Statistical Analysis Plan: I8H-MC-BDCW

A Phase 3, Parallel-Design, Open-Label, Randomized Control Study to Evaluate the Efficacy and Safety of LY3209590 Administered Weekly Using a Fixed Dose Escalation Compared to Insulin Glargine in Insulin-Naïve Adults with Type 2 Diabetes

NCT05662332

Approval Date: 26-Jul-2024

Title Page

Protocol Title: A Phase 3, Parallel-Design, Open-Label, Randomized Control Study to Evaluate the Efficacy and Safety of LY3209590 Administered Weekly Using a Fixed Dose Escalation Compared to Insulin Glargine in Insulin-Naïve Adults with Type 2 Diabetes

Protocol Number: I8H-MC-BDCW (a)

Compound Number: LY3209590

Short Title: Efficacy and Safety of LY3209590 Administered Weekly Using a Fixed Dose Escalation Compared to Insulin Glargine in Adults with Type 2 Diabetes Who Are Starting Basal Insulin Therapy for the First Time

Acronym: QWINT-1

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND	129390

Document ID: VV-CLIN-135884

Approved on 26 Jul 2024 GMT

Table of Contents

Approved on 26 Jul 2024 GMT

Abbreviations and Definitions

Term	Definition
ADA	antidrug antibody
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
CI	confidence interval
CN	conventional
CRF	case report form
CV	coefficient of variation
DID-EQ	Diabetes Injection Device Experience Questionnaire
DKA	diabetic ketoacidosis
DMC	Data Monitoring Committee
DTSQ-S	Diabetes Treatment Satisfaction Questionnaire – Status Version
EAS1	Efficacy Analysis Set 1
EAS2	Efficacy Analysis Set 2
eGFR	estimated glomerular filtration rate
FBG	fasting blood glucose
GGT	Gamma glutamyl transferase
GLP-1 RA	glucagon-like peptide-1 receptor agonist
GV	glucose variability
HbA1c	hemoglobin A1c
HLT	high level term
IP	investigational product

ISR	injection site reaction
LLT	lowest level term
LS	least squares
MACE	major adverse cardiovascular event
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified Intent-to-Treat
MMRM	mixed model repeated measurement
MRD	minimum required dilution
NIM	noninferiority margin
PD	pharmacodynamic
PK	pharmacokinetic
SAP	Statistical Analysis Plan
SD	standard deviation
SI	System International
SIM-Q	Simplicity Questionnaire
SMBG	self-monitoring of blood glucose
SMQ	standardized MedDRA query
SoA	Schedule of Activities
SS	Safety Analysis Set
T1D	type 1 diabetes
T2D	type 2 diabetes
TBL	total bilirubin
TE-ADA	treatment-emergent antidrug antibody
TEAE	treatment-emergent adverse event
TIA	transient ischemic attack
TRIM-D	Treatment-Related Impact Measure – Diabetes
ULN	upper limit of normal

Version history

The Statistical Analysis Plan (SAP) Version 1 was approved on 02 Sep 2022, prior to the first participant visit of the study. The SAP Version 2 was approved on 23 Jan 2024, prior to the database lock.

This SAP for Study I8H-MC-BDCW (BDCW) is the third version and approved prior to the database lock. This version is based on the initial protocol and protocol amendment (a) approved on 18 May 2022 and 31 May 2023, respectively.

Table 1. SAP Version History Summary

Version	Section # and Name	Description of Change	Brief Rationale
Version 2	1.1. Objectives, Endpoints, and Estimands	Clarified the hypoglycemia endpoint included in the tertiary composite endpoints are level 2 or 3	For clarity
	3. Analysis Sets	For EAS2, updated data cutoff for glargine arm	Based on PK profile
	4.1. General Considerations	Updated the definitions of baseline and post-baseline observations for different analysis	For clarity and consistency with data collection
	4.3.3.and 4.4.1.3. Sensitivity Analysis	1. Updated the two-way tipping point analysis in 4.3.3. 2. Added sensitivity analyses by including inadvertently enrolled participants for primary and key secondary efficacy analyses	To address regulatory feedback
	4.5.1. Tertiary Efficacy Endpoints	Used EAS1 for the binary outcome of HbA1c and the composite endpoint of HbA1c and hypoglycemia and updated the analysis details	To address regulatory feedback
	4.6. Safety Analyses	1. For laboratory analysis, removed shift analysis, treatment emergent high/low and add elevated or low values meeting specified levels	Per PSAP update

Version	Section # and Name	Description of Change	Brief Rationale
		2. Used risk difference and 95% for safety categorical data analysis	
	4.7.1. Immunogenicity	Simplified the section and reduced analysis at the study level	Per PSAP update
	Throughout SAP	Minor changes and reorganization	For clarity
Version 3	4.1 General Considerations	Clarified the SAP language regarding the use of unplanned measurements for post-baseline observations for HbA1c analysis in Section 4.1.	To clarify FDA feedback.
	4.1. General Consideration 4.4.2.1. Other Efficacy endpoints 4.5.1. Tertiary Efficacy Endpoints	Added treatment regimen analysis for the selected secondary and tertiary endpoints (fasting glucose from SMBG and fasting serum glucose)	According to FDA feedback
	4.5.1. Tertiary Efficacy Endpoints	1. Modified missing data handling approach for hypoglycemia in the composite endpoints of HbA1c and hypoglycemia 2. Updated the logistic regression for binary outcomes with unconditional treatment group effect for binary endpoints	According to FDA feedback
	4.5.3. Other Efficacy Endpoints	Added missing data handling approach for SMBG 6-point profiles analysis	Clarification
	4.6.3.1. Participant-Reported	Added a sensitivity analysis that considers all hypoglycemic events as one hypoglycemic event until a	According to FDA feedback

Version	Section # and Name	Description of Change	Brief Rationale
	Hypoglycemic Events	subsequent glucose value is ≥ 70 mg/dL	
	4.6.4. Device Product Complaints	Added analysis for device product complaints	According to the new SAP template requirement
	4.7.2. Subgroup Analyses	Updated the subgroups for race	According to FDA feedback
	4.7.2. Subgroup Analyses	Removed subgroup analysis for Age (<65 years, ≥65 years to < 75 and ≥75)	Due to the small number of subjects in the ≥75 subgroup.

1. Introduction

The study protocol contains a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints. There are no changes to the analyses described in the protocol. This SAP includes the analyses details for efficacy, safety measures, and patient-reported outcomes. PK/PD analyses will be conducted by the PK/PD group and will be described in the PK analysis plan. The specifications for 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 for the treatment of T2D in adults	Change in HbA1c from baseline to Week 52
Key Secondary (Gated)	
Demonstrate superiority of LY3209590 compared to insulin glargine	Change in HbA1c from baseline to Week 52
Other Secondary	
Compare the effect of LY3209590 to insulin glargine in parameters of glycemic control	• Change from baseline to Week 26 for HbA1c • Change from baseline to Weeks 16, 26, and 52 for fasting glucose measured by SMBG • Insulin dose at Weeks 16, 26, and 52
Compare the effect of LY3209590 to insulin glargine on safety endpoints	• Incidence and rate of composite of Level 2 and 3 hypoglycemia events during the treatment period • Incidence and rate of composite of Level 2 and 3 nocturnal hypoglycemia events during the treatment period • Change from baseline to Weeks 26 and 52 in body weight
Compare the effect of LY3209590 to insulin glargine on patient-reported outcomes questionnaires	• Change from baseline to Weeks 26 and 52 for: o TRIM-D, and o DTSQ • Treatment experience at Weeks 26 and 52 for: o DID-EQ, and o SIM-Q
Tertiary/Exploratory	
Compare the effect of LY3209590 to insulin glargine for efficacy parameters	• Percentage of participants at Week 52 achieving: o HbA1c <7% o HbA1c <7% without Level 2 or 3 nocturnal hypoglycemia, and o HbA1c ≤6.5% without Level 2 or 3 hypoglycemia • Change from baseline to Weeks 26 and 52 for fasting serum glucose
Evaluate the percentage of participants receiving LY3209590 that require a transition to the LY3209590 prefilled pen	• Percentage of participants that require transition to the LY3209590 prefilled pen at Weeks 26 and 52
Evaluate the percentage of participants at each fixed LY3209590 dose level	• Percentage of participants at each fixed dose level at Weeks 26 and 52

Objectives	Endpoints
Compare the effect of LY3209590 to insulin glargine for safety parameters	• Incidence and rate of Level 2 hypoglycemia events during treatment period • Incidence and rate of Level 3 hypoglycemia events during treatment period • Incidence of positive LY3209590 treatment-emergent antidrug antibodies
Characterize the PK/PD of LY3209590	• LY3209590 PK and concentration-response relationships to key safety and efficacy measures
Compare the effect of LY3209590 to insulin glargine on patient-reported outcomes questionnaires	Frequency of responses to Basal Insulin Experience of "likelihood of incorporating into routine"

Abbreviations: DID-EQ = Diabetes Injection Device Experience Questionnaire; DTSQ = Diabetes Treatment Satisfaction Questionnaire; HbA1c = hemoglobin A1c; PK/PD = pharmacokinetics/pharmacodynamics; SIM-Q = Simplicity Questionnaire; SMBG = self-monitoring of blood glucose; T2D = type 2 diabetes; TRIM-D = Treatment-Related Impact Measure – Diabetes.

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 HbA1c change from baseline to Week 52, in study eligible participants, regardless of treatment discontinuation for any reason and regardless of initiation of rescue medication?

The *treatment regimen estimand* will be used for the primary objective and the attributes are described in the following table.

Treatment Regimen Estimand Attribute	Description
Population	Targeted study population. See Section 3 for details.
Endpoint	HbA1c change from baseline to Week 52.
Remaining intercurrent events	None. The 2 intercurrent events, treatment discontinuation for any reason and initiation of rescue medication, are both addressed by the treatment condition of interest attribute.
Treatment condition	The randomized treatment regardless of treatment discontinuation and use of rescue medications.
Population-level summary	Difference in mean changes between treatment conditions.

Abbreviation: HbA1c = hemoglobin A1c.

Rationale for estimand: The treatment regimen estimand estimates how participants with T2D are treated in clinical practice and considers 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 HbA1c change from baseline after 52 weeks of treatment, in

study eligible participants who adhere to the randomized treatment without intercurrent events during the study treatment period?

The *efficacy estimand* will be used for the primary objective. This table describes the efficacy estimand attributes.

Efficacy Regimen Estimand Attribute	Description
Population	Targeted study population. See Section 3 for details.
Endpoint	HbA1c change from baseline to Week 52.
Remaining intercurrent events	None. The 2 intercurrent events, treatment discontinuation for any reason and initiation of rescue medication, are both handled by the hypothetical strategy, for example, the potential outcome for those participants if the intercurrent events have not occurred will be estimated.
Treatment condition	The randomized treatment.
Population-level summary	Difference in mean changes between treatment conditions.

Abbreviation: HbA1c = hemoglobin A1c.

Rationale for estimand: The efficacy estimand supports the interpretation of the treatment effect as participants adhere to study treatment and free from the confounding effect of rescue medications.

Secondary estimand (for gated objective)

The superiority test in change from baseline to Week 52 (Visit 33) in HbA1c will also be based on the primary estimands described above.

1.2. Study Design

- This is a Phase 3, parallel-design, open-label, randomized control, treat-to-target study to evaluate the efficacy and safety of LY3209590 administered once-weekly compared to insulin glargine administered daily in adult participants with T2D that are insulin naïve.
- The study includes a 3-week screening and lead-in period, a 52-week treatment period, and a 5-week safety follow-up period after the last visit in the treatment period.
- Participants will be randomly assigned to receive daily insulin glargine or weekly insulin LY3209590. Stratification will be by country, HbA1c stratum at Visit 1 (Week -3) (<8.0%, ≥8.0%), and GLP-1 RA use at randomization, regardless of oral or injectable administration.
    - **Participants randomly assigned to insulin glargine**
      Participants randomly assigned to insulin glargine will start on 10 U/day and will have weekly dose adjustments based on FBG and hypoglycemia events.
    - **Participants randomly assigned to receive LY3209590**
      Participants randomly assigned to receive LY3209590 will be on a fixed dose escalation for at least the first 16 weeks of the study. Participants will start LY3209590 on a fixed dose of 100 U per week and will escalate the LY3209590 dose every 4 weeks as required to the set doses of 150, 250, and 400 U per week.

Treat-to-target dose adjustments will be based on FBG of 80 to 130 mg/dL (4.4 to 7.2 mmol/L) and hypoglycemia events.
Dose administration will be via a fixed dose autoinjector.
After Week 16, if a participant who has received the fixed dose of 400 U per week for at least 4 weeks has a median FBG level >130 mg/dL, then the investigator should transition the participant to titratable doses using a prefilled pen. This will provide more dosing flexibility and allow dose titrations above 400 U per week.
If a participant transitions from the fixed dose autoinjector to a prefilled pen, they must continue using the prefilled pen for the remainder of the study.

- Rescue therapy will be considered during the treatment period if the participants meet the protocol criteria of severe, persistent hyperglycemia.
  If a participant is transitioning to a prefilled pen, investigators should:
  - ensure that the transition does not occur at the same time as initiation of rescue therapy, and
  - wait for at least 4 weeks after the first dose with the LY3209590 prefilled pen.
- If study intervention is permanently discontinued, the participant will remain in the study and follow procedures for remaining study visits, as shown in the SoA in the protocol.

2. Statistical Hypotheses

Primary Hypothesis

The primary objective of this study is to test the hypothesis that LY3209590 is not inferior to insulin glargine on glycemic control as measured by change in HbA1c from baseline to Week 52 (Visit 33) in participants with T2D who are starting basal insulin for the first time.

The null hypothesis (H_0) is the difference between LY3209590 versus insulin glargine in the change from baseline to Week 52 (Visit 33) in HbA1c is greater than the non-inferiority margin (NIM).

The NIMs of 0.4% and 0.3% will both be tested to meet regulatory requirements. The 2-sided 95% CI will be used for testing the noninferiority.

Secondary Hypothesis

The key secondary (gated) objective is to test the hypothesis that LY3209590 is superior to insulin glargine for change from baseline to Week 52 (Visit 33) in HbA1c.

H_0: the difference (LY3209590-insulin glargine) $\geq$0.0.

These hypotheses will be tested using a strategy to control the overall type I error.

2.1. Multiplicity Adjustment

A gatekeeper approach (Westfall and Krishen 2001) will be used to control for type I error for the primary and gated secondary objective. The noninferiority test for the primary objective will be based on the 2-sided 95% CI. The primary analysis has been prespecified to use the NIM of 0.4% for the US submission using the full 0.05 alpha level. The NIM of 0.3% will be evaluated to meet the requirement from other regulatory agencies at full 0.05 alpha level. Once the upper limit of the 2-sided 95% CI is below the NIM, the non-inferiority is achieved, the gated secondary objective will be conducted at a 2-sided 0.05 significance level.

3. Analysis Sets

This table defines the populations for the purpose of analysis.

Analysis Populations or Datasets	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 rescue medications.
Efficacy Analysis Set 2 (EAS2) for efficacy estimand 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 medication. • the data cutoff for participants who had intercurrent events is defined by the earliest date from below dates for individual participants except for the analysis on study dose: o the date of last study dose + 10 days for LY3209590, or +1 day for glargine o the start date of the first rescue medication • the data cutoff for analysis on study dose is defined by the earliest date from below dates for individual participants: o the date of last study dose o the start date of the first rescue medication
Safety Analysis Set (SS)	The data will include: • mITT Population, and • all measurements regardless of the use of study treatment or rescue medications.

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 the SAP or the final clinical study report. Additional exploratory analyses of data will be conducted as deemed appropriate.

Unless otherwise stated, the efficacy analyses will be conducted on either EAS1 or EAS2. Primary and key secondary (gated) objectives will be based on both treatment regimen estimand and efficacy estimand, while the other secondary and tertiary efficacy measures will be based on efficacy estimand. For treatment regimen estimand, EAS1 (see the definition in Section 3) will be used for analysis. For efficacy estimand, other secondary and tertiary efficacy measures, EAS2 (see the definition in Section 3) will be used for analysis.

Unless otherwise noted, the safety analyses will be conducted on the SS. Percentages will be calculated using the mITT 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.

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05 and 2-sided 95% CIs will be calculated.

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 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. 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 CN and 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 Level 2 hypoglycemia and 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.

Below table describes the definition of baseline and postbaseline observations for different analyses.

Analysis	**Baseline Observations**	**Postbaseline Observations**
HbA1c (treatment regimen estimand)	The baseline is the last non-missing assessment prior to or at the first day of study treatment.	Planned measurements at Visit 33 (primary endpoint, Week 52) and Visit 25 (secondary endpoint, Week 26) in EAS1. Use unplanned measurements (collected on the same visit day) if no planned measurement. Multiple imputation approach will be used to impute missing observations at Visit 33 (Week 52) and Visit 25 (Week 26).
HbA1c (efficacy estimand)	The baseline is the last non-missing assessment prior to or at the first day of study treatment.	• Visits 6, 8, 12, 16, 18, 25, 28, and 33 (Weeks 2, 4, 8, 12, 16, 26, 36, and 52) in EAS2 for MMRM • For sensitivity analysis to compare the treatment effect of LY3209590 with autoinjector and insulin glargine: o LY3209590: Visits prior to or at the time of switch to prefilled pen in EAS2 for MMRM o Glargine: Visits 6, 8, 12, 16, 18, 25, 28, and 33 (Weeks 2, 4, 8, 12, 16, 26, 36, and 52) in EAS2 for MMRM Planned measurements at scheduled visits will be included. Use unplanned measurements (collected on the same visit day) if no planned measurements.

Analysis	Baseline Observations	Postbaseline Observations
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 Visit 33 (Week 52) in EAS2 for MMRM (efficacy estimand). • 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 next visit. Data at Visit 33 (Week 52), Visit 25 (Week 26), and Visit 18 (Week 16) in EAS1 for ANCOVA with multiple imputation for missing data (treatment regimen estimand). For sensitivity analysis to compare the treatment effect of LY3209590 with autoinjector and insulin glargine: • LY3209590: after the day of first dose of study treatment up to the visit prior to or at the time of switch to prefilled pen in EAS2 for MMRM (efficacy estimand). • Glargine: after the day of first dose of study treatment up to Week 52 (Visit 33) in EAS2 for MMRM (efficacy estimand).

Analysis	Baseline Observations	Postbaseline 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 Visit 33 (Week 52) in EAS2 for MMRM (efficacy estimand) Planned measurements at scheduled visits will be included. Data at Visit 33 (Week 52) and Visit 25 (Week 26) in EAS1 for ANCOVA with multiple imputation for missing data (treatment regimen estimand) For sensitivity analysis to compare the treatment effect of LY3209590 with autoinjector and insulin glargine: • LY3209590: after the day of first dose of study treatment up to the visit prior to or at the time of switch to prefilled pen in EAS2 for MMRM (efficacy estimand). • Glargine: after the day of first dose of study treatment up to Visit 33 (Week 52) in EAS2 for MMRM (efficacy estimand).
Insulin dose during the treatment period	N/A	All scheduled visits between Visit 3 and Visit 33 in EAS2 for MMRM. The averages of weekly basal insulin doses between visits for individual participants will be used in the analysis. (Section 4.4.2.1)

Analysis	Baseline Observations	Postbaseline Observations
Participant-reported hypoglycemia	The baseline period is the lead-in period prior to the first dose date of study treatment.	• Treatment period starts at or after the first dose date of study treatment and ends o At Visit 33 (Week 52) if completed treatment 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 Visit 33 (Week 52) +1 day if completed treatment 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 (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 first dose date of study treatment. Planned measurements at scheduled visits will be included.	• All scheduled visits after the first dose date up to Visit 802 for MMRM or ANCOVA Planned measurements at scheduled visits will be included.
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 measurements at scheduled and unscheduled visits will be included. The baseline for the 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 measurements at scheduled and unscheduled visits in the specified analysis period will be included.
Anti-LY3209590 antibody	Refer to PSAP.	Refer to PSAP.

Analysis	Baseline Observations	Postbaseline Observations
Patient-reported outcomes	The baseline will be the data collected at Visit 3.	• All scheduled visits after Visit 3 • Last collection after Visit 3 • For sensitivity analysis to compare the treatment effect of LY3209590 with autoinjector and insulin glargine: o LY3209590: Last collection prior to or at the time of switch to prefilled pen o Glargine: Last collection after Visit 3

Abbreviations: ANCOVA = analysis of covariance; CRF = case report form; EAS1 = efficacy analysis set 1; EAS2 = efficacy analysis set 2; HbA1c = hemoglobin A1c; MMRM = mixed model repeated measurement; N/A = not applicable; 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/discontinued from the study/treatment will be summarized by treatment using the Randomized Population. The individual reasons for discontinuation will also be included in the summary. Comparison will be conducted using the Fisher's exact test.

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 to warrant a summary) will be presented as a figure.

4.3. Primary Endpoint Analysis

4.3.1. Definition of Endpoint(s)

The primary endpoint of this study is the HbA1c change from baseline to Week 52 (Visit 33). 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 (NGSP 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

either of the 2 estimands: ***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 one non-missing postbaseline scheduled measure
Analysis Data	All non-missing observations at baseline and Week 52 (Visit 33) regardless of the use of study intervention or rescue medications.	All non-missing observations at baseline and all scheduled postbaseline timepoints during treatment period (that is, Weeks 2, 4, 8, 12, 16, 26, 36, and 52) prior to the date of last study dose + 10 days for LY3209590 (+1 day for Glargine), or the initiation of rescue medication, whichever is earlier for participants with intercurrent events.
Missing Data	There may be missing values at Week 52 (Visit 33) 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 52 but have non-missing measures at Week 52 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 52 will be imputed by return-to-baseline multiple imputation approach (Qu and Dai, 2022).	There may be missing values due to the early discontinuation of study intervention or use of rescue medication. 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 and GLP-1 RA treatment 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 GLP-1 RA treatment at randomization), visit and treatment-by-visit interaction as fixed effects, and baseline 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; EAS1 = efficacy analysis set 1; EAS2 = efficacy analysis set 2; GLP-1 RA = glucagon-like peptide-1 receptor agonist; HbA1c = hemoglobin A1c; MMRM = mixed model repeated measurement.

The 2-sided 95% CI of the LS mean difference (LY3209590 – insulin glargine) in the HbA1c change from baseline to Week 52 (Visit 33) 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 (measured in %) 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 corresponding p-value of 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 using treatment regimen estimand (Section 4.3.2) for the LY3209590 group only. The ANCOVA model for treatment regimen estimand will 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 both treatment regimen estimand and efficacy estimand by including inadvertent enrolled participants using mITT population.

4.3.3.3. Other Sensitivity Analysis

The following sensitivity analyses will be performed using the same MMRM model as for the efficacy estimand described in Section 4.3.2 based on the EAS2 data.

- Sensitivity analysis will be conducted excluding the data after the change to LY3209590 with a prefilled pen to compare the HbA1c change from baseline to Week 52 (Visit 33) between LY3209590 with the autoinjector and insulin glargine.
- To balance the exposure time for LY3209590 with autoinjector and insulin glargine, sensitivity analysis will be conducted to compare the HbA1c change from baseline to Week 52 (Visit 33) for participants receiving LY3209590 with autoinjector for the whole treatment period and participants receiving insulin glargine.

4.3.4. Supplementary Analyses

Additional analysis may be conducted as needed.

4.4. Secondary Endpoint(s) Analysis

4.4.1. Gated Secondary Endpoint

A gatekeeper approach (Westfall and Krishen 2001) will be used to control for type I error for the primary and gated secondary objective. 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 gated secondary objective will be conducted at a 2-sided 0.05 significance level.

4.4.1.1. Definition of Endpoint(s)

See Section 4.3.1 for the HbA1c change from baseline to Week 52 (Visit 33).

4.4.1.2. Main Analytical Approach

The superiority test in change from baseline to Week 52 (Visit 33) in HbA1c will be based on the same primary endpoint analysis described in Section 4.3.2.

4.4.1.3. Sensitivity Analysis

The analyses described in Section 4.4.1.2 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 the other supportive efficacy endpoints of insulin dose, and patient-reported outcomes will be based on the EAS2 data. The analysis of change from baseline for HbA1c at Week 26 and fasting glucose by SMBG at Week 16, 26 and 52 will be performed for both treatment regimen estimand and efficacy estimand.

For treatment regimen estimand, only participants with an observation at baseline or at the endpoint visit will be included in the analysis. The missing baseline will be imputed using multiple imputation under assumption of missing at random. Missing data at the endpoint visit will be imputed using either retrieved dropout or return-to-baseline multiple imputation approach determined by the criterion described for the primary efficacy endpoint in Section 4.3.2. ANCOVA analysis will be conducted similar to that for the primary endpoint. An additional term of baseline HbA1c stratum (<8.0%, ≥8.0%) will be added into the model for the endpoints other than HbA1c.

For efficacy estimand, participants with baseline and at least one post baseline observation will be included in the analysis. The longitudinal observations of actual and change from baseline in HbA1c to Week 26 will be analyzed using the same MMRM model as for the efficacy estimand described in Section 4.3.2. The analyses of other continuous efficacy measures (fasting glucose measured by SMBG) will use an MMRM model similar to that for the primary endpoint with an additional term of baseline HbA1c stratum (<8.0%, ≥8.0%). If a subject does not indicate which BG measurement was fasting, the FBG value will be assigned programmatically by designating the FBG as the first measurement between 5am-10am. These programmatically derived FBG values will also be included in the analysis. The variance-covariance structure in the MMRM models for fasting glucose from SMBG and insulin dose will be based on compound symmetry.

The analysis of study insulin dose will use an MMRM model on the EAS2 data with treatment, strata (country, GLP-1 RA treatment at randomization, and baseline HbA1c stratum [<8.0%, ≥8.0%]), visit, and treatment-by-visit interaction as fixed effects. The average weekly dose during each visit for individual participants will be used in the analysis. In the LY3209590 group, the average dose of each visit (that is, average of weekly doses since last visit) will be used as the average weekly dose. In the insulin glargine group, the average weekly dose will be calculated by the average daily dose multiplying by 7. The analysis will be repeated for average daily dose in U and daily dose in U/kg. The body weight at the visit will be used to calculate the daily dose in U/kg at the same visit.

4.4.2.2. Other Safety Endpoints

For other safety measures, the details are provided in Section 4.6.

4.4.2.3. Patient-Reported Outcome

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

4.4.2.3.1. Treatment-Related Impact Measure – Diabetes (TRIM-D)

The TRIM-D is a self-administered instrument, which assesses the impact of diabetes treatment on participants' functioning and well-being across available diabetes treatments (Brod et al. 2009). The TRIM-D consists of 28 items each assessed on a 5-point scale, where higher scores indicate a better health state, with a recall period of "over the past 2 weeks." In addition to an overall score, the TRIM-D items assess 5 domains of impact:

- Treatment Burden (6 items: Question 1a, Question 2a to 2e)
- Daily Life (5 items: Question 3a to 3b, Question 5a to 5c)
- Diabetes Management (5 items: Question 4a to 4e)
- Compliance (4 items: Question 6a, b, c, e)
- Psychological Health (8 items: Question 6d, Question 7a to 7g)

Items within each domain are summed to obtain a raw domain score, which is 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, Diabetes Management = ((sum of raw score -5)/(25-5))×100. If answer for the 5 questions in Diabetes Management (Question 4a to 4e) are all 1, then the sum of raw score is 5 and the transformed domain score is 0. All items can also be summed and transformed to obtain a transformed total score in a similar way.

Summary statistics for the transformed scores of each domain will be provided by study treatment. The scores at scheduled visits and the change from baseline will be analyzed by the MMRM model similar to the other efficacy endpoints (see Section 4.4.2.1). The last non-missing postbaseline observations will be analyzed by an ANCOVA model using treatment, strata (country, GLP-1 RA treatment at randomization, and baseline HbA1c stratum), and the baseline value of the dependent variable as independent variables.

In addition, sensitivity analysis will be conducted using the last non-missing postbaseline observations prior to or at the time of switch to LY3209590 with a prefilled pen to compare the patient-reported outcomes between LY3209590 with autoinjector and insulin glargine. The same ANCOVA model mentioned above will be used for the sensitivity analysis.

4.4.2.3.2. Diabetes Treatment Satisfaction Questionnaire – Status Version (DTSQ-s)

Description of DTSQ-s

The DTSQ-s (Bradley and Lewis 1990; Bradley 1999) is a diabetes-specific, patient-reported outcome instrument that assesses the overall treatment satisfaction and perceived frequency of hyperglycemia and hypoglycemia. It is appropriate for use in both T1D and T2D. The DTSQ-s consists of 8 items that assess treatment satisfaction as well as concerns about hyperglycemia and hypoglycemia over the past few weeks, prior to the visit. Each item is rated on a 7-point Likert scale. Items 1, 4, 5 through 7, and 8 are rated from 0 (very dissatisfied) to 6 (very satisfied) and

can be summed up to produce a treatment satisfaction score. Items 2 and 3 evaluate the perceived frequency of hyperglycemia and hypoglycemia and are rated from 0 (none of the time) to 6 (most of the time).

Summary statistics for overall treatment satisfaction score, perceived frequency of hyperglycemia score, and perceived frequency of hypoglycemia score collected in DTSQ-s at baseline will be provided by study treatment. Treatment comparison will be conducted by Wilcoxon rank sum test.

4.4.2.3.3. Diabetes Treatment Satisfaction Questionnaire – Change Version (DTSQ-c)

Description of DTSQ-c

The DTSQ-c (Bradley 1999) was designed to overcome potential ceiling effects in the status version. The DTSQ-c has the same 8 items as the status version but is reworded slightly to measure the change in treatment satisfaction rather than absolute treatment satisfaction. Each item is scored on a scale of -3 to +3. For Items 2 and 3, representing perceived frequency of hyperglycemia, and perceived frequency of hypoglycemia, the lower the score represents better the perception. For the remaining items:

- the higher the score represents the greater the improvement in treatment satisfaction
- the lower the score represents the greater the deterioration in treatment satisfaction, and
- a score of 0 represents no change.

Summary statistics for overall treatment satisfaction score, perceived frequency of hyperglycemia score, and perceived frequency of hypoglycemia score collected in DTSQ-c at postbaseline visits will be provided by study treatment. Wilcoxon rank-sum test will be used for treatment comparison. In addition, the last non-missing postbaseline observations prior to or at the time of switch to LY3209590 with a prefilled pen will be summarized and compared to the last non-missing postbaseline observations for insulin glargine by Wilcoxon rank sum test.

4.4.2.3.4. Simplicity Questionnaire (SIM-Q) Single Medication Status Version

The SIM-Q is a brief 10-item measure developed to assess the simplicity and complexity of treatment for T2D. This version of the instrument assesses the simplicity and complexity of a single medication. This measure asks participants to consider only the assigned study intervention when completing each item on a 5-point scale ranging from “Very complex” to “Very simple.” In Study BDCW, only the last 2 questions of the SIM-Q will be completed – “How simple or complex is your medication treatment for diabetes?” and “Overall, how simple or complex is it to manage your diabetes, including medication, checking your blood glucose levels, diet, and any other aspects of diabetes treatment?”

The frequency and proportion of the response at Week 26 (Visit 25) and Week 52 (Visit 33) for each question will be summarized by study treatment. Treatment comparison will be conducted by Wilcoxon rank sum test. In addition, the last non-missing postbaseline observations prior to or at the time of switch to LY3209590 with a prefilled pen will be summarized and compared to the last non-missing postbaseline observations for insulin glargine by Wilcoxon rank sum test.

4.4.2.3.5. 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 non-insulin 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 non-insulin 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 with the possible range from 0 to 100 using the same calculation method in Section 4.4.2.3.1.

In addition, there are 3 global items:

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

These 3 global items are each scored separately.

Summary statistics for the Device Characteristics subscale will be provided by study treatment. The scores at scheduled visits will be analyzed by the MMRM model with treatment, strata (country, GLP-1 RA treatment at randomization, and baseline HbA1c stratum [<8.0%, ≥8.0%]), visit and treatment-by-visit interaction as fixed effects. The last non-missing postbaseline observations will be analyzed by an ANCOVA model using treatment, strata (country, GLP-1 RA treatment at randomization, and baseline HbA1c stratum) as independent variables. In addition, sensitivity analysis will be conducted using the last non-missing postbaseline observations prior to or at the time of switch to LY3209590 with a prefilled pen to compare the patient-reported outcomes between LY3209590 with autoinjector and insulin glargine. The same ANCOVA model mentioned above will be used for the sensitivity analysis.

The frequency and proportion of the 3 global items at Week 26 (Visit 25) and Week 52 (Visit 33) will be summarized by study treatment. Wilcoxon rank sum test will be used for treatment comparison. In addition, the last non-missing postbaseline observations prior to or at the time of switch to LY3209590 with a prefilled pen will be summarized and compared to the last non-missing postbaseline observations for insulin glargine by Wilcoxon rank sum test.

4.5. Tertiary/Other Endpoints Analysis

4.5.1. Tertiary Efficacy Endpoints

For continuous endpoints, similarly to what is described in section 4.4.2.1 for supportive secondary efficacy endpoints, either ANCOVA model using data from EAS1 or MMRM model using data from EAS2 will be performed.

For the binary outcome endpoints of percentage of participants at Week 52 achieving HbA1c <7%, HbA1c <7% without nocturnal level 2 or 3 hypoglycemia, HbA1c ≤6.5% without level 2 or 3 hypoglycemia, and HbA1c <7% without level 2 or 3 hypoglycemia, EAS1 will be used for the analysis and the details are provided in the table below:

Analysis population	All participants in EAS1 with non-missing baseline measure
Analysis data	All non-missing observations at Week 52 (Visit 33) during treatment period regardless of the use of study intervention or rescue therapy
Endpoint	• Binary outcome of HbA1c < 7% with 1 indicating achieving HbA1c target • The composite of o binary outcome of HbA1c <7% at Week 52 with 1 indicating achieving HbA1c target. o binary outcome of no nocturnal hypoglycemia (<54 mg/dL or severe) during 52-week treatment period with 1 indicating no occurrence of nocturnal hypoglycemia. • The composite of o binary outcome of HbA1c ≤6.5% at Week 52 with 1 indicating achieving HbA1c target. o binary outcome of no hypoglycemia (<54 mg/dL or severe) during 52-week treatment period with 1 indicating no occurrence of hypoglycemia. • The composite of o binary outcome of HbA1c <7% at Week 52 with 1 indicating achieving HbA1c target. o binary outcome of no hypoglycemia (<54 mg/dL or severe) during 52-week treatment period with 1 indicating no occurrence of hypoglycemia.
Missing data handling	• For HbA1c, missing values at Week 52 (Visit 33) will be imputed using the same method for the primary endpoint. The binary outcomes of HbA1c <7% or ≤6.5% will be based on the imputed data. • For nocturnal hypoglycemia or hypoglycemia that are included in the composite endpoints, a participant who discontinued the treatment period before Week 52 (Visit 33) is considered as a non-responder (i.e. experienced the event) and the binary outcome value is 0.
Analysis model	Each endpoint will be analyzed using a logistic regression model including treatment, strata (country, and GLP-1 RA treatment at randomization), and baseline HbA1c value. The odds ratio between LY3209590 and insulin glargine will be used for treatment comparison. The unconditional treatment group effect will be assessed based on a robust variance estimator for g-computation estimators (Ye et al. 2023). Multiple imputation will be performed, and the statistical inference will be based on the multiple imputation framework by Rubin (1987)

The analysis of fasting serum glucose will be performed for both treatment regimen estimand and efficacy estimand. The analyses of fasting serum glucose by laboratory will use same MMRM model as for the efficacy estimand and same ANCOVA model as for the treatment regimen estimand described in Section 4.4.2.1.

The frequency and proportion of participants who switch to the LY3209590 prefilled pen prior to Week 26 (Visit 25) and Week 52 (Visit 33) will be summarized. A Kaplan-Meier curve will be constructed for the time to switch. Furthermore, the frequency and proportion of participants at each fixed LY3209590 dose level at Week 26 (Visit 25) and Week 52 (Visit 33) will be summarized.

4.5.2. Tertiary Patient Reported Outcome

4.5.2.1. 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 responses at Week 52 (Visit 33) will be summarized by study treatment. Wilcoxon rank sum test will be used for treatment comparison. In addition, the last non-missing postbaseline observations prior to or at the time of switch to LY3209590 with a prefilled pen will be summarized and compared to the last non-missing postbaseline observations for insulin glargine by Wilcoxon rank sum test.

4.5.3. Other Efficacy Endpoints

- Participants will perform SMBG 6-point profiles over a 24-hour period on 2 nonconsecutive days during the 7-day period prior to visits shown in the SoA of the protocol. The 6-point profile consists of pre-meal and 2-hour postprandial SMBG measurements for the morning, midday, and evening meals during 1 day. The following SMBG variables will be analyzed and summarized based on the EAS2 data. Only participants with an observation at baseline or at the postbaseline visits will be included in the analysis. The missing baseline and postbaseline visits will be imputed using multiple imputation under assumption of missing at random. MMRM analysis will be conducted similar to that for the primary endpoint with an additional term of baseline HbA1c stratum (<8.0%, ≥8.0%). The statistical inference will be based on the multiple imputation framework by Rubin (1987).Between-day GV measured by the CV and SD of the FBG by SMBG
- Average of the SMBG values at each time point from 6-point SMBG profiles (pre-meal and 2-hour postprandial SMBG measurements for the morning, midday, and evening meals) collected during the 7-day period prior to visits
- Within-day and between-day GV measured by the CV and SD of 6-point SMBG. The between-day GV from SMBG was calculated in 2 steps: (1) GV at each time point calculated from values collected during the 7-day period prior to visits; (2) between-day GV calculated by averaging GVs over SMBG time points

4.5.4. Tertiary Safety Endpoints

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

4.6. Safety Analyses

Safety measures include treatment exposure, 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.

Percentages will be calculated using the mITT 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 and follow-up period will be provided. For 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 value by MMRM model using treatment, visit, and treatment by visit as fixed effect, baseline of the dependent variable as a covariate, and compound symmetry as the variance-covariance structure.

The incidence and event rate of participant-reported hypoglycemia will be summarized by treatment and analysis period for different types of hypoglycemia. Analysis details are provided in Section 4.6.3.1.

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, ≥365 days
- >0 and <30 days, ≥30 and <90 days, ≥90 and <180 days, ≥180 and <365 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 final 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 IP or reported to worsen in severity from baseline will be considered TEAEs. The MedDRA 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 System Organ Class (SOC)	Number and percentage of participants with TEAEs 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. 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): • TEAEs (occurring in ≥1% before rounding in LY3209590 group) • 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 AEs or 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: 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 CRF.

A summary of severe hypoglycemia by decreasing risk difference will be provided. The incidence and event rate during the treatment period will be analyzed using the method described in Section 4.6.3.1. 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 and will be based on two approaches to identify persistent-recurrent hypoglycemic events:

1. using investigator assessment and clinical judgment to determine if repeated hypoglycemia may have contributed to a hypoglycemic event that degenerated with poor outcomes, and
2. using a prespecified criteria to derive events from e-diary database (see definition in Appendix 10 (Section 6.9).

Identification of P-R hypoglycemia based on investigator reporting is precipitated by a hypoglycemic event a participant or their caregiver inform 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 them 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 blood glucose values and hypoglycemia reports to determine the best answer. If 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.

P-R hypoglycemia events and participants who present P-R hypoglycemia will be identified by one or both approaches explained above. 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 a CRF designed to record additional information about AEs suggestive of a hypersensitivity reaction. Analyses are based on the following:

- Anaphylactic reaction SMQ (20000021; narrow terms)

- Hypersensitivity SMQ (20000214; narrow terms)
- Angioedema SMQ (20000024; narrow terms)

The number and percentage of participants reported with 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 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

ISRs 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
- Lipodystrophies and localized amyloidosis, as represented by PTs of:
  - Lipoatrophy
  - Lipodystrophy acquired
  - Partial lipodystrophy
  - Lipohypertrophy
  - Sclerema
  - Cutaneous amyloidosis

The summary will present

- The number of participants reported at least one AE meeting any of the above categories
- The number of participants reported any AE in each category
- The number of participants reported any AE for each PT within a specific category

The PTs will be listed for summary within each category in decreasing risk difference between 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 for the following SMQ:

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

A summary will presented

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

4.6.2.1.6. Diabetic Ketoacidosis (DKA)

The 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
- Lactic acidosis

4.6.2.1.7. Diabetic Retinopathy or Maculopathy

The diabetic retinopathy or maculopathy will be searched by MedDRA PTs from all TEAEs. The list of PTs for the search is provided in Appendix 6 (Section 6.6). A summary of treatment-emergent retinopathy or maculopathy by PT will be provided.

4.6.2.1.8. Peripheral Edema

The peripheral edema will be searched by MedDRA PTs (see Section 6.7) from all TEAEs. 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 TEAEs meeting the SMQ narrow search criteria will be provided by PT.

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 medication 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 deviations (IPD) indicative of multiple dose, according to the Trial Issue Management Plan (TIMP). 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 individual trial 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. This listing will indicate if severe hypoglycemia or P-R hypoglycemia occurred after the MEI.

4.6.3. Additional Safety Assessments

4.6.3.1. Participant-Reported Hypoglycemic Events

The following types of hypoglycemia 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.

The hypoglycemia will also be further classified into

- nocturnal hypoglycemia (occurs between midnight and 0600)
- 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

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

of the multiple hypoglycemic events.

The combined event starts from the first records with Level 1, 2 or 3, and ends when there are no more events for at least 60 minutes.

The evaluation of potential persistent-recurrent hypoglycemia events will be through assessment of the clinically significant events confirmed by investigators based on clinical judgment and through events identified by a pre-specified criteria described in 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 below table.

Endpoint	Analysis Period*	Statistical Method
Event rate of Level 1 hypoglycemia events (events/participant/year): • All documented	Baseline, 0-16, 0-26, 0-52, 16-26, 26-52 weeks, post-treatment period	Negative binomial regression with treatment and baseline HbA1c as covariates, log (exposure/365.25 days) as the offset in the model.
Event rate of Level 2 hypoglycemia events (events/participant/year): • All documented	Baseline, 0-16, 0-26, 0-52, 16-26, 26-52 weeks, post-treatment period	Negative binomial regression with treatment 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 8 [Section 6.8] for details) will be used for treatment comparison.
Event rate of Level 3 hypoglycemia events (events/participant/100 year): • All documented	Baseline, 0-26, 0-52, 26-52 weeks, post-treatment period	Negative binomial regression with treatment and baseline HbA1c as covariates, log (exposure/36525 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 8 [Section 6.8] for details) will be used for treatment comparison.

Endpoint	Analysis Period*	Statistical Method
Event rate of Level 2/3 hypoglycemia events (events/participant/year): • All documented • Nocturnal • Non-nocturnal	Baseline, 0-16, 0-26, 0-52, 16-26, 26-52 weeks, post-treatment period	Negative binomial regression with treatment 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 8 [Section 6.8] for details) will be used for treatment comparison. 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 hypoglycemia events: • All documented	Baseline, 0-16, 0-26, 0-52, 16-26, 26-52 weeks, post-treatment period	Logistic regression with treatment and baseline HbA1c as covariates.
Incidence of Level 2 hypoglycemia events: • All documented	Baseline, 0-16, 0-26, 0-52, 16-26, 26-52 weeks, post-treatment period	Logistic regression with treatment and baseline HbA1c as covariates.
Incidence of Level 3 hypoglycemia events: • All documented	Baseline, 0-26, 0-52, 26-52 weeks, post-treatment period	Logistic regression with treatment and baseline HbA1c as covariates.
Incidence of Level 2/3 hypoglycemia events: • All documented • Nocturnal • Non-nocturnal	Baseline, 0-16, 0-26, 0-52, 16-26, 26-52 weeks, post-treatment period	Logistic regression with treatment and baseline HbA1c as covariates.

Endpoint	Analysis Period*	Statistical Method
Potential persistent-recurrent hypoglycemia events • identified by investigator • identified by pre-criteria (defined in Appendix 9 [Section 6.9])	Safety analysis period (see definition in Section 4.6)	The number of participants with at least one event will be summarized and compared by Fisher's exact test. The number of events will also be provided.

Abbreviations: HbA1c = hemoglobin A1c; Level 2/3= Level 2 and Level 3 composite; mITT = modified Intent-to-Treat.

Note: The hypoglycemia yearly rate during defined period is calculated by the number of hypoglycemia within the period/number of days patient at risk within the period times 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 events within the period (Yes/No).

A sensitivity analysis will be done for selected hypoglycemic endpoints where all hypoglycemic events are considered one hypoglycemic event until a subsequent glucose value is ≥ 70 mg/dL.

4.6.3.2. 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.
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.

Analysis	Details
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.
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.

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.

4.6.3.3. 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 patients with elevated or low values meeting specified levels (see Appendix 10, Section 6.10 at any time postbaseline (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 using Level 2 definitions in Appendix 10, Section 6.10) and qualitatively (abnormal during postbaseline) will be provided, including participant identification, treatment group, laboratory sample collection day (that is, days from start of study drug), analyte name, abnormal result, reference low or high limits, and level 2 cut-off value if applicable.

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. Analysis Data 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.

4.6.3.4. 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. ○ Systolic BP (mm Hg): ▪ Low: Level 1: <90 ▪ High: Level 1: ≥90, Level 2: ≥120, Level 3: ≥140, Level 4: ≥160, Level 5: ≥180 ○ 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.

Analysis Type	Analysis Details
Participants meeting CTC grade changes in weight	For weight, cutoffs informed by CTCAE version 5 (Grades 1-3) will be used: ○ (Loss) decrease: Level 1: ≥5%, Level 2: ≥10%, Level 3: ≥20% ○ (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

4.6.4. 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 nonmissing baseline ADA result, and at least 1 nonmissing 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 × 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.

Refer to PSAP for all definitions and additional details for immunogenicity.

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. Separate analysis without the terms related with the subgroup will be performed for each subpopulation. Subgroup analyses will be conducted as defined in this section. Additional subgroup analysis may also be performed as appropriate.

4.7.2.1. Subgroup Analyses in HbA1c

Subgroup analyses in HbA1c and change in HbA1c from baseline to Week 52 will be conducted with subgroups defined as:

- Age (<65 years and ≥65 years)
- Baseline HbA1c stratum (<8.0% and ≥8.0%)
- Ethnicity
- Gender
- Race (White, American Indian or Alaska Native, Asian, and Black or African American) – Native Hawaiian or Other Pacific Islander, and Multiracial are not included in the subgroup analysis because the number of subjects is too small.
- Region (US and non-US)
- Region (North America and South America)
- GLP-1 RA use at randomization
- Estimated Glomerular Filtration Rate (eGFR) at baseline (<60, 60 - <90, and ≥ 90 mL/min/1.73 m^2)
- Duration of diabetes (< median and ≥ median)
- Insulin Dose (>400 U at anytime and <= 400 U at anytime), where
  - ">400 U at anytime" group includes participants receiving insulin glargine with weekly dose >400 U at anytime up to week 52 (Visit 33) or participants who switch to prefilled pen for LY3209590, and
  - "<= 400 U at anytime" group includes participants receiving insulin glargine with weekly dose <=400 U at anytime up to Week 52 (Visit 33) or participants receiving LY3209590 with auto injector for the whole treatment period.

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 and covariate 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.

Another subgroup analysis for HbA1c will be conducted for treatment regimen estimand. 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 chi-square test based on estimated treatment differences within each subgroup. (See details in Appendix 11 [Section 6.11]).

4.7.2.2. Subgroup Analyses in Hypoglycemia

Subgroup analyses in documented Level 2/3 hypoglycemia, non-nocturnal, and nocturnal hypoglycemia rates during 0 to 52 weeks will be conducted with subgroups defined as:

- Baseline HbA1c stratum (<8.0% and ≥8.0%)
- Region (North America and South America)
- Region (US and non-US)
- GLP-1 RA use at randomization
- eGFR at baseline (<60, 60 - <90, and ≥ 90 mL/min/1.73 m^2)

The hypoglycemia rates will be analyzed using a negative binomial regression including the same independent variables for hypoglycemia event analyses (see Section 4.6.3.1) plus factors of subgroup, 2-way interaction of subgroup and treatment. Separate analysis without the terms related with the subgroup will be performed for each subpopulation.

4.8. Interim Analyses

4.8.1. Data Monitoring Committee (DMC)

An independent external 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 to 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 the DMC are included in the DMC charter.

4.9. Changes to Protocol-Planned Analyses

There are no changes but additional details are provided to the analyses described in the protocol.

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 52 (Visit 33), approximately 284 participants on LY3209590 and 284 participants on insulin glargine will complete 52 weeks of treatment.

The number of completers will provide at least 99% statistical power to show noninferiority between LY3209590 and insulin glargine using the upper limit of a 2-sided 95% CI and these assumptions:

- NIM of 0.4%
- no true difference between treatment groups, and
- an SD of 1.1%.

This sample size has at least 90% statistical power to show noninferiority between LY3209590 and insulin glargine using a 0.3% NIM at Week 52 (Visit 33).

The number of completers will provide 90% statistical power to show superiority between LY3209590 and insulin glargine for change in HbA1c from baseline to Week 52 (Visit 33), assuming an SD of 1.1% and true mean difference of -0.3%, and alpha of 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), 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%), fasting serum glucose (mmol/L and mg/dL), GLP-1 RA treatment at randomization, and number of non-insulin antihyperglycemic medications in use 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 within SOC. Similar summary will also be provided for preexisting conditions.

6.2. Appendix 2: Treatment Compliance

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

The study protocol provides dosing algorithms for both study treatments. The investigator will calculate the algorithm recommended dose based on participant's fasting blood glucoses 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 a reason for not following the algorithm recommended dose from a picklist. The number and percentage of investigator prescribed doses different from algorithm recommended doses will be provided to evaluate investigator's adherence to the dosing algorithm. The reasons for not following the algorithm recommended dose will also be summarized. The number and percentage of investigator prescribed doses that are not equal to participant administered dose will also be provided.

6.3. Appendix 3: Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the CTR requirements.

Analyses provided for the CTR requirements include the following:

- Summary of AEs, 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 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 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 Medication

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 ATC Level. The concomitant medications will be ordered by decreasing frequency of LY3209590 within each ATC level.

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: MedDRA PT for Diabetic Retinopathy or Maculopathy

The following PTs 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

6.7. Appendix 7: 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

6.8. Appendix 8: 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

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

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

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

6.9. Appendix 9: Definition for Persistent-Recurrent Hypoglycemia

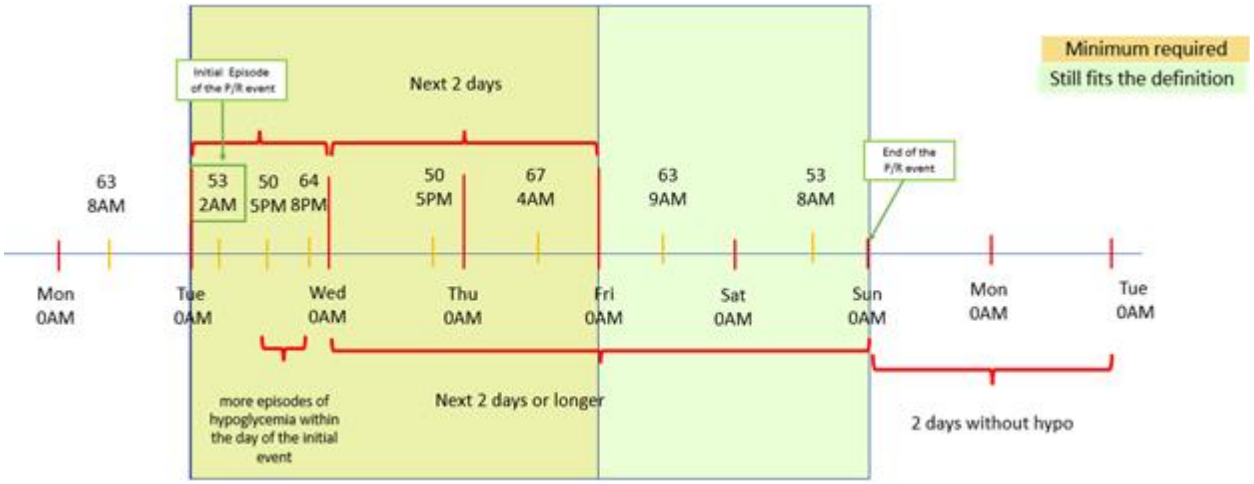
A P-R hypoglycemia based on programming search in the e-diary of hypoglycemic events that meet prespecified criteria is defined as a set of hypoglycemic episodes that:

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 one 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 the search criteria for a P-R hypoglycemia event is illustrated in the figure below.



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/1.73m2)	≥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

Parameter	Level 1	Level 2	Level 3
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

Abbreviations: eGFR = estimated glomerular filtration rate; HDL = high-density lipoprotein; LDL = low-density lipoprotein; WBC = white blood cell.

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

6.11. Appendix 11: 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: $\hat{\theta}_1 - \hat{\theta}_2 \sim N\left(\theta_1 - \theta_2\ , \left[se(\hat{\theta}_1)\right]^2 + \left[se(\hat{\theta}_2)\right]^2\right)$

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

For k groups (k ≥ 2),

let $\hat{\boldsymbol{\theta}} = \left(\hat{\theta}_1\ , \hat{\theta}_2\ , \ldots , \hat{\theta}_k\right)$ and $Var\left(\hat{\boldsymbol{\theta}}\right) = diag\left(\left[se(\hat{\theta}_1)\right]^2, \left[se(\hat{\theta}_2)\right]^2, \ldots , \left[se(\hat{\theta}_k)\right]^2\right)$.

A chi-square test (with degrees of freedom = k-1) can be constructed as

$$T = \left(C\widehat{\boldsymbol{\theta}}\right)'(CVC')^{-1}\left(C\widehat{\boldsymbol{\theta}}\right) \sim \chi^2_{k-1}$$

where C is a matrix of contrast such that

$$C = \begin{bmatrix} -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{bmatrix}.$$

7. References

[ADA] American Diabetes Association. Diabetes technology: standards of medical care in diabetes – 2021. Diabetes Care. 2021;45(Suppl 1):S85-S99. https://doi.org/10.2337/dc21-S007

Bradley C. Diabetes treatment satisfaction questionnaire: change version for use alongside status version provides appropriate solution where ceiling effects occur. *Diabetes Care*. 1999;22(3):530-532. https://doi.org/10.2337/diacare.22.3.530

Bradley C, Lewis KS. Measures of psychological well-being and treatment satisfaction developed from the responses of people with tablet-treated diabetes. *Diabet Med*. 1990;7(5):445-451. https://doi.org/10.1111/j.1464-5491.1990.tb01421.x

Brod M, Hammer M, Christensen T, et al. Understanding and assessing the impact of treatment in diabetes: the Treatment-Related Impact Measures for Diabetes and Devices (TRIM-Diabetes and TRIM-Diabetes Device). *Health Qual Life Outcomes*. 2009;7(83). https://doi.org/10.1186/1477-7525-7-83

[FDA 2022] STANDARD SAFETY TABLES AND FIGURES: INTEGRATED GUIDE. https://downloads.regulations.gov/FDA-2022-N-1961-0046/attachment_1.pdf

Matza LS, Boye KS, Stewart KD, et al. Development of the Diabetes Injection Device Experience Questionnaire (DID-EQ) and Diabetes Injection Device Preference Questionnaire (DID-PQ). *J Patient Rep Outcomes*. 2018;2:43. https://doi.org/10.1186/s41687-018-0068-z

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

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

[NGSP]. Harmonizing Hemoglobin A1c Testing. IFCC Standardization of HbA1c. Published 2010. Accessed May 16, 2022. http://www.ngsp.org/ifccngsp.asp

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

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

Westfall PH, Krishen A. Optimally weighted, fixed sequence and gatekeeper multiple testing procedures. *J Stat Plan Inference*. 2001;99(1):25-40. https://doi.org/10.1016/S0378-3758(01)00077-5

Signature Page for VV-CLIN-135884 v2.0

Approval	PPD 26-Jul-2024 18:09:24 GMT+0000

Signature Page for VV-CLIN-135884 v2.0