last non-missing assessment prior to or on the day of the first dose of study treatment (CRF is used to determine).	Planned measurements at Week 26 (primary endpoint, Visit 23) in EAS1. Use unplanned measurements if no planned measurement. Multiple imputation approach will be used to impute missing observations at Week 26 (Visit 23).
HbA1c (efficacy estimand)	The baseline is the last non-missing assessment prior to or on the day of the first dose of study treatment.	Weeks 2, 4, 12, 16 and 26 (Visit 5, 7, 15, 17 and 23) in EAS2 for MMRM. Planned measurements at scheduled visits will be included. Use unplanned measurements if no planned measurement
CGM parameters	The baseline will be derived from the data collected during CGM session prior to the first dose date of study treatment.	The CGM sessions prior to Weeks 4, 12, 26 (Visit 7, 15, 23) in EAS2 for MMRM. The CGM session prior to Week 26 (Visit 23) in EAS1 for ANCOVA with multiple imputation for missing data.
Basal insulin dose	The baseline period is the screening/lead-in period prior to the initiation of study treatment. The following variables will be derived: • daily dose in U • weekly dose in U • dose in U/kg/day See detailed derivation of variables in Section 4.4.2.1.	All scheduled visits between Week 0 (Visit 3) and Week 26 (Visit 23) in EAS2 for MMRM. The following variables will be derived: • daily dose in U • weekly dose in U • dose in U/kg/day See detailed derivation of variables in Section 4.4.2.1.

Analysis	Baseline Observations	Post-Baseline Observations
Prandial (bolus) insulin dose	The baseline period is the screening/lead-in period prior to the initiation of study treatment. The following variables will be derived: • daily dose in U • weekly dose in U • dose in U/kg/day See detailed derivation of variables in Section 4.4.2.1.	All scheduled visits after Week 0 (Visit 3) and up to Week 26 (Visit 23) in EAS2 for MMRM. The following variables will be derived: • daily dose in U • weekly dose in U • dose in U/kg/day See detailed derivation of variables in Section 4.4.2.1.
Total insulin dose	The baseline period is the screening/lead-in period prior to the initiation of study treatment. The following variables will be derived: • daily dose in U • weekly dose in U • dose in U/kg/day See detailed derivation of variables in Section 4.4.2.1.	All scheduled visits after Week 0 (Visit 3) and up to Week 26 (Visit 23) in EAS2 for MMRM. The following variables will be derived: • daily dose in U • weekly dose in U • dose in U/kg/day See detailed derivation of variables in Section 4.4.2.1.
Basal/total insulin dose ratio	The baseline period is the screening/lead-in period prior to the initiation of study treatment. See detailed derivation of variables in Section 4.4.2.1.	All scheduled visits after Week 0 (Visit 3) and up to Week 26 (Visit 23) in EAS2 for MMRM. See detailed derivation of variables in Section 4.4.2.1.
Fasting glucose by SMBG	The baseline period is the lead-in period up to the day of first dose of study treatment. Baseline will be derived as the average of all fasting glucose measurements between V2 date and the first dose date.	All available data after the day of first dose of study treatment up to Week 26 (Visit 23) in EAS2 for MMRM. Values at each visit will be derived as average of all fasting glucose measurements from the day post prior visit up to the day of the next visit. Data at Week 26 (Visit 23) in EAS1 for ANCOVA with multiple imputation for missing data.

Analysis	Baseline Observations	Post-Baseline Observations
Fasting serum glucose	The baseline is the last non-missing assessment prior to or on the day of the first dose of study treatment.	All scheduled visits after the day of first dose up to Week 26 (Visit 23) for MMRM Planned measurements at scheduled visits will be included. Data at Week 26 (Visit 23) in EAS1 for ANCOVA with multiple imputation for missing data.
Participant-reported hypoglycemia	The baseline period is the screening/lead-in period prior to the day of first dose of study treatment.	• Treatment period starts on the day of first dose of study treatment and ends o at Week 26 (Visit 23) if completed treatment and Visit 23 o on the last dose date of study treatment + 10 days for LY3209590, +1 day for glargine if discontinued study treatment early • Post-treatment period starts from o Week 26 (Visit 23)+1 day if completed treatment and Visit 23 o last dose date of treatment +11 days for LY3209590, +2 days for glargine if discontinued study treatment early and ends on the last date in the study
TEAEs	The baseline period includes the screening/lead-in period up to the first dose of study treatment (AE Start Relative to Exposure Assessment in CRF is used to determine).	Safety analysis period starts after the first dose and ends at the last visit in the study including safety follow-up period.
Safety laboratory tests, vital signs and body weight	The baseline will be the last non-missing assessment prior to or on the day of first dose of study treatment Planned measurements at scheduled visits will be included.	All scheduled visits after the day of first dose up to Visit 802 for MMRM or ANCOVA Planned measurements at scheduled visits will be included.

Analysis	Baseline Observations	Post-Baseline Observations
laboratory values elevated or low, vital signs and body weight categorical measures	Starts from the screening visit and ends prior to or on the day of first dose of study treatment. All available measurement at scheduled and unscheduled visits will be included. The baseline for weight will be the last non-missing value during the baseline period	Starts after the day of first dose of study treatment and ends at the last visit in the study including both treatment period and follow-up period. All available measurement at scheduled and unscheduled visits in the specified analysis period will be included.
Anti-LY3209590 antibody	The baseline will be the last non-missing assessment prior to or on the day of first dose of study treatment	All scheduled and unscheduled measurement after the day of first dose of study treatment .
Patient-reported outcomes	The baseline will be the data collected at Visit 3	Last collection after Visit 3

Abbreviations: ANCOVA = analysis of covariance; CGM = continuous glucose monitoring; CRF = case report form; EAS1 = efficacy analysis set 1; EAS2 = efficacy analysis set 2; MMRM = mixed model repeated measurement; TEAE = treatment-emergent adverse event.

All analyses will be implemented using SAS Enterprise Guide Version 7.1 or above.

4.2. Participant Dispositions

Reasons for discontinuation prior to randomization including screen failure will be summarized for all participants who sign the informed consent form.

The number and percentage of participants who have completed or discontinued from the study or treatment will be summarized by treatment for the Randomized Population. The individual reasons for discontinuation will also be included in the summary. A Fisher's exact test will be conducted for comparison between LY3209590 and insulin glargine.

A listing of the reasons for study/treatment discontinuations will be generated for the randomized Population. A listing of the randomized treatment for this study will also be provided.

Time to permanent discontinuation of study treatment, time to study discontinuation, and time to early discontinuation of study treatment due to AEs (if there is a sufficient number of participants to warrant a summary) will be presented as a figure.

4.3. Primary Endpoint Analysis

4.3.1. Definition of Endpoint

The primary endpoint of this study is change in HbA1c from baseline at Week 26 (Visit 23). The HbA1c is reported in unit of % by central laboratory and will be converted to the unit of mmol/mol using the following formula: HbA1c in mmol/mol = 10.93 * HbA1c in % - 23.5 (IFCC 2010). The HbA1c analysis will be conducted for both units.

4.3.2. Main Analytical Approach

The primary objective is to test the hypothesis that LY3209590 is noninferior to insulin glargine on glycemic control in the targeted study population. The noninferiority test will be based on the

- ***treatment regimen estimand*** for the US FDA submission, and
- ***efficacy estimand*** for registrations in other countries.

The full significance level of 0.05 will be used for each estimand.

This table provides the details of treatment regimen estimand and efficacy estimand.

	Treatment Regimen Estimand	**Efficacy Estimand**
Analysis Population	All participants in EAS1 with non-missing baseline measure	All participants in EAS2 with non-missing baseline measure and at least 1 non-missing post-baseline scheduled measure
Analysis Data	All non-missing observations at baseline and Week 26 (Visit 23) regardless of the use of study intervention or rescue therapy.	All non-missing observations at baseline and all scheduled post-baseline timepoints during treatment period (that is, Weeks 2, 4, 12, 16, and 26) prior to the date of last study dose + 10 days for LY3209590 (+1 day for glargine), or the initiation of rescue therapy, whichever is earlier for participants who had intercurrent events

	Treatment Regimen Estimand	Efficacy Estimand
Missing Data	There may be missing values at Week 26 (Visit 23) due to early study discontinuation. The missing values will be imputed using multiple imputation by the retrieved dropout approach. The retrieved dropout participants are those who discontinue study intervention prior to Week 26 but have non-missing measures at Week 26 in the same treatment arm. If there are only a limited number of retrieved participants that leads to a failure in performing the multiple imputation analysis, such as the model cannot converge, or the number of retrieved dropout participants is small (i.e., at least 1 arm has < 8 participants who discontinued study treatment early and have endpoint visit measurements), the missing HbA1c at Week 26 will be imputed by return-to-baseline multiple imputations approach (Qu and Dai, 2022).	The MMRM model will be used, and the missing values will be handled in the MMRM analysis under the assumption of missing at random.

	Treatment Regimen Estimand	Efficacy Estimand
Analysis Model	After the imputation, the observed and imputed data will be analyzed by the ANCOVA model using treatment, strata (country, routine use of personal CGM or FGM at randomization), and baseline value of the dependent variable as independent variables. The statistical inference will be based on the multiple imputation framework by Rubin (1987).	The MMRM model will include treatment, strata (country, and routine use of personal CGM or FGM at randomization), visit and treatment-by-visit interaction as fixed effects, and baseline value of the dependent variable as a covariate. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom for the MMRM models. An unstructured covariance structure will be used to model the within-participant errors. If this structure fails to converge, the following covariance structures will be used in order until one converges: 1. Toeplitz with heterogeneity 2. Autoregressive with heterogeneity 3. compound symmetry with heterogeneous variances 4. Toeplitz 5. Autoregressive 6. Compound symmetry without heterogeneous variances.

Abbreviations: ANCOVA = analysis of covariance; CGM = continuous glucose monitoring; EAS1 = efficacy analysis set 1; EAS2 = efficacy analysis set 2; FGM = flash glucose monitoring; HbA1c = hemoglobin A1c; MMRM = mixed model repeated measurement.

The 2-sided 95% CI of the LS mean difference (LY3209590 – insulin glargine) in the HbA1c (in %) change from baseline at Week 26 (Visit 23) will be estimated. For both estimands, LY3209590 will be declared noninferior to insulin glargine if the upper limit of the 2-side 95% CI for the LS mean difference in the HbA1c change from baseline is below NIM of +0.4%. In addition, the 95% CI for the LS mean difference will be compared to an alternative NIM of +0.3%.

4.3.3. Sensitivity Analysis

4.3.3.1. Two-way Tipping Point Analysis

To confirm the robustness of the primary endpoint results, a 2-way tipping point analysis represents varying assumptions for missing data from both treatment groups will be conducted. This sensitivity analysis will focus on missing data at the primary endpoint. Penalty parameters for imputed missing values will be added for both treatment arms. The ANCOVA model for treatment regimen estimand will be conducted after penalty parameters are added. The multiple imputation framework by Rubin (1987) will be used to summarize the results.

The non-inferiority test will be shown by color scale in the figure.

Additionally, imputation under the noninferiority null hypothesis will be conducted by adding 0.4 (NIM) to the imputed data used for treatment regimen estimand (Section 4.3.3) for the LY3209590 group only (Koch 2008). The ANCOVA model for treatment regimen estimand will

be rerun using the adjusted data. The multiple imputation framework by Rubin (1987) will be used to summarize the results.

4.3.3.2. Including Inadvertently Enrolled Participants

The primary efficacy analysis will be repeated for the mITT population, including participants inadvertently enrolled, for both treatment regimen estimand and efficacy estimand.

4.3.4. Supplementary Analyses

Additional analysis may be conducted as needed.

4.4. Secondary Endpoints Analysis

4.4.1. Key Secondary Endpoints

A graphical approach will be used to control the overall type I error for the primary objective and the key secondary objectives. The key secondary objectives will test the superiority of LY3209590 compared with insulin glargine for the following endpoints.

1) change in HbA1c from baseline at Week 26 (Visit 23)
2) the percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26 (Visit 23), and
3) the event rate of nocturnal hypoglycemia (<54 mg/dL or severe) during the treatment phase up to Week 26 (Visit 23).

4.4.1.1. Definition of Endpoint(s)

See Section 4.3.1 for change in HbA1c from baseline at Week 26 (Visit 23).

The hypoglycemia events will be based on participant entry into an e-diary, which receives all blood glucose (BG) measurements performed by the participant and transmitted via Bluetooth from the study-provided glucometer. The nocturnal period is defined by midnight to 6 AM. The event rate of nocturnal hypoglycemia will be based on the count of the hypoglycemia episodes in the nocturnal period and the corresponding exposure of study intervention.

Participants achieving HbA1c<7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L], or severe) during treatment phase up to Week 26 will be a composite measure based on HbA1c measured at Week 26 and nocturnal hypoglycemia reported by participants during the 26-week treatment period.

4.4.1.2. Main Analytical Approach

The superiority test for change in HbA1c from baseline at Week 26 (Visit 23) will be based on the same primary endpoint analysis described in Section 4.3.2.

The percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L], or severe) during treatment phase up to Week 26 will be analyzed by a composite estimand. This table describes the details.

	Treatment Regimen Estimand	Efficacy Estimand
Analysis population	All participants in Efficacy Analysis Set 1 with non-missing baseline measure.	All participants in Efficacy Analysis Set 2 with non-missing baseline measure
Analysis data	All non-missing observations at Week 26 (Visit 23) during treatment period regardless of the use of study intervention or rescue therapy.	All non-missing observations at Week 26 (Visit 23) during treatment period without discontinuation of study treatment or initiation of rescue therapy.
Endpoint	The composite • binary outcome of HbA1c <7% at Week 26 with 1 indicating achieving HbA1c target • binary outcome of no nocturnal hypoglycemia (<54 mg/dL or severe) during 26-week treatment period with 1 indicating no occurrence of nocturnal hypoglycemia.	Same method as for treatment regimen estimand
Missing data handling	• For HbA1c, missing values at Week 26 (Visit 23) will be imputed using the same method for the primary endpoint. The binary outcome of HbA1c <7% will be based on the imputed data. • For nocturnal hypoglycemia, participant who discontinued the treatment period before Week 26 (Visit 23) is considered as a non-responder (i.e. experiencing the event) and the binary outcome value is 0.	For HbA1c, missing values at Week 26 (Visit 23) will be imputed using multiple imputation with assumption of missing at random. For nocturnal hypoglycemia, participant who discontinued the treatment period before Week 26 (Visit 23) is considered as a non-responder (i.e. experiencing the event) and the binary outcome value is 0.
Analysis model	The composite endpoint will be analyzed using a logistic regression model including treatment, strata (country and routine use of personal CGM or FGM at randomization), baseline HbA1c value, and baseline incidence of nocturnal hypoglycemia. Odds ratio (LY3209590 vs. insulin glargine) and 95% confidence interval, and p-value will be used for treatment comparison. Multiple imputation will be performed, and the statistical inference will be based on the multiple imputation framework by Rubin (1987).	Same model as for treatment regimen estimand

The rate of participant-reported nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) will be analyzed by a negative binomial regression. The analysis details are provided in Section 4.6.2.2.1.

4.4.1.3. Sensitivity Analysis

The analyses described in Section 4.4.1.3 will be repeated for both the treatment regimen estimand and efficacy estimand by including inadvertent enrolled participants using mITT population.

4.4.2. Supportive Secondary Endpoints

4.4.2.1. Other Efficacy Endpoints

The analyses for CGM variability parameters, insulin dose and patient reported outcomes will be based on the efficacy estimand using EAS2 data.

The analysis of CGM data (except for variability parameters) and fasting glucose from SMBG will be done for both treatment regimen estimand and efficacy estimand. Analysis of covariance for treatment regimen estimand and MMRM model for efficacy estimand similar to those for the primary endpoint with an additional term of baseline HbA1c stratum (<8.0%, ≥8.0%) will be used.

The variance-covariance structure in the MMRM models for fasting glucose from SMBG and insulin dose will be based on compound symmetry.

For treatment regimen estimand, only subjects with an observation at baseline or at Week 26 (Visit 23) will be included in the analysis. The missing baseline will be imputed using multiple imputation with assumption of missing at random. Missing data at Week 26 (Visit 23) will be imputed as described for the primary HbA1c endpoint.

For the efficacy estimand, only participants with baseline and at least one post baseline observation will be included in the analysis, missing data will be handled by MMRM model.

4.4.2.1.1. Analysis for CGM Parameters

The time in glucose ranges will be based on the percentage of CGM readings within the glucose ranges, and glucose variability will be based on the readings during the specific CGM session. Section 6.6 (Appendix 6) provides the derivation of CGM parameters.

Longitudinal logistic regression model with treatment, strata (country, routine use of personal CGM or FGM at randomization and baseline HbA1c stratum (<8.0% and ≥8.0%)), time, and treatment-by-time interaction as fixed effects, baseline value as a covariate will be used for analysis for CGM targets of glycemic control. See Section 4.3.2 for variance-covariance structure selection.

Time in glucose range between 70 and 180 mg/dL (3.9 and 10.0 mmol/L) inclusive collected during the CGM session that occurs between Week 22 and Week 26 is not a key secondary endpoint. To be consistent with other studies in the same program, this CGM parameter will also be analyzed using an ANCOVA model for treatment regimen estimand similar to that for the primary endpoint. The missing data will be imputed by multiple imputation with the approach similar to the imputation used for the primary endpoint.

The time point for primary CGM data analysis is CGM session. In addition, the average time in glucose range, hypoglycemia range and hyperglycemia range by week (by visit for treatment period) will be summarized by treatment. Average daily time since last dose by CGM Session during treatment period will also be summarized for LY3209590. Missing data patterns will be summarized at daily, weekly (visit) and CGM session levels.

4.4.2.1.2. Study Insulin Dose Analysis

For study basal insulin dose , both average daily and weekly dose during each visit for individual participants will be analyzed. In the LY3209590 group, the average weekly dose of each visit (that is, average of weekly doses since last visit) will be used as the average weekly dose, and the average daily dose will be the average weekly dose divided by 7. For insulin glargine group, the average daily dose since last visit will be computed for each visit, and the average weekly dose will be calculated as average daily dose × 7. For participants using typical once-daily pre-study basal insulin, the average pre-study basal dose during the screening/lead-in period will be the corresponding baseline daily dose. For participants using pre-study U-300 or Neutral Protamine Hagedorn twice daily (NPH BID), the average pre-study basal dose during the lead-in period multiplying by 0.8 will be the baseline daily dose. The baseline daily dose multiplying by 7 will be the baseline weekly basal insulin dose. The corrected baseline basal insulin dose will be controlled in the model for the insulin dosing analysis.

The daily prandial insulin dose will be calculated as the sum of the prandial insulin doses for morning meal, midday meal, evening meal, and additional dose of the day. All data including 0 dose and single dose of a day are included for analysis. On the first day of study treatment, a participant may have used both pre-study prandial insulin and study provided insulin lispro. Therefore, Week 0 (first day of study treatment) will be excluded from calculating the average daily prandial insulin for baseline and post-baseline period. The weekly prandial insulin dose at each visit is the average daily prandial insulin dose at the same visit×7.

The weekly total insulin dose is the sum of average weekly basal insulin dose and average weekly prandial insulin dose. The daily total insulin dose is weekly total insulin dose divided by 7.

The basal/total insulin dose ratio at each visit is the ratio of average weekly basal insulin dose and average weekly total insulin dose at the same visit.

If either weekly basal insulin dose or weekly prandial insulin dose is missing, the total weekly insulin dose and basal/total insulin dose ratio will be set as missing for analysis. Similarly, the total daily insulin dose is also set as missing.

The analysis will be done for

- both daily dose and weekly dose in U for each of the basal, prandial and total insulin, respectively.
- dose in U/kg/day for basal, prandial and total insulin, respectively
- basal/total insulin dose ratio

4.4.2.2. Other Safety Endpoints

The safety measures based on CGM data will be analyzed as described in Section 4.6.2.2.2. For other safety measures, the details are provided in Section 4.6.

4.4.2.3. Participant Reported Outcomes

Participant-reported outcome analyses will be based on the EAS2 data.

4.4.2.3.1. Diabetes Injection Device Experience Questionnaire (DID-EQ) Version 1.0

Description of DID-EQ

The DID-EQ (Matza et al. 2018) is a self-administered, 10-item questionnaire designed to assess participants' perceptions of diabetes injection delivery systems for T2D. Each item is rated on a 4-point scale with higher scores indicating more positive perceptions of the injection device. Items 1 to 7 focus on specific characteristics of diabetes injection device, and these 7 items comprise the Device Characteristics subscale. To compute the Device Characteristics subscale score, the seven individual item scores are first summed, resulting in a subscale raw score and then transformed to a scale of 0 to 100 to obtain a transformed domain score.

The transformed domain score can be calculated by

$$\frac{(\text{sum of raw score - sum of lowest possible raw score})}{(\text{sum of highest possible raw score - sum of lowest possible raw score})} \times 100$$

For an example, DID-EQ = ((sum of raw score -7)/ (28-7)) ×100. If answer for the 7 questions (Question 1 to 7) are all 1, then the sum of raw score is 7 and the transformed domain score is 0.

In addition, there are 3 global items:

- Item 8 assessing overall satisfaction
- Item 9 ease of use, and
- Item 10 convenience of diabetes injection devices.

These 3 global items are each scored separately.

Summary statistics for the Device Characteristics subscale will be provided by study treatment. The last non-missing postbaseline measures up to Week 26 (Visit 23) will be analyzed by an ANCOVA model using treatment, strata (country, routine use of personal CGM or FGM at randomization, and baseline HbA1c stratum (<8.0%, ≥8.0%)) as independent variables.

The frequency and proportion of the 3 global items at Week 26 (Visit 23) will be summarized by study treatment. Wilcoxon rank sum test will be used for treatment comparison.

4.5. Tertiary Endpoint Analysis

4.5.1. Tertiary Efficacy Endpoints

Analysis for fasting serum glucose will be based on both EAS1 and EAS2 similar to those described in Section 4.4.2.1. Patient reported outcomes will be based on EAS2 .

Treatment comparison for the proportion of participants achieving HbA1c target (<7% or ≤6.5%) will be analyzed using the same method for the key secondary endpoint.

4.5.2. Tertiary Safety Endpoints

For tertiary safety measures, the details are provided in Section 4.6 and the following subsections.

4.5.3. Participant Reported Outcome

The analyses for participant-reported outcomes will be based on the EAS2 data.

4.5.3.1. Basal Insulin Experience

Likelihood of incorporating into routine

This is a Lilly-developed, participant-completed question to understand the participant's likelihood of incorporating their study insulin into their diabetes management routine. The question is rated on a 5-point scale with responses ranging from "very unlikely" to "very likely."

The frequency and proportion of the last non-missing responses up to Week 26 (Visit 23) will be summarized by study treatment and using Wilcoxon rank sum test for treatment comparison.

Preference

This is a Lilly-developed, participant-completed question to understand the participant's preference for their pre-study or current study treatment. The question is rated on a 5-point scale with responses ranging from "strongly prefer the study insulin" to "strongly prefer my previous insulin." The question also includes a "not applicable" option for participants that stayed on the same insulin in the treatment phase.

The frequency and proportion of the last non-missing responses up to Week 26 (Visit 23) will be summarized by study treatment and using Wilcoxon rank sum test for treatment comparison.

4.5.3.2. EQ-5D-5L

The EQ-5D-5L (EuroQol 2019) is a standardized 5-item self-administered instrument for use as a measure of health outcome. It provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care as well as population health surveys. The EQ-5D-5L assesses 5 dimensions of health:

- mobility
- selfcare
- usual activities
- pain/discomfort
- anxiety/depression

The 5L version, scores each dimension at 5 levels:

- no problems
- slight problems
- moderate problems
- severe problems
- unable to perform/extreme problems

In addition to the health profile, a single health state index value can be derived based on a formula that attaches weights to each of the levels in each dimension. This index value ranges between less than 0 (where 0 is a health state equivalent to death; negative values are valued as worse than dead) to 1 (perfect health).

The second part of the questionnaire consists of EQ Visual Analog Scale (VAS) which records the respondent's self-rated health status. The participant rates his or her perceived health from 0 (the worst imaginable health) to 100 (the best imaginable health). In conjunction with the health state data, it provides a composite picture of the respondent's health status.

The frequency and proportion of each level for a given item will be summarized by study treatment for the scheduled visit. The last non-missing postbaseline observations up to Week 26 (Visit 23) on EQ-5D-5L index, EQ VAS score and their change from baseline will be analyzed by an ANCOVA model using the independent variables of treatment, strata (country, baseline HbA1c stratum (<8.0%, ≥8.0%), and routine use of personal CGM or FGM at randomization), and baseline value of the dependent variable.

4.6. Safety Analyses

Safety measures include treatment exposure, adverse event (AE), vital signs, weight, hypoglycemia, laboratory measures and immunogenicity. All safety analyses will be based on the SS. Unless otherwise specified, safety analysis period will include both treatment period and follow-up period as specified in Section 4.1. Percentages will be calculated using the safety population as the denominator. For events that are gender-specific, the denominator and computation of the percentage will include only participants from the given sex. Unless otherwise noted, Fisher's exact test will be used for treatment comparison, risk difference and 95% confidence intervals will be provided.

For continuous safety variables (for example, laboratory measures, vital signs, and weight), descriptive statistics for the observed values and change from baseline at scheduled visits during the treatment period and follow-up period will be provided. For some selected laboratory measures (i.e. liver enzyme tests, lipid measures), observed values, change from baseline and percentage change from baseline will be analyzed for the log-transformed values by the MMRM model using treatment, visit, and treatment by visit as fixed effects, baseline of the dependent variable as a covariate, and compound symmetry as the variance-covariance structure.

The incidence and event rate of participant-reported and CGM-based hypoglycemia will be summarized by treatment and analysis period for different types of hypoglycemia. Analysis details are provided in Section 0.

4.6.1. Extent of Exposure

Duration of exposure to study treatment will be summarized. No p-values will be reported in these tables as they are intended to describe the study populations, rather than test hypotheses about them. Total patient-years of exposure will be reported. The number and proportion of participants falling into the following different exposure categories will also be summarized by study treatment

- >0, ≥30 days, ≥90 days, ≥180 days
- >0 and < 30 days, ≥30 and < 90 days, ≥90 and <180 days, ≥180 days

Exposure on study treatment will be calculated as

- LY3209590: date of the last treatment administration – date of first treatment administration + 7 days
- Insulin glargine: date of the last treatment administration – date of first treatment administration + 1 day

Total patient-years of exposure will be calculated by the sum of duration of exposure in days divided by 365.25. The following summary statistics will be provided: n, mean, standard deviation, median, minimum, maximum, interquartile range, and total exposure (that is, total patient-years).

All participants who complete the study treatment period are required to complete a safety follow-up period and participants who discontinue the study treatment prematurely are encouraged to remain in the study for safety monitoring. The duration on study from the first dose of study treatment to the study disposition date will also be summarized by treatment.

A listing of exposure to study treatment will be provided.

4.6.2. Adverse Events

Events that are newly reported after the first dose of investigational product (IP) or reported to worsen in severity from baseline will be considered treatment emergent adverse events (TEAEs). The Medical Dictionary for Regulatory Activities (MedDRA) lowest level term (LLT) will be used in the treatment-emergent assessment. The maximum severity for each LLT during the baseline period (see the table in Section 4.1) will be used as baseline severity.

The table below describes the analysis related to AEs.

Analysis	Details
Overview of AEs	Number and percentage of participants who experienced: • SAE • Death • Discontinuation from study treatment due to an AE • Discontinuation from study due to an AE • TEAE • TEAE related to study treatment
Summary by PT within SOC	Number and percentage of participants with AEs using MedDRA PT nested within SOC: • TEAE • Maximum Severity TEAEs • SAE • AE leading to permanent discontinuation of study treatment Events will be ordered by decreasing risk difference within SOC. SOCs will be listed by decreasing risk difference . For each participant and TEAE, the maximum severity for the MedDRA PT is the maximum postbaseline severity observed from all associated LLTs mapping to the MedDRA PT. The maximum severity will be determined based on the non-missing severities.
Summary by PT (within SMQ when applicable)	Number and percentage of participants with TEAEs using MedDRA PT (irrespective of SOC):

Analysis	Details
	• TEAEs occurring in ≥1% before rounding in LY3209590 group. Events will be ordered by decreasing risk difference and by decreasing frequency in LY3209590 group, respectively • TEAE of safety topic of interest by PT (within SMQ when applicable). Events will be ordered by decreasing risk difference.
Listing	Separate listings for the following events will be provided: • SAE including death • AEs leading to study treatment discontinuation • Severe hypoglycemia • Events sent to the external adjudicator for MACE adjudication • Participants who receive rescue therapy due to severe/persistent hyperglycemia • Persistent-recurrent hypoglycemia reported by investigators • Persistent-recurrent hypoglycemia identified by programing • Medication errors of interest

Abbreviations: AE = adverse event; LLT = lowest level term; MACE = major adverse cardiovascular event; MedDRA = Medical Dictionary for Regulatory Activities; SAE = serious adverse event; SOC = System Organ Class; PT = preferred term; TEAE = treatment emergent adverse event.

4.6.2.1. Safety Topics of Interest

4.6.2.1.1. Severe Hypoglycemia

Severe hypoglycemia is a severe event characterized by altered mental and/or physical status requiring assistance for treatment of hypoglycemia. The determination of a hypoglycemic event as an episode of severe hypoglycemia is made by the investigator based on the medical need of the participant to have required assistance and is not predicated on the report of a participant simply having received assistance. If a hypoglycemic event meets the criteria of severe hypoglycemia, the investigator must record the event as serious on the AE case report form (CRF).

A summary of severe hypoglycemia by preferred term (PT) and ordered by decreasing risk difference between LY3209590 and glargine will be provided. Section 4.6.2.2.1 describes the method for analyzing the incidence and event rate during the treatment period. A listing of severe hypoglycemia events will also be provided.

4.6.2.1.2. Persistent-Recurrent Hypoglycemia

The potential risk of persistent-recurrent hypoglycemia (P-R hypoglycemia) will be assessed from the first dose date up to the end of the study. P-R hypoglycemia events will be identified using

1. investigator assessment and clinical judgment to determine if repeated hypoglycemia may have contributed to a hypoglycemic event that degenerated with poor outcomes, and
2. a prespecified criteria to derive events from e-diary database by programming (see definition Appendix 9, Figure BDCV. 6.9.).

Identification of P-R hypoglycemia based on investigator reporting is precipitated by a hypoglycemic event a participant or their caregiver informed as having resulted in a poor

outcome. This information will trigger an e-mail alert notifying the investigator to contact the study participant, obtain more information about the specific hypoglycemic event, and provide the participant clinical guidance, if appropriate. Hypoglycemic events that trigger alerts to investigators are those participants or caregivers report in the e-diary have required treatment with glucagon or IV glucose, resulted in coma, motor vehicle accident or other trauma, hospitalization, or emergency medical care.

Upon receiving the e-mail notification, investigator will access the e-diary database and answer the following question: *"In your clinical judgement, is this hypoglycemia event associated with repeated hypoglycemia events?"*. Investigator should select *"Yes"* or *"No"* as an answer.

During protocol training, investigators are trained to consult the e-diary database and review participant's BG values and hypoglycemia reports to determine the best answer. If the investigator judges the hypoglycemic event that triggered an alert is related to repeated episodes of hypoglycemia, the participant will be identified as having presented a P-R hypoglycemic event reported by the investigator.

Summary statistics and a listing of the events identified by both methods will be provided.

4.6.2.1.3. Systemic Hypersensitivity Reactions

Hypersensitivity reactions are exaggerated or inappropriate immunologic responses occurring in response to an antigen or allergen. These can be systemic or localized. At all visits, participants will be evaluated by the investigator for signs and symptoms suggestive of hypersensitivity. Investigators will complete an CRF designed to record additional information about AEs suggestive of a hypersensitivity reaction. The TEAEs of hypersensitivity reactions were identified using

- *Anaphylactic reaction* SMQ (20000021; narrow terms)
- *Hypersensitivity* SMQ (20000214; narrow terms), and
- *Angioedema* SMQ (20000024; narrow terms).

The number and percentage of participants who reported a TEAE for the following will be analyzed

- Any narrow term from any 1 of the 3 SMQs indicated above (that is, combined search across narrow portions of all 3 SMQs)
- Any narrow term within each SMQ, separately (that is, narrow SMQ search)

Individual PTs that satisfied the query will appear in the summary in decreasing order of risk difference between the LY3209590 and insulin glargine group.

The analyses above are the starting point for medical interpretation of any apparent differences between treatment groups. For notable events, case review will be applied to make the final determination of whether an event is most accurately described as a potential hypersensitivity reaction to study treatment, or another event that is not clearly associated with study treatment administration. This judgment will be on the basis of totality of information available, including the content of a follow-up CRF collected for potential hypersensitivity events.

4.6.2.1.4. Injection Site Reactions

Injection site reactions (ISR) are AEs localized to the immediate site of the administration of a medication. The evaluation of ISRs will be through the unsolicited reporting of ISR TEAEs and through the use of an Injection Site Reaction Follow-up Form completed by the investigator for each incidence of ISR. A summary of the number of participants with reported events meeting any of the following categories will be provided

- MedDRA HLT of Injection site reactions
- MedDRA HLT of Administration site reactions NEC
- Lipodystrophies and localized amyloidosis, as represented by PTs of
    - Lipoatrophy
    - Lipodystrophy acquired
    - Partial lipodystrophy
    - Lipohypertrophy
    - Sclerema, and
    - Cutaneous amyloidosis.

The summary will present the number of participants who reported

- at least 1 AE meeting any of the above categories
- any AE in each category, and
- any AE for each PT within a specific category.

The PTs will be listed for summary within each category in decreasing order of risk difference between the LY3209590 and insulin glargine group.

The additional data collected on the ISR follow-up forms will be summarized in 2 distinct ways:

1. at the participant-level (each participant contributes to totals at most once, regardless of the number of ISR events the participant experienced), and
2. at the event level (each separate event contributes to totals, regardless of whether the participant experienced multiple ISR events).

4.6.2.1.5. Neoplasms

The TEAEs of neoplasms will be identified by narrow search using the MedDRA SMQ

- *Malignant tumours* SMQ (20000194, narrow terms), and
- *Tumours of unspecified malignancy* SMQ (20000195, narrow terms).

A summary will present any narrow term

- from any 1 of the 2 SMQs indicated above (that is, combined search across narrow portions of both SMQs), and
- within each SMQ, separately (that is, narrow SMQ search).

4.6.2.1.6. Diabetic Ketoacidosis (DKA)

DKA will be searched by MedDRA PTs from all TEAEs. The number and percentage of participants experiencing treatment-emergent DKA will be summarized. The TEAEs of DKA will be identified using MedDRA PTs of

- Diabetic ketoacidosis
- Ketoacidosis
- Euglycaemic diabetic ketoacidosis
- Ketonuria
- Diabetic ketosis
- Diabetic ketoacidotic hyperglycaemic coma
- Ketosis
- Urine ketone body present
- Blood ketone body
- Blood ketone body increased
- Urine ketone body
- Blood ketone body present, and
- Lactic acidosis.

4.6.2.1.7. Diabetic Retinopathy or Maculopathy

Diabetic retinopathy or maculopathy will be searched by MedDRA PTs from all TEAEs. Section 6.7 (Appendix 7) lists the PTs that will be used to identify TEAEs of diabetic retinopathy or maculopathy. A summary of retinopathy or maculopathy by PT will be provided.

4.6.2.1.8. Peripheral Edema

Peripheral edema TEAEs will be identified by MedDRA PTs (See Section 6.8). The number and percentage of participants experiencing treatment-emergent peripheral edema will be summarized by PT.

4.6.2.1.9. Hypokalemia

The TEAEs of hypokalemia will be identified by narrow terms in *Hypokalaemia* SMQ (20000233). A summary of the number of participants with reported TEAEs meeting the SMQ narrow search criteria by PT will be provided.

4.6.2.1.10. Hyperglycemia

The study treatments were designed as the treatment of hyperglycemia for diabetes patients. Therefore, the hyperglycemia is usually not reported as an AE in diabetes studies. However, if a participant develops severe, persistent hyperglycemia after randomization, a rescue therapy will be considered. A listing of participants who receive rescue therapy will be provided.

4.6.2.1.11. Major Adverse Cardiovascular Events (MACE)

Potential cerebrocardiovascular events will be identified by the investigative site or by a medical review conducted by the sponsor or designee. A blinded external Clinical Event Committee will adjudicate the events in a consistent and unbiased manner. Events include

- death
- cardiac ischemic events including
  - myocardial infarction,
  - hospitalization for unstable angina
- cerebrovascular events including
  - stroke
  - transient ischemic attack.
- hospitalization for heart failure, and
- coronary revascularization procedure.

Only confirmed MACE by the adjudication committee will be considered as AESIs. A listing of MACE events reported by investigator, including reported term and adjudication results, will be provided.

4.6.2.1.12. Medication Error of Interest

Medication Errors of Interest (MEI) are defined as medication error AEs (SMQ 20000224 - narrow and broad terms) that meet the criteria of important protocol deviation (IPD) indicative of multiple doses, according to the Trial Issue Management Plan. These events are considered IPDs and of special interest because of their potential to impact participant's safety.

MEI AEs are categorized as IPDs of "Investigational Medicinal Product and/or Investigational Device." Screening and identification of MEI AEs will occur during routine review of the protocol deviations and trial level safety reviews.

The number and percentage of participants reported with MEI will be analyzed.

A listing of MEI will be provided. The listing will indicate if severe hypoglycemia or P-R hypoglycemia occurred after the MEI.

4.6.2.2. Hypoglycemic Events

4.6.2.2.1. Participant-Reported Hypoglycemic Events

The following types of hypoglycemic events will be derived in the analysis data sets: documented hypoglycemia as Level 1, Level 2, and Level 3 (severe hypoglycemia) according to definitions based on the American Diabetes Association criteria where

- Level 1: glucose <70 mg/dL (3.9 mmol/L) and ≥54 mg/dL (3.0 mmol/L).
- Level 2: glucose <54 mg/dL (3.0mmol/L).
- Level 3: severe hypoglycemia (confirmed by the investigator to be an event that required assistance for treatment).

Level 2 and Level 3 events are considered clinically significant hypoglycemia. Therefore, the analysis on a composite of Level 2 and Level 3 (denoted as Level 2/3) hypoglycemia will also be conducted.

Hypoglycemic events will also be further classified into

- nocturnal hypoglycemia (occurs between midnight and 0600), and
- non-nocturnal hypoglycemia (occurs between 0600 and midnight).

If a hypoglycemic event is within 60 minutes of another hypoglycemic event, it is considered as a continuation of the previous event. If there are multiple hypoglycemic events within 60 minutes of each other, then all events will be combined into a single event, which has the

- earliest date time
- minimum glucose value, if applicable
- maximum severity (Level 1, 2 or 3)
- combined symptoms and outcomes
- time of nocturnal if any of the events is nocturnal

of the multiple hypoglycemic events.

The combined event starts from the first record with Level 1, 2 or 3, and ends when there are no more events for at least 60 minutes.
A sensitivity analysis will be done for selected hypoglycemic endpoints where all hypoglycemic events are considered one hypoglycemic event until a succeeding glucose value is ≥ 70 mg/dL.

The evaluation of potential persistent-recurrent hypoglycemic events will be through assessment of the clinically significant events confirmed by investigators based on clinical judgment and through events identified by a prespecified criteria (Section 4.6.2.1.2 and Appendix 9, Section 6.9) using information based on the participant-reported hypoglycemia.

The details of planned analyses are provided in the table below.

Endpoint	Analysis Period	Statistical Method
Event rate of Level 1 hypoglycemic events (events/participant/year): • All documented	Baseline, 0-6, 0-12, 0-26, 12-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline HbA1c and baseline Level 1 hypoglycemia rate as covariates, log (exposure/365.25 days) as the offset in the model.
Event rate of Level 2 hypoglycemic events (events/participant/year): • All documented	Baseline, 0-6, 0-12, 0-26, 12-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline HbA1c and baseline Level 2 hypoglycemia rate as covariates, log (exposure/365.25 days) as the offset in the model.

Endpoint	Analysis Period	Statistical Method
Event rate of Level 3 hypoglycemic events (events/participant/year): • All documented	Baseline, 0-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline Level 3 hypoglycemia rate and baseline HbA1c as covariates, log (exposure/365.25 days) as the offset in the model. If the number of events is too small to run the negative binomial regression, exposure adjusted rate calculated by total number of events divided by total exposure for individual patients) will be provided and the empirical method (see Appendix 11 (Section 6.11) for details) will be used for treatment comparison.
Event rate of Level 2/3 hypoglycemic events (events/participant/year): • All documented • Nocturnal • Non-nocturnal	Baseline, 0-6, 0-12, 0-26, 12-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline HbA1c and baseline hypoglycemia rate of the same hypoglycemia type as covariates, log (exposure/365.25 days) as the offset in the model. The plots of the mean cumulative functions (MCFs) by each treatment arm will also be created. The population mean for cumulative number of events up to time t, M(t), will be estimated using a nonparametric estimation method described by Nelson (2003).
Incidence of Level 1 hypoglycemic events: • All documented	Baseline, 0-6, 0-12, 0-26, 12-26 weeks, post-treatment period	Logistic regression with treatment, baseline HbA1c and baseline Level 1 hypoglycemia incidence as covariates.
Incidence of Level 2 hypoglycemic events: • All documented	Baseline, 0-6, 0-12, 0-26, 12-26 weeks, post-treatment period	Logistic regression with treatment, baseline HbA1c and baseline Level 2 hypoglycemia incidence as covariates.
Incidence of Level 3 hypoglycemic events: • All documented	Baseline, 0-26 weeks, post-treatment period	Logistic regression with treatment and baseline HbA1c as covariates.

Endpoint	Analysis Period	Statistical Method
Incidence of Level 2/3 hypoglycemic events: • All documented • Nocturnal • Non-nocturnal	Baseline, 0-6, 0-12, 0-26, 12-26 weeks, post-treatment period	Logistic regression with treatment, baseline HbA1c and baseline hypoglycemia incidence of the same hypoglycemia type as covariates.
Potential Persistent-recurrent hypoglycemic events • Identified by investigator • Identified by prespecified criteria (defined in Section 6.11)	Safety analysis period (see definition in Section 4.6)	The number of participants with at least 1 event will be summarized and compared by Fisher's exact test. The number of events will also be provided. Listings of the events will also be provided.

Abbreviations: HbA1c = hemoglobin A1c; Level 2/3= Level 2 and Level 3 composite.

Note: The yearly hypoglycemia rate during defined period is calculated by number of hypoglycemic events within the period/number of days patient at risk within the period *365.25. For rare events, 100 year rate will be provided. The hypoglycemia incidence during defined period indicates if the patient has at least 1 hypoglycemia event within the period (Yes/No).

Note: Group comparisons for hypoglycemia incidence and rate at baseline , baseline HbA1c is not included in the model.

Note: The hypoglycemia rate during the defined period is calculated by the number of hypoglycemia events within the period/number of days a participant is at risk within the period. The hypoglycemia incidence during the defined period indicates if the participant has at least 1 hypoglycemia event within the period (yes/no).

4.6.2.2.2. Hypoglycemic Events derived from CGM

Level 2, Level 2 ending with BG value >=70 mg/dL, and Level 1 or Level 2 hypoglycemic events collected from the sponsor provided CGM are defined in Appendix 6, Section 6.6.2. The analysis for incidence, event rate, and duration are described in the table below. All data are included in the analysis. Missing data will be handled as described in Appendix 6.

Endpoint	CGM Session*	Statistical Method
Event rate of Level 1 or Level 2 hypoglycemic events (events/participant/year): • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline HbA1c and baseline Level 1 or Level 2 hypoglycemia rate as covariates, log (exposure/365.25 days) as the offset in the model.
Event rate of Level 2 hypoglycemic events (events/participant/year): • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline HbA1c and baseline Level 2 hypoglycemia rate as covariates, log (exposure/365.25 days) as the offset in the model.

Endpoint	CGM Session*	Statistical Method
Event rate of Level 2 hypoglycemic events ending with >=70 mg/dL (events/participant/year): • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	Negative binomial regression with treatment, baseline HbA1c and baseline rate for Level 2 hypoglycemia ending with >=70 mg/dL as covariates, log (exposure/365.25 days) as the offset in the model.
Incidence of Level 1 or Level 2 hypoglycemic events: • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	Logistic regression with treatment, baseline HbA1c and baseline Level 1 or Level 2 hypoglycemia rate as covariates. The total number of episodes, number of episodes that includes Level 2 hypoglycemia episodes, and number of episodes that include glucose readings <54 mg/dL(3.0 mmol/L) but no Level 2 hypoglycemia episodes will be summarized for each CGM session.
Incidence of Level 2 hypoglycemic events: • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	Logistic regression with treatment, baseline HbA1c and baseline Level 2 hypoglycemia incidence as covariates.
Incidence of Level 2 hypoglycemic events ending with >=70 mg/dL: • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	Logistic regression with treatment, baseline HbA1c and baseline incidence for Level 2 hypoglycemia ending with >=70 mg/dL as covariates.

Endpoint	CGM Session*	Statistical Method
Duration of Level 1 or Level 2 hypoglycemic events: • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period	**For 0-4, 8-12, 22-26 weeks:** The MMRM model will include treatment, strata (country, routine use of personal CGM or FGM at randomization, baseline HbA1c stratum (<8.0%, ≥8.0%)), time and treatment-by-time interaction as fixed effects, and baseline duration as a covariate. **For 0-26 weeks:** ANCOVA model model will include treatment, strata (country, routine use of personal CGM or FGM at randomization, baseline HbA1c stratum (<8.0%, ≥8.0%)), and baseline duration. **For Baseline and post-treatment period:** ANOVA models only include treatment.
Duration of Level 2 hypoglycemic events: • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period.	Same models as for Level 1 or Level 2 hypoglycemic events. A listing will be provided for participants with Level 2 hypoglycemia episodes lasting >360 minutes
Duration of Level 2 hypoglycemic events ending with BG >=70 mg/dL: • 24-hour	Baseline, 0-4, 8-12, 22-26, 0-26 weeks, post-treatment period.	Same models as for Level 1 or Level 2 hypoglycemic events.

Abbreviations: HbA1c = hemoglobin A1c.

Note: The hypoglycemic rate during defined period is a yearly rate and calculated by number of hypoglycemic events within the period/number of days patient at risk within the period*365.25. The hypoglycemia incidence during defined period indicates if the patient has at least 1 hypoglycemic event within the period (Yes/No).

4.6.2.3. Laboratory and Adverse Event for Hepatic Safety

Hepatic labs include

- alanine aminotransferase (ALT)
- aspartate aminotransferase (AST)
- total bilirubin (TBL)
- direct bilirubin (DBL)

- serum alkaline phosphatase (ALP), and
- gamma-glutamyltransferase (GGT).

When criteria are met for hepatic evaluations, investigators will conduct close monitoring of hepatic symptoms and liver tests, perform a comprehensive evaluation for alternative causes of abnormal liver tests, and complete follow-up hepatic safety CRFs.

The table below lists summary TFLs for the analysis of hepatic laboratory data.

Analysis	Details
Abnormal postbaseline categories – hepatic safety parameters	ALT: The number and percentage of participants with a measurement greater than or equal to 1 time (1X), 3 times (3X), 5 times (5X), 10 times (10X), and 20 times (20X) the performing laboratory ULN during the postbaseline period will be summarized for all participants with a postbaseline value. AST: The number and percentage of participants with a measurement greater than or equal to 1 time (1X), 3 times (3X), 5 times (5X), 10 times (10X), and 20 times (20X) the performing laboratory ULN during the postbaseline period will be summarized for all participants with a postbaseline value. ALP: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) and 3 times (3X) the performing laboratory ULN during the postbaseline period will be summarized for all participants with a postbaseline value. TBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2X), 5 times (5X), and 8 times (8X) the performing laboratory ULN during the postbaseline period will be summarized for all participants with a postbaseline value. DBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) and 5 times (5X) the performing laboratory ULN during the postbaseline period will be summarized for all participants with a postbaseline value. GGT: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) the performing laboratory ULN during the postbaseline period will be summarized for all participants with a postbaseline value.

Analysis	Details
Treatment-emergent potentially drug-related hepatic disorders	Potentially drug-related hepatic disorders are defined using a custom query based on the following SMQs: • Broad and narrow terms in the Liver-related investigations, signs and symptoms SMQ (20000008) • Broad and narrow terms in the Cholestasis and jaundice of hepatic origin SMQ (20000009) • Broad and narrow terms in the Hepatitis non-infections SMQ (20000010) • Broad and narrow terms in the Hepatic failure, fibrosis and cirrhosis and other liver damage SMQ (20000013) • Narrow terms in the Liver-related coagulation and bleeding disturbances SMQ (20000015) These SMQs are a subset of the sub-SMQs comprising the full Hepatic Disorders SMQ. Only the sub-SMQs considered applicable to capturing potentially drug-related hepatic disorders are included. The percentage of study participants with at least one of any of the MedDRA preferred terms from any of the above SMQs will be summarized in addition to the percentages for each MedDRA preferred term.
Hepatocellular drug-induced liver injury screening plot (TBL vs ALT or AST)	Each participant's data is plotted based on their maximum postbaseline TBL (y-axis) and transaminase (ALT or AST, whichever is higher), regardless of the time between the 2 maximum values. Lines represent TBL and transaminase cutoffs of 2X ULN and 3X ULN, respectively. A potential Hy's law case is circled and is defined as having a maximum postbaseline TBL equal to or exceeding 2X ULN within 30 days after maximum postbaseline ALT or AST equal to or exceeding 3X ULN, without cholestasis (defined as ALP less than 2X ULN).
Hepatocellular drug-induced liver injury screening table	The percentages of study participants falling in each of the 3 relevant quadrants of the plot (right upper, left upper, right lower) will be summarized in a table.
Cholestatic drug-induced liver injury screening plot (TBL vs ALP)	Each participant's data is plotted based on their maximum postbaseline TBL (y-axis) and ALP (x-axis), regardless of the time between the 2 maximum values. Lines represent TBL and ALP cutoffs of 2X ULN and 3X ULN, respectively. A potential cholestatic liver injury case is circled and is defined as having a maximum postbaseline TBL equal to or exceeding 2X ULN within 30 days after maximum postbaseline ALP equal to or exceeding 3X ULN.
Cholestatic drug-induced liver injury screening table	The percentages of study participants falling in each of the 3 relevant quadrants of the plot (right upper, left upper, right lower) will be summarized in a table.
List of Participants with potential hepatocellular drug-induced liver injury	Includes participants falling in the right upper quadrant in the Hepatocellular Drug-Induced Liver Injury Screening plot. Variables to include are unique subject ID, age, sex, race, treatment, max AST, max ALT, max ALP, max TBL.

Analysis	Details
List of Participants with potential cholestatic drug-induced liver injury	Includes participants falling in the right upper quadrant in the Cholestatic Drug-Induced Liver Injury Screening plot. Variables to include are unique subject identifier, age, sex, race, treatment, max AST, max ALT, max ALP, max TBL.

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CRF = case report form; DBL = direct bilirubin; GGT = gamma-glutamyl transferase; MedDRA = Medical Dictionary for Regulatory Activities; SMQ = Standardized MedDRA Query; TBL = total bilirubin; ULN = upper limit of normal.

Planned and unplanned measurements will be included. The measurements do not need to be taken at the same blood draw.

4.6.2.4. Clinical Laboratory Evaluations

For the following selected laboratory measures:

- Liver enzyme tests: ALT, ALP, AST, GGT, DBL and Total Bilirubin
- Lipid measures: Triglycerides, total cholesterol, LDL-C, and HDL-C (results from fasting samples))

MMRM model (as described in Section 4.6) will be used for the observed values, change from baseline and percentage change from baseline, for which log-transformation will be applied. Geometric LS means will be provided. Analyses will be provided in both international units (SI) and conventional units (CN) if they are different.

Box plots with descriptive statistics for the observed values and change from baseline will be provided by treatment group and visit.

For other laboratory measures, descriptive summaries will be provided for the observed values and change from baseline by treatment group and visit.

The percentages of participants with elevated or low values meeting specified levels (see Appendix 10, Section 6.10) at any time post-baseline (including scheduled and unscheduled measurements) will be summarized and compared between treatment groups using risk difference and 95% confidence interval.

A listing of abnormal laboratory analytes collected quantitatively (high or low during postbaseline period using Level 2 definitions in Appendix 10) and qualitatively (abnormal during postbaseline period) will be provided, including participant identification, treatment group, laboratory sample collection day (that is, days from start of study drug), analyte name, abnormal result, and reference low or high limits (Level 2 cut-off value).

Scatter plots of maximum-by-maximum measurements and minimum-by-minimum measurements will not be created a-priori. They may be created if warranted after review of the planned tables and figures, using Figures 6.3 and 6.4 from the Analysis and Displays for Labs white paper (PHUSE 2022) as the model. ADaM datasets will include variables to enable the creation of scatter plots for use in either an interactive tool or for ad-hoc figures.

4.6.2.5. Vital Signs and Physical Characteristics

The planned summaries are provided in the Table below. The measurements analyzed for vital signs and physical characteristics include systolic BP, diastolic BP, pulse, weight, and BMI.

Analysis Type	Analysis Details
Observed values change by visit	• Includes all participants in the safety population who have both a baseline and at least 1 postbaseline observation • MMRM model (as described in Section 4.6) will be used. See also: Table 6.2 from the Analyses and Displays for Labs white paper (PHUSE 2022)
Summary by category	• Definitions provided in Tables 31-33 from FDA's September 2022 Standard Safety Tables and Figures document will be used for the numerator. o Systolic BP (mm Hg): - Low: Level 1: <90 - High: Level 1: ≥90, Level 2: ≥120, Level 3: ≥140, Level 4: ≥160, Level 5: ≥180 o Diastolic BP (mm Hg): - Low: Level 1 <60 - High: Level 1: >=60, Level 2: >=90, Level 3: >=110, Level 4: ≥120 • Includes participants with at least one postbaseline measurement. • Statistical comparisons (using methods described in Section 4.6) will be included.
Participants meeting CTC grade changes in weight	For weight, cutoffs informed by CTCAE version 5 (Grades 1-3) will be used: o (Loss) decrease: Level 1: ≥5%, Level 2: ≥10%, Level 3: ≥20% o (Gain) increase: Level 1: ≥5%, Level 2: ≥10%, Level 3: ≥20% Includes participants with both a baseline and at least 1 postbaseline observation. • Statistical comparisons (using methods described in Section 4.6) will be included.

Scatter plots to support vital sign evaluations

Scatter plots of maximum-by-maximum measurements and minimum-by-minimum measurements will not be created a-priori. They may be created if warranted after review of the planned tables and figures, using Figures 6.3 and 6.4 from the Analysis and Displays for Labs

white paper (PHUSE 2022) as the model. ADaM datasets will include variables to enable the creation of scatter plots for use in either an interactive tool or for ad-hoc figures.

4.6.3. Device Product Complaints

A summary of all product complaints, inclusive of device product complaints that lead to an AE and/or SAE will be included by category. Additional summaries may be provided as deemed appropriate.

4.7. Other Analyses

4.7.1. Immunogenicity

A participant is evaluable for TE ADA if the participant has a non-missing baseline ADA result, and at least 1 non-missing postbaseline ADA result.

A participant who is evaluable for TE ADA is TE ADA+ if either of the following holds:

- Treatment-induced ADA: the participant has baseline status of ADA Not Present and at least 1 postbaseline status of ADA Present with titer ≥ 1:40, which is 2 × minimum required dilution (MRD) of the ADA assay (MRD=1:20).
- Treatment-boosted ADA: the participant has baseline status of ADA Present and at least 1 postbaseline status of ADA Present with the titer being ≥2 dilutions (4-fold) of the baseline titer. That is, the participant has baseline (B) status of ADA Present, with titer 1:B, and at least 1 postbaseline (P) status of ADA Present, with titer 1:P and P/B≥4.

All analyses will be based on all evaluable TE ADA participants. The baseline ADA status will be summarized by treatment group. The number and percentage of participants who are TE ADA+ will be summarized by treatment group. The summary will include the number and percentage of participants with treatment-induced ADA and treatment boosted ADA. A summary of titer values will be provided for participants who are TE ADA positive. This analysis will be performed for the following periods:

- The treatment period up to treatment discontinuation
- The entire postbaseline period, including safety follow-up

A number of additional analyses of the immunogenicity data from this study will be presented in an integrated summary document, alongside data from other studies. The analyses to be performed are described in the PSAP.

4.7.2. Subgroup Analyses

The interaction effects will be evaluated using a significance level of 0.05, unadjusted. Subgroup analyses will be conducted as defined in this section. Additional subgroup analysis may also be performed as appropriate.

4.7.2.1. Subgroups Analysis for HbA1c:

The subgroups for analyzing HbA1c and change in HbA1c from baseline to Week 26 will be defined as:

- Baseline HbA1c stratum (<8.0% and ≥8.0%)
- Routine use of personal CGM or FGM at randomization (Yes or No)
- Region
  - US and non-US
  - North America, South America, Europe, Asia
- Type of pre-study basal insulin (typical once-daily basal insulin and U-300/NPH BID regimens)
- Prandial insulin dosing plan (carbohydrate counting: Yes or No)
- Use of correction factor (Yes or No)
- Age group: <65 years and ≥65 years
- Sex (Male and Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported)
- Race (White, Asian, American Indian or Alaska Native, Black or African American)
- Glucagon-like peptide 1 (GLP-1) usage at baseline (Yes and No)
- Estimated Glomerular Filtration Rate (eGFR) at baseline (~~<30, ≥30 to~~ <60, ≥60 to <90 and ≥90 mL/min/1.73 m^2), and
- Duration of diabetes (< median and ≥ median).

Analyses for HbA1c and its change will be performed within each subgroup using the same MMRM model for efficacy estimand described for the primary analysis in Section 4.3.2. In addition, the interaction effects will be assessed using the model including the same fixed effects given for the primary analysis model plus factors of subgroup, 2-way interaction of subgroup and treatment, 2-way interaction of subgroup and visit, and 3-way interaction of treatment, visit and the subgroup.

The ANCOVA analysis for HbA1c for the treatment regimen estimand as described in Section 4.3.2 will be performed within each subgroup. The statistical inference will be based on the multiple imputation framework by Rubin (1987). The p-value for treatment by subgroup interaction will be calculated using a z-test for variables with 2 subgroups and chi-square test for variables with 3 subgroups based on estimated treatment differences within each subgroup. (See details in Appendix 12 [Section 6.12]).

4.7.2.2. Subgroup Analysis for Participant-Reported Hypoglycemic Events

The subgroups, for analyzing documented Level 2/3 hypoglycemia, non-nocturnal and nocturnal hypoglycemia rates during 0-26 weeks, will be defined as:

- Baseline HbA1c stratum (<8.0% and ≥8.0%)
- Routine use of personal CGM or FGM at randomization (Yes or No)
- Region
  - US and non-US
  - North America, South America, Europe, Asia
- Type of pre-study basal insulin (typical once-daily basal insulin and U-300/NPH BID regimens)
- Age group: <65 years and ≥65 years
- Estimated Glomerular Filtration Rate (eGFR) at baseline : (<60, ≥60 to <90 and ≥90 mL/min/1.73 m^2)

- Glucagon-like peptide 1 (GLP-1) usage at baseline (Yes and No)
- Number of prandial insulin injection/day at baseline (2 and >2)

The hypoglycemia rates will be analyzed using a negative binomial regression including the same independent variables for hypoglycemia event analyses (see Section 4.6.2.2) plus factors of subgroup, 2-way interaction of subgroup and treatment.

4.8. Interim Analyses

4.8.1. Data Monitoring Committee

An independent external data monitoring committee (DMC) will be responsible for reviewing unblinded data during the study. The committee will include 4 clinicians and 1 statistician who are independent experts not involved in the study. The DMC will review unblinded safety data to ensure the safety of study participants and some efficacy data to confirm a reasonable risk-benefit profile. A subset of analyses described above in Sections 4.3, 4.4, 4.5, and 4.6 will be provided for the DMC review. The external Statistical Analysis Center statistician/analyst will generate the unblinded reports and confidentially distribute the unblinded reports to DMC members. Study team will remain blinded to study treatment until the planned unblinding occurs. The DMC will be conducted to maintain study integrity. Details of DMC is included in the DMC charter.

4.9. Changes to Protocol-Planned Analyses

The changes to the analyses described in the protocol are listed in the Version History section.

5. Sample Size Determination

Approximately 670 participants will be randomly assigned to LY3209590 and insulin glargine in a 1:1 ratio. With the assumption of 15% dropout at Week 26, approximately 284 participants on LY3209590 and 284 participants on insulin glargine will complete 26 weeks of treatment.

The primary objective of this study is to test the hypothesis that LY3209590 is noninferior to insulin glargine on glycemic control as measured by change from baseline at Week 26 in HbA1c in participants with T2D currently on basal insulin and at least 2 injections per day of prandial insulin.

Assuming an NIM of 0.4%, no true difference between treatment groups, and a SD of 1.1%, 568 completers (284 on each of LY3209590 and insulin glargine) will provide greater than 99% statistical power to show noninferiority of LY3209590 compared to insulin glargine using the upper limit of a 2-sided 95% CI of (LY3209590 – insulin glargine). This sample size also has at least 90% statistical power to show noninferiority of LY3209590 compared to insulin glargine using a 0.3% NIM at Week 26.

The 568 completers will provide 90% statistical power to demonstrate the superiority (LY3209590 vs insulin glargine) of change in HbA1c from baseline at Week 26 (assuming a SD of 1.1% and true mean difference is -0.3%) using t-test at alpha of 0.05.

The 568 completers will provide 90% statistical power to demonstrate the superiority (LY3209590 vs insulin glargine) on percentage of participants achieving HbA1c <7% at Week 26 without nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26 using a Fisher's Exact test assuming 28.7% of insulin glargine patients achieve this target and 42% of LY3209590 patients achieve this target.

The 568 participants will provide approximately 83% statistical power to show the superiority of the event rate of clinically significant nocturnal hypoglycemia (<54 mg/dL [3.0 mmol/L] or severe) during treatment phase up to Week 26 (assuming event rate of 0.55 [SD=2.6] and 1.1 [SD=2.6] events per participant per year for LY3209590 and insulin glargine, respectively, representing a 50% decrease in event rate for LY3209590 over insulin glargine) using a negative binomial distribution at alpha = 0.05.

6. Supporting Documentation

6.1. Appendix 1: Demographic and Baseline Characteristics

Demographic and baseline characteristics including but not limited to age (years), age groups (<65, ≥65 and <75, ≥75 and <85, ≥85 years), sex, ethnicity, race, country, region, height, weight (kg), body mass index (BMI: kg/m2), BMI groups (<25, ≥25 and <30, ≥30 and <35, ≥35 kg/m^2), eGFR groups (<30, ≥30 and <60, ≥60 and <90, ≥90 mL/min/1.73 m^2), duration of diabetes (years), HbA1c at screening, HbA1c stratum at screening (<8.0% and ≥8.0%), baseline HbA1c, baseline HbA1c stratum (<8.0% and ≥8.0%), baseline fasting serum glucose (mmol/L and mg/dL), type and category of pre-study basal insulin type of pre-study prandial insulin (type, category and frequency), prandial insulin dosing plan (carbohydrate counting: Yes or No), use of prandial insulin correction factor, GLP-1 RA use, baseline noninsulin antihyperglycemic medications (type and number), baseline FBG<120 mg/dL (based on the study-provided glucometer), and routine use of personal CGM or FGM at randomization will be summarized by treatment group using the mITT and Randomized Population (if different from the mITT).

Continuous measures will be summarized using descriptive statistics and treatment difference will be analyzed using the analysis of variance. Categorical measures will be summarized using sample size, frequency, and percentage and treatment difference will be analyzed using Chi-squared test.

The by-participant listing of demographic and baseline characteristics will be provided for Randomized Population.

Historical conditions are conditions that end prior to inform consent and preexisting conditions are conditions that are still ongoing at inform consent. The number and percentage of participants with historical conditions will be summarized by treatment group using MedDRA PT using the mITT and Randomized Population (if different from the mITT). Events will be ordered by decreasing frequency. Similar summary will also be provided for preexisting conditions.

6.2. Appendix 2: Treatment Compliance

Treatment compliance will be summarized using mITT population excluding inadvertently enrolled participants .

The study protocol provides dosing algorithm for both study treatments (LY3209590 and insulin glargine). The investigator will calculate algorithm recommended dose based on the participant fasting blood glucose and hypoglycemia occurrence reported in e-diary. If the investigator does not agree with the algorithm recommended dose, the investigator will prescribe another dose for the participant and choose the reason for not following the algorithm recommended dose from a pick-list. The number and percentage of investigator prescribed doses different from algorithm recommended dose will be provided to evaluate the adherence of algorithm dose. The reason of not following the algorithm recommended dose will also be summarized.

Study personnel or the participant will administer the first study dose of LY3209590 at the site. Insulin glargine can be taken at the site or after the visit depending upon whether the study visit timing coincides with the participant's usual time to administer basal insulin. The subsequent doses are self-administered. To access the adherence of treatment administration, the number and

percentage of investigator prescribed doses that are not equal to participant administered dose will be provided.

6.3. Appendix 3: 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 adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Non-Serious Adverse Events are summarized: by treatment group, by MedDRA PT.
- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE 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
    - the number of events experienced.
- For each Serious AE, these additional terms are provided for European Union Drug Regulating Authorities Clinical Trials (EudraCT):
    - the total number of occurrences causally related to treatment
    - the total number of deaths
    - the total number of deaths causally related to treatment.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may be excluded if a 5% threshold is chosen. Allowable thresholds include 0% (all events), 1%, 2%, 3%, 4%, and 5%.
- AE reporting is consistent with other document disclosures for example, the clinical study report (CSR), manuscripts, and so forth.
- Demographic table including the following age ranges required by EudraCT: in utero, preterm newborn infants (gestational age <37 weeks), newborns (0-27 days), infants and toddlers (28 days – 23 months, children (2-11 years), adolescents (12-27 years), adults (18-64 years), 65-85 years, and 85 years and over.

6.4. Appendix 4: Concomitant Therapy

Concomitant therapy is defined as the therapy that starts before, on, or after the first day of study treatment and before the last dose date in the treatment period, and continues into the treatment period, that is, either no end date (the therapy is ongoing) or an end date on or after the first day of study treatment.

The number and percentages of participants who take concomitant medication will be summarized by treatment using PTs nested within Anatomical Therapeutic Chemical (ATC) codes. The concomitant medications will be ordered by decreasing frequency of LY3209590 within each ATC.

6.5. Appendix 5: Protocol Deviations

IPDs are the deviations from the study protocol that may compromise the data integrity and patients' safety. The IPD category and details of IPD identification are provided in the trial issue management plan.

The number and percentage of participants with any reported IPDs will be summarized by treatment group and IPD category. The IPDs identified by site monitoring and clinical database will be integrated. If the IPD is identified by both methods, only the site monitoring IPD will be presented.

6.6. Appendix 6: Derivation of CGM Variables

Blinded CGM will be worn by participants at the following designated weeks:

- Two weeks in screening/lead-in period with the data download at Week 0 (Visit 3)
- Four weeks after the initiation of study treatment with the data download at Week 4 (Visit 7)
- Four weeks near the end of weekly titration period with the data download at Week 12 (Visit 15)
- Four weeks near the primary endpoint with the data download at Week 26 (Visit 23)
- Four weeks after the end of study treatment period with the data download at Week 31 (Visit 802).

Linear interpolation will be used to impute the missing glucose readings during intervals >7 minutes and ≤ 15 minutes. Missing data will be imputed at a 5-minute interval using the BG value before and the one after the interval. For example, the BG reading before the interval is 50 mg/dL at time 100 min, and the BG reading after interval is 70 at time 115 min, then the missing records will be imputed as:

- at time 105 min, BG reading=50+5*(70-50)/(115-100)=56.7,

- at time 110 min, BG reading=50+10*(70-50)/(115-100)=63.3

Glucose reading intervals >15 minutes will be treated as missing data and not be counted in the length of the analysis periods.

For the primary analysis, CGM data must meet minimal completeness criteria for the period and the session in order to be included as a valid time period and session.

As shown in the table below, all the CGM derivations are based on the data from valid CGM periods of a day and sessions unless otherwise specified.

Minimum Data Available to Define a Valid Period of the Day and CGM Session

Period	Definition	Minimum Valid CGM Time Period for a Session
Valid CGM Day (00:00-23:59)	≥70% of expected values available (≥202 of 288 values)	3 valid CGM days define a CGM session for 24-hour

Valid CGM Nighttime (00:00-05:59)	≥70% of expected values available (≥50 of 72 values)	3 valid CGM days define a CGM session for Nighttime
Valid CGM Daytime (06:00-23:59)	≥70% of expected values available 151 of 216 values)	3 valid CGM days define a CGM session for Daytime

Sensitivity analysis will be done for selected parameters including all CGM data.

The ambulatory glucose profile during the 24-hour period will be generated with interquartile ranges, at treatment-group level by CGM session, based upon the observed and imputed CGM measures.

6.6.1. Glucose in Target Ranges, Hypoglycemia or Hyperglycemia

The following variables of time in range, hypoglycemia, hyperglycemia during each CGM session will be derived:

- Percentage and Duration (in minutes) of time per day where glucose values are within a hypoglycemic range (defined as <54 mg/dL [3.0 mmol/L]) during the nighttime period (defined as midnight to 0600 hours), the daytime period (defined as 0600 hours to 2400 hours) and a 24-hour period.
- Percentage and Duration (in minutes) of time per day where glucose values are within a hypoglycemic range (defined as <70 mg/dL [3.9 mmol/L] and ≥54 mg/dL [3.0 mmol/L]) during the nighttime period, the daytime period, and a 24-hour period.
- Percentage and Duration (in minutes) of time per day where glucose values are within a hypoglycemic range (defined as <70 mg/dL [3.9 mmol/L]) during the nighttime period, the daytime period and a 24-hour period.
- Percentage and Duration (in minutes) of time per day glucose values are within a hyperglycemic range (defined as >180 mg/dL [10.0 mmol/L] and ≤ 250 mg/dL [13.9 mmol/L]) during the nighttime period, the daytime period and a 24-hour period.
- Percentage and Duration (in minutes) of time per day glucose values are within a hyperglycemic range (defined as >250 mg/dL [13.9 mmol/L]) during the nighttime period, the daytime period and a 24-hour period.
- Percentage and Duration (in minutes) of time per day glucose values are within a hyperglycemic range (defined as >180 mg/dL [10.0 mmol/L]) during the nighttime period, the daytime period and a 24-hour period.
- Percentage and Duration (in minutes) of time per day glucose values are within a glucose range (defined as between 70 mg/dL and 180 mg/dL [3.9 and 10.0 mmol/L]) inclusive during the nighttime period, the daytime period and a 24-hour period.
- Percentage and Duration (in minutes) of time per day glucose values are within a glucose range (defined as between 70 mg/dL and 140 mg/dL [3.9 and 7.8 mmol/L] inclusive) during the nighttime period, the daytime period and a 24-hour period.

The percentage of time within a glucose range (target, hypoglycemia, or hyperglycemia ranges) will be calculated as the number of observations within the specified range divided by the

number of observations in the time interval (for example, 24-hour period). The average percentage of time among valid CGM days for the corresponding time interval during each CGM session will be used in the analysis.

The duration (in minutes) within the glucose range will then be calculated as the average percentage of time within the glucose range times the length of the period (24-hour, 18-hour, and 6-hour, for the periods of 24-hour, daytime or nighttime, respectively).

According to the guidance (Battelino et al. 2019), the following CGM targets of glycemic control will also be derived during a 24-hour period.

- The percentage of time within a normal glucose range (defined as between 70 mg/dL and 180 mg/dL [3.9 and 10.0 mmol/L] inclusive) >70%
- The percentage of time within a hypoglycemia range (defined as <70 mg/dL [3.9 mmol/L]) <4%
- The percentage of time within a hypoglycemia range (defined as <54 mg/dL [3.0 mmol/L]) <1%
- The percentage of time within a hyperglycemic range (defined as >180 mg/dL [10.0 mmol/L]) <25%
- The percentage of time within a hyperglycemic range (defined as >250 mg/dL [13.9 mmol/L]) <5%

In addition, according to the guidance in 2023 (Battelino et al. 2023), 2 composite endpoints will be derived during a 24-hour period:

- >70% time in range 70 -180 mg/dL (3.9 - 10.0 mmol/L inclusive) and <4% time below range <70 mg/dL (<3.9 mmol/L)
- >70% time in range 70 -180 mg/dL (3.9 - 10.0 mmol/L inclusive) and <1% time below range <54 mg/dL (<3.0 mmol/L)

The daily duration (in minutes) of time within a glucose range (defined as <54 mg/dL [3.0 mmol/L]; <70 mg/dL [3.9 mmol/L] and ≥54 mg/dL [3.0 mmol/L]; <70 mg/dL [3.9 mmol/L]; >=70 mg/dL[3.9 mmol/L] and <=140 mg/dL [7.8 mmol/L]; >=70 mg/dL[3.9 mmol/L] and <=180 mg/dL [10.0 mmol/L]; >180 mg/dL[10.0 mmol/L] and <=250 mg/dL[13.9 mmol/L]; >250 mg/dL[13.9 mmol/L]) will be merged with the dose administration data to get the summary of daily time in each glucose range since dose administration in each CGM session in the treatment period. For LY3209590, days relative to dose administration will be derived as 0 (dosing day), and then 1, 2, …, up to 6 days after most recent dose administration but before the next dose administration. The average daily time in each glucose range for the given day (0 to 6) relative to the dose administration among valid CGM days (with at least 70% of the data of each day) during each CGM session will be used in the analysis.

6.6.2. Hypoglycemic Episode

According to the International Consensus Statement (Battelino et al, 2023), the CGM-determined hypoglycemic episodes for Level 1 or 2 and Level 2 are defined as below.

- Level 2 hypoglycemia (BGs<54 mg/dL)
    - Starting time: is the time of the 1st BG of BGs<54 mg/dL for >=15 consecutive min (regardless of number of BG readings)
    - Ending time: is the time of the last BG of the BGs>=54 mg/dL for 15 consecutive min (regardless of number of BG readings)
    - Duration: Time of the last BG <54 mg/dL – the time of the 1st BG<54 mg/dL
- Level 1 or 2 hypoglycemia (BGs<70 mg/dL)
    - Starting time: is the time of the 1st BG of BGs<70 mg/dL for >=15 consecutive min (regardless of number of BG readings)
    - Ending time: is the time of the last BG of the BGs>=70 mg/dL for 15 consecutive min (regardless of number of BG readings)
    - Duration: Time of the last BG <70 mg/dL - the time of the 1st BG<70 mg/dL
- Level 2 hypoglycemia ending with BG >=70 mg/dL (start with BG<54mg/dL, end with BG>=70 mg/dL)
    - Starting time: is the time of the 1st BG of BGs<54 mg/dL for >=15 consecutive min (regardless of number of BG readings)
    - Ending time: is the time of the BG that ensures the BGs>=70 mg/dL for 15 consecutive min (regardless of number of BG readings)
    - Duration: Time of the last BG <70 mg/dL – the time of the 1st BG<54 mg/dL

If the truncated time interval is >7 and <=15 minutes, use linear interpolation to impute the missing data at 5-minute intervals as described in Section 6.6.

If the truncated time interval is >15 minutes

- at any time, which make the starting time undeterminable, then do not count this time interval and no episodes start.
- If it is after an episode started, which makes the ending time undeterminable, then the episode ends at the starting time of the truncation.
- If an episode started and continued until the end of a CGM session, then the episode ends at the end of CGM session.

The average duration, incidence, event rate of the Level 1 or 2 and Level 2 hypoglycemia episodes will be analyzed for each CGM session as planned in Section 4.6.2.2.2. The average duration will be calculated by dividing the sum of the duration of individual episodes during the given CGM session by the number of episodes and used in the analysis. The hypoglycemic event

rate (events/participant/year) will be calculated by dividing the number episodes by the number of CGM days × 365.25 days.

The duration (minute) for a day= the time of last CGM value - the time of the 1st CGM value - the sum of intervals that are >15 minutes within each CGM day.

6.6.3. Mean Glucose and Glucose Management Indicator

The average glucose within a time period (a 24-hour period, daytime or nighttime) for each valid CGM day will be calculated first and then the average of daily average for a CGM session will be used as the mean glucose of the CGM session in the analysis.

The glucose management indicator (GMI) is a new parameter estimating A1c from CGM. The GMI is based on the above mean glucose (24-hour period) by CGM using the below formula (Bergenstal et al. 2018):

$$\text{GMI}(\%) = 3.31 + 0.02392 \times \text{mean glucose (mg/dL)}$$

6.6.4. Glycemic Variability

Glycemic variability will be derived using the notation below:

i represents a time point within a time period (a 24-hour period, daytime or nighttime).

n represents the number of time points within the time period.

k represents a valid CGM day within a visit.

m represents the number of valid CGM days in the specific time period at a visit.

$BG_{k,i}$ represents the glucose value at time point *i* on day *k* unless otherwise specified.

Sections 6.6.4.1 and 6.6.4.2 provide the derivation method for variables assessing within-day and between-day glucose variability based on CGM readings.

6.6.4.1. Within-Day Variability

For variables assessing within-day variability, first determine the variability within each valid CGM day, then average across days within a CGM session.

Within-day glucose SD (Rodbard 2009):

$$SD = \frac{1}{m}\sum_{k=1}^{m} \text{SD}_k = \frac{1}{m}\sum_{k=1}^{m} \sqrt{\frac{\sum_{i=1}^{n}(BG_{k,i} - \left\{\frac{\sum_{i=1}^{n} BG_{k,i}}{n}\right\})^2}{n-1}}$$

Within-day glucose CV (Clarke and Kovatchev 2009):

$$CV = \frac{1}{m}\sum_{k=1}^{m} CV_k = \frac{1}{m}\sum_{k=1}^{m} \frac{SD_k}{\left(\frac{\sum_{i=1}^{n} BG_{k,i}}{n}\right)} \times 100$$

The low blood glucose index (LBGI), high blood glucose index (HBGI), and blood glucose risk index (BGRI) will be calculated using the following standard formulas (Kovatchev et al. 2006).

The LBGI, HBGI, and BGRI will be derived for each valid CGM day of a visit and then average across days within a CGM session. The calculations of LBGI, HBGI, and BGRI take the following steps:

1. For each blood glucose (BG [mg/dL]) at the i^{th} time point, compute the following:

$$f(BG_i) = 1.509 \times [(\ln(BG_i))^{1.084} - 5.381]$$

2. Compute BG risk for each reading

$$rl(BGi) = 10 \times f(BG_i), \text{ if } f(BG_i) < 0; \text{ otherwise } rl(BG_i) = 0$$
$$rh(BGi) = 10 \times f(BG_i), \text{ if } f(BG_i) > 0; \text{ otherwise } rh(BG_i) = 0$$

3. Compute LBGI and HBGI

$$LBGI = \frac{1}{n}\sum_{i=1}^{n} rl(BG_i)$$

$$HBGI = \frac{1}{n}\sum_{i=1}^{n} rh(BG_i)$$

4. Compute BGRI

$$BGRI = LBGI + HBGI$$

6.6.4.2. Between-Day Variability

For variables assessing between-day variability, first determine the variability for each time point across days within a CGM session then average across all time points.

Between-day glucose SD (Rodbard 2009):

$$SD = \frac{1}{n}\sum_{i=1}^{n} SD_i = \frac{1}{n}\sum_{i=1}^{n} \sqrt{\frac{\sum_{k=1}^{m} (BG_{k,i} - \left\{\frac{\sum_{k=1}^{m} BG_{k,i}}{m}\right\})^2}{m-1}}$$

Between-day glucose CV (Kovatchev et al. 2009):

$$CV = \frac{1}{n}\sum_{i=1}^{n} CV_i = \frac{1}{n}\sum_{i=1}^{n} \frac{SD_i}{\left(\frac{\sum_{k=1}^{m} BG_{k,i}}{m}\right)} \times 100$$

Mean of daily differences (MODD): this parameter is calculated as the mean of absolute differences between glucose values at corresponding time points of consecutive days.

$$MODD = \frac{1}{m-1}\sum_{k=1}^{m-1} \frac{\sum_{i=1}^{n} \left|BG_{k+1,i} - BG_{k,i}\right|}{n}$$

6.7. Appendix 7: MedDRA PT for Diabetic Retinopathy or Maculopathy

The following PT will be used to identify TEAEs of diabetic retinopathy or maculopathy (see Section 4.6.2.1.7):

- Amaurosis
- Amaurosis fugax
- Arteriosclerotic retinopathy
- Blindness
- Blindness transient
- Blindness unilateral
- Choroidal neovascularisation
- Cystoid macular oedema
- Detachment of macular retinal pigment epithelium
- Detachment of retinal pigment epithelium
- Diabetic blindness
- Diabetic eye disease
- Diabetic retinal oedema
- Diabetic retinopathy
- Diabetic uveitis
- Diplopia
- Exudative retinopathy
- Eye laser surgery
- Fundoscopy
- Fundoscopy abnormal
- Intra-ocular injection
- Macular detachment
- Macular oedema

- Maculopathy
- Noninfective chorioretinitis
- Noninfective retinitis
- Phacotrabeculectomy
- Retinal aneurysm
- Retinal arteriovenous malformation
- Retinal artery embolism
- Retinal artery occlusion
- Retinal artery stenosis
- Retinal collateral vessels
- Retinal cryoablation
- Retinal detachment
- Retinal exudates
- Retinal haemorrhage
- Retinal laser coagulation
- Retinal neovascularisation
- Retinal oedema
- Retinal operation
- Retinal thickening
- Retinal vascular disorder
- Retinal vascular occlusion
- Retinal vein occlusion
- Retinitis
- Retinopathy
- Retinopathy haemorrhagic
- Retinopathy hypertensive
- Retinopathy hyperviscosity
- Retinopathy proliferative
- Scintillating scotoma
- Sudden visual loss
- Venous stasis retinopathy
- Vision blurred
- Visual acuity reduced
- Visual acuity reduced transiently
- Visual impairment
- Vitrectomy

6.8. Appendix 8: MedDRA PT for Peripheral Edema

The analysis of peripheral edema (Section 4.6.2.1.8) will be based on the TEAEs in the following terms:

- Acute pulmonary oedema
- Ascites
- Brain oedema
- Bronchial oedema
- Capillary leak syndrome
- Cerebral oedema management
- Compression garment application
- Cytotoxic oedema
- Effusion
- Fluid retention
- Gastrointestinal oedema
- Generalised oedema
- Gravitational oedema
- Hydraemia
- Hypervolaemia
- Hypoosmolar state
- Lipoedema
- Lymphoedema
- Negative pressure pulmonary oedema
- Non-cardiogenic pulmonary oedema
- Non-pitting oedema
- Oedema
- Oedema blister
- Oedema due to cardiac disease
- Oedema due to hepatic disease
- Oedema due to renal disease
- Oedema mucosal
- Oedema peripheral
- Pelvic fluid collection
- Pericardial effusion
- Perinephric collection
- Perinephric oedema
- Peripheral swelling
- Pleural effusion
- Pulmonary oedema
- Retroperitoneal effusion
- Retroperitoneal oedema
- Skin oedema
- Skin swelling
- Subdural effusion
- Swelling

- Visceral oedema

6.9. Appendix 9: Definition for Persistent-Recurrent Hypoglycemia by Programming

A P-R hypoglycemia based on programming search in the e-diary database for hypoglycemic events that meet prespecified criteria is defined as a set of hypoglycemic episodes that:

a) starts with the occurrence of a level 3, or level 2 hypoglycemic episode (<54mg/dL [3.0 mmol/L]) and is followed by more episodes of hypoglycemia (<70mg/dL [3.9 mmol/L]), within the day of the initial episode,

AND

b) is followed by at least 1 episode of hypoglycemia (<70mg/dL [3.9 mmol/L]) per day, in the next 2 days or longer, and that ends when no hypoglycemia episode occurs for at least 2 days.

An example of a set of hypoglycemic episodes meeting prespecified criteria of a P-R hypoglycemia event is illustrated in the figure below.

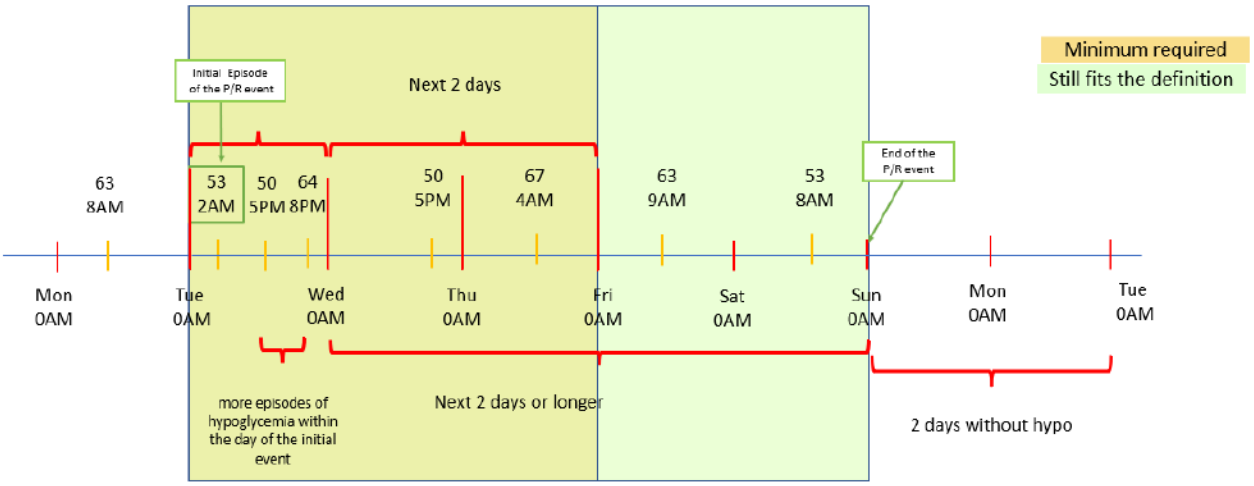


Figure BDCV 6.9. Illustration of persistent-recurrent hypoglycemia by programming

.

6.10. Appendix 10: Abnormality Level Criteria for Chemistry and Hematology Laboratory Results

Parameter	Level 1	Level 2	Level 3
General Chemistry			
Sodium, low (mEq/L)	<132	<130	<125
Sodium, high (mEq/L)	>150	>155	>160
Potassium, low (mEq/L)	<3.6	<3.4	<3.0
Potassium, high (mEq/L)	>5.5	>6	>6.5
Chloride, low (mEq/L)	<95	<88	<80
Chloride, high (mEq/L)	>108	>112	>115
Bicarbonate, low (mEq/L)	<20	<18	<15
Bicarbonate, high (mEq/L)	N/A	N/A	>30
Blood urea nitrogen, high (mg/dL)	>23	>27	>31
Calcium, low (mg/dL)	<8.4	<8.0	<7.5
Calcium, high (mg/dL)	>10.5	>11.0	>12.0
Phosphate, low (mg/dL)	<2.5	<2.0	<1.4
Protein (total), low (g/dL)	<6.0	<5.4	<5.0
Albumin, low (g/dL)	<3.1	<2.5	<2.0
Uric Acid (urate), high (mg/dL)	>7.0	NA	NA
Kidney Function			
Creatinine, increase (mg/dL)	≥1.5 x baseline	≥2.0 x baseline	≥3.0 x baseline
eGFR, decrease (ml/min/1.73m$_2$)	≥25% decrease	≥50% decrease	≥75% decrease
Lipids			
Cholesterol (total), high (mg/dL)	>200	>240	>300
HDL, low (mg/dL), males	<40	<30	<20
HDL, low (mg/dL), females	<50	<40	<20
LDL, high (mg/dL)	>130	>160	>190
Triglycerides, high (mg/dL)	>150	>300	>500
Hematology			
Complete Blood Count			
WBC, low (cells/µL)	<3500	<3000	<1000
WBC, high (cells/µL)	>10,800	>13,000	>15,000
Hemoglobin, decrease (g/dL)	N/A	>1.5 dec. from baseline	>2 dec. from baseline
Hemoglobin, increase (g/dL)	N/A	>2 inc. from baseline	>3 inc. from baseline
Platelets, low (cells/µL)	<140,000	<125,000	<100,000
Hemoglobin, low (g/dL), male	12.5-13.5	<12.5	<10.5
Hemoglobin, low (g/dL), female	11.0 – 12.0	<11	<9.5
WBC Differential			
Lymphocytes, low (cells/µL)	<1000	<750	<500
Lymphocytes, high (cells/µL)	>4000	>10000	>20000
Neutrophils, low (cells/µL)	<2000	<1000	<500

Eosinophils, high (cells/µL)	>650	>1500	>5000
Coagulation Studies			
prothrombin time, increase (sec)	>1.1 x ULN	>1.3 x ULN	>1.5 x ULN

Note: For Liver enzymes, Lilly defined categories will be used.

6.11. Appendix 11: Empirical Estimation of Relative Event Rate

Traditionally, Poisson distribution has been assumed to draw inference for the rate of rare events. When the event is rare and the sample size is large, it is known that the overall number of events is approximately from Poisson distribution. However, for some not very rare events such as severe hypoglycemic events in T2D patients, the total number of events may not be distributed from Poisson and may be over-dispersed. Assuming Poisson distribution may significantly underestimate the variance, and therefore may reduce the overage probability and inflate the Type-I error. An empirical method in estimating the variance of the relative event rate without assuming any distribution on the number of events will be provided in this appendix.

Let X_{ij} denote the count response variable for patient j in treatment group i. Let $Y_i = \sum_j X_{ij}$ be the total number of events for treatment group i, and T_i denote the exposure for treatment group i. Let $i = 0$ for the control group and $i = 1$ for the experimental treatment group. The event rate for treatment group i can be calculated as

$$\hat{r}_i = \frac{Y_i}{T_i}$$

The empirical variance of $\hat{r}_i$ is

$$\widehat{Var}(\hat{r}_i) = T_i^{-2}\widehat{Var}(Y_i) = T_i^{-2} n_i S_i^2,$$

where S_i^2 is the variance of X_{ij} for treatment group i. Using the delta-method, the variance of $\log(\hat{r}_i)$ can be estimated as

$$\widehat{Var}(\log(\hat{r}_i)) = Y_i^{-2} n_i S_i^2$$

The relative rate of the experimental treatment versus the control treatment is estimated as

$$\hat{\lambda} = \frac{\hat{r}_1}{\hat{r}_0}$$

The variances of $\hat{\lambda}$ and $\log(\hat{\lambda})$ are

$$\widehat{Var}(\hat{\lambda}) = \hat{\lambda}^2 \widehat{Var}(\log(\hat{\lambda}))$$

$$\widehat{Var}(\log(\hat{\lambda})) = \widehat{Var}(\log(\hat{r}_0)) + \widehat{Var}(\log(\hat{r}_1)) = Y_0^{-2} n_0 S_0^2 + Y_1^{-2} n_1 S_1^2$$

Assuming $\log(\hat{\lambda})$ is asymptotically from a normal distribution, the $100(1-\alpha)\%$ confidence interval for $\log(\hat{\lambda})$ can be constructed as

$$\left[\log(\hat{\lambda}) - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{Var}(\log(\hat{\lambda}))}, \log(\hat{\lambda}) + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{Var}(\log(\hat{\lambda}))}\right]$$

Then, the $100(1-\alpha)\%$ confidence interval for $\hat{\lambda}$ is

$$\left[\hat{\lambda}\exp\left(-z_{1-\frac{\alpha}{2}}\sqrt{\widehat{Var}(\log(\hat{\lambda}))}\right),\quad \hat{\lambda}\exp\left(z_{1-\frac{\alpha}{2}}\sqrt{\widehat{Var}(\log(\hat{\lambda}))}\right)\right] \quad (1)$$

The p-value for testing the null hypothesis of $H_0: \lambda = 1$ is calculated as

$$p = 2\Phi\left(|\log(\hat{\lambda})|/\sqrt{\widehat{Var}(\log(\hat{\lambda}))}\right) \quad (2)$$

The missing values will be imputed using multiple imputation by the retrieved dropout approach. The retrieved dropout participants are those who discontinue study intervention prior to Week 26 but have non-missing measures at Week 26. If there are only a limited number of retrieved participants that leads to a failure in performing the multiple imputation analysis, such as the model cannot converge, or the number of retrieved dropout participants is small, the missing HbA1c at Week 26 will be imputed by return-to-baseline multiple imputations.

6.12. Appendix 12: Interaction Effect for Subgroup Analysis – Treatment Regimen Estimand

The ANCOVA analysis will be performed within each subgroup with multiple imputation of missing primary measures. Statistical inference over multiple imputation of missing data will be guided by Rubin (1987) to obtain $\hat{\theta}_1$ and $se(\hat{\theta}_1)$ for the treatment difference in subgroup 1, and $\hat{\theta}_2$ and $se(\hat{\theta}_2)$ in subgroup 2. Then, the distribution of treatment by subgroup interaction (difference of treatment effects) is:

$\hat{\theta}_1 - \hat{\theta}_2 \sim N\left(\theta_1 - \theta_2, [se(\hat{\theta}_1)]^2 + [se(\hat{\theta}_2)]^2\right)$

A z-statistic can be contrasted such that $z = \frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{[se(\hat{\theta}_1)]^2 + [se(\hat{\theta}_2)]^2}} \sim N(0,1)$ under the null hypothesis of no treatment by subgroup interaction.

For more than 2 groups (K groups), let $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_K)$ and $Var(\hat{\theta}) = diag\left([se(\hat{\theta}_1)]^2, [se(\hat{\theta}_2)]^2, \ldots, [se(\hat{\theta}_K)]^2\right)$. A chi-square test (df=K-1) can be constructed as

$T = (C\hat{\theta})'(CVC')^{-1}(C\hat{\theta}) \sim \chi^2_{K-1}$

Where C is a matrix of contrast such that $C = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$.

7. References

[ADA] American Diabetes Association. Diabetes technology: standards of medical care in diabetes – 2021. *Diabetes Care*. 2021;44(Suppl 1):S85-S99. https://doi.org/10.2337/dc21-S007

Battelino T, Danne T, Bergenstal RM, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the International Consensus on Time in Range. *Diabetes Care*. 2019;42(8):1593-1603. https://doi.org/10.2337/dci19-0028

Battelino T, Alexander C, Amiel S. Continuous glucose monitoring and metrics for clinical trials: an international consensus statement. Lancet Diabetes Endocrinol 2023; 11: 42–57.

https://doi.org/10.1016/ S2213-8587(22)00319-9Bergenstal RM, Beck RW, Close KL, et al. Glucose management indicator (GMI): a new term for estimating a1c from continuous glucose monitoring. *Diabetes Care*. 2018;41(11):2275-2280. https://doi.org/10.2337/dc18-1581

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med.* 2009;28(4):586-604. https://doi.org/10.1002/sim.3495

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J.* 2011;53(6):894-913. https://doi.org/10.1002/bimj.201000239

Clarke W, Kovatchev B. Statistical tools to analyze continuous glucose monitor data. *Diabetes Technol Ther*. 2009;11(suppl 1):S45-S54. https://doi.org/10.1089/dia.2008.0138

[EuroQol] EuroQol Research Foundation. EQ-5D-5L user guide, version 3.0. Updated September 2019. https://euroqol.org/publications/user-guides

[FDA 2022] STANDARD SAFETY TABLES AND FIGURES: INTEGRATED GUIDE.

https://downloads.regulations.gov/FDA-2022-N-1961-0046/attachment_1.pdf

[PHUSE 2022] https://phuse.s3.eu-central-1.amazonaws.com/Deliverables/Safety+Analytics/WP068.pdf

Hsu J, Berger RL. Stepwise confidence intervals without multiplicity adjustment for dose response and toxicity studies. *J Am Stat Assoc*. 1999;94:468–482.

Koch GG. Comments on 'Current issues in non-inferiority trials'. *Stat Med*. 2008;27(3):333-342. https://doi.org/10.1002/sim.2923

Kovatchev BP, Otto E, Cox D, et al. Evaluation of a new measure of blood glucose variability in diabetes. *Diabetes Care*. 2006;29(11):2433–2438. https://doi.org/10.2337/dc06-1085

Kovatchev BP, Shields D, Breton M. Graphical and numerical evaluation of continuous glucose sensing time lag. *Diabetes Technol Ther*. 2009;11(3):139-143. https://doi.org/10.1089/dia.2008.0044

[IFCC] International Federation of Clinical Chemistry (IFCC) Standardization of HbA1c. Harmonizing Hemoglobin A1c Testing: A better A1C Test means better diabetes care. 2010. Accessed 19May2022. http://www.ngsp.org/ifccngsp.asp

Nelson WB. Recurrent Events Analysis for Product Repairs, Disease Recurrences, and Other Applications, *The ASA-SIAM Series on Statistics and Applied Probability*. 2003.

Qu Y, Dai B. Return-to-baseline multiple imputation for missing values in clinical trials. *Pharm Stat.* 2022; 21: 641-653. https://doi.org/10.1002/pst.2191

Qu Y, Luo J. Estimation of group means when adjusting for covariates in generalized linear models. *Pharm Stat*. 2015;14(1):56-62. https://doi.org/10.1002/pst.1658

Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharm Stat.* 2013;12(6):337-347. https://doi.org/10.1002/pst.1549

Rodbard D. New and improved methods to characterize glycemic variability using continuous glucose monitoring. *Diabetes Technol Ther*. 2009;11(9):551-565. https://doi.org/10.1089/dia.2009.0015

Rubin, DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons Inc.; 1987.

Torres C. A Tipping Point Method to Evaluate Sensitivity to Potential Violations in Missing Data Assumptions. 2019 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop. https://ww2.amstat.org/meetings/biop/2019/onlineprogram/ViewPresentation.cfm?file=301002.pdf

Signature Page for VV-CLIN-149255 v1.0

Approval	PPD 18-Mar-2024 13:27:41 GMT+0000

Signature Page for VV-CLIN-149255 v1.0