Statistical Analysis Plan I8F-MC-GPIM (2.0)

A Randomized, Phase 3, Double-blind Trial Comparing the Effect of the Addition of Tirzepatide versus the Addition of Placebo to Titrated Basal Insulin on Glycemic Control in Chinese Participants with Type 2 Diabetes (SURPASS-CN-INS)

NCT05691712

Approval Date: 08-Jan-2025

Protocol Title: A Randomized, Phase 3, Double-blind Trial Comparing the Effect of the Addition of Tirzepatide versus the Addition of Placebo to Titrated Basal Insulin on Glycemic Control in Chinese Participants with Type 2 Diabetes (SURPASS-CN-INS)

Protocol Number: I8F-MC-GPIM

Compound Number: LY3298176

Short Title: Comparing the Effect of the Addition of Tirzepatide or Placebo to Titrated Basal Insulin on Glycemic Control in Chinese Participants with Type 2 Diabetes

Acronym: SURPASS-CN-INS

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Lilly Corporate Center, Indianapolis, Indiana 46285, USA

Regulatory Agency Identifier Number(s): Not Applicable

Document ID: VV-CLIN-159774

Approved on 08 Jan 2025 GMT

Table of Contents

Table of Contents

Table **Page**

Approved on 08 Jan 2025 GMT

Table of Contents

Figure **Page**

Version history

This Statistical Analysis Plan (SAP) for study I8F-MC-GPIM is based on the protocol amendment (b) dated 17Nov2023.

Table GPIM.1.1. SAP Version History

SAP Version	Approval Date	Change	Rationale
1	16 Jul 2024	Not Applicable	Original version
2	17 Dec 2024	Section 1.1 2 places of "HbA1c ≤5.7%" were mistakenly typed in "Endpoints/Estimands". Corrected as "HbA1c<5.7%". Section 2 Figure GPIM.2.1 "5mg" was mistakenly typed in the chunk on top of "$H_{15,6}$". Corrected as "15mg".	Typos corrected.

1. Introduction

Study GPIM is a multicenter, randomized, double-blind, parallel, placebo-controlled Phase 3 study which will assess the efficacy and safety of the addition of 5 mg, 10 mg, or 15 mg tirzepatide, as compared with placebo in participants with T2DM as an add-on to titrated basal insulin alone or in combination with metformin with or without SGLT-2i over a 40-week treatment. The purpose of this study is to compare the effect of the addition of tirzepatide or placebo to titrated basal insulin on glycemic control in Chinese participants with type 2 diabetes.

Approved on 08 Jan 2025 GMT

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints/Estimands
Primary	
• To demonstrate superiority of QW tirzepatide 10 mg and/or 15 mg versus placebo when added to titrated basal insulin alone or in combination with metformin with or without SGLT-2i, with respect to glycemic control at 40 weeks for:	• Mean change in HbA1c from baseline
Key Secondary (controlled for type 1 error)	
• To demonstrate superiority of QW tirzepatide 5 mg versus placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i, with respect to glycemic control at 40 weeks for:	• Mean change in HbA1c from baseline
• To demonstrate superiority of QW tirzepatide 5 mg, 10 mg, and/or 15 mg versus placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i, at 40 weeks for:	• Mean change in body weight from baseline • Proportion of participants with HbA1c <7.0% (53 mmol/mol), ≤6.5% (48 mmol/mol) • Mean change in fasting serum glucose (central laboratory) from baseline
• To demonstrate superiority of QW tirzepatide 10 mg and/or 15 mg versus placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i, at 40 weeks for:	• Proportion of participants with HbA1c <5.7% (39 mmol/mol)

Objectives	Endpoints/Estimands
Additional Secondary (not controlled for type 1 error)	
<u>Efficacy</u> • To compare QW tirzepatide 5 mg 10 mg and/or 15 mg versus placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i, at 40 weeks for:	• Proportion of participants with HbA1c <5.7% (39 mmol/mol) • Mean change in daily average 7-point self-monitored blood glucose profiles from baseline • Proportion of participants who achieved weight loss of ≥5%, ≥10%, and ≥15% from baseline • Percentage change from baseline in daily mean insulin glargine dose • Proportion of participants with HbA1c <7.0%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.0 mmol/L [54 mg/dL] or severe hypoglycemia) • Proportion of participants with HbA1c ≤6.5%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.0 mmol/L [54 mg/dL] or severe hypoglycemia) • Proportion of participants with HbA1c <7.0%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.9 mmol/L [70 mg/dL]) • Proportion of participants with HbA1c ≤6.5%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.9 mmol/L [70 mg/dL]) • Proportion of participants with HbA1c <5.7%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.0 mmol/L [54 mg/dL]) • Proportion of participants with HbA1c <5.7%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.9 mmol/L [70 mg/dL])

Approved on 08 Jan 2025 GMT

Objectives	Endpoints/Estimands
Safety • To compare QW tirzepatide 5 mg, 10 mg, and 15 mg to placebo to the end of safety follow up for:	• TEAEs • Early discontinuation of study drug due to AEs • Adjudicated pancreatitis • Serum calcitonin • Incidence of allergic and hypersensitivity reactions • Mean change in systolic and diastolic blood pressure and pulse rate from baseline • Hypoglycemic episodes • Incidence of initiation of rescue therapy for severe, persistent hyperglycemia
Tertiary	
• To compare QW tirzepatide 5 mg, 10 mg, and 15 mg to placebo at 40 weeks for:	• Percentage change in lipids (total cholesterol, HDL, LDL, VLDL, non-HDL and TG) • Percentage change in UACR • Mean change in waist circumference • Changes from baseline in mean body mass index • Patient reported outcomes o Ability to Perform Physical Activities of Daily Living o Impact of Weight on Self-Perception o Diabetes Treatment Satisfaction Questionnaire status/Diabetes Treatment Satisfaction Questionnaire change o European Quality of Life 5 Dimensions 5 level

Abbreviations: AE = adverse events; HbA1c = hemoglobin A1c; HDL = high-density lipoprotein; LDL = low-density lipoprotein; QW = once weekly; SGLT-2i = sodium-glucose transport protein 2 inhibitor; TEAE = treatment emergent adverse events; TG = triglycerides; UACR = urine albumin/creatinine ratio; VLDL = very low-density lipoprotein.

There will be 2 estimands evaluated in this study:

- The efficacy estimand focuses on the treatment effect if participants who underwent randomization continued to receive the study treatment without rescue antihyperglycemic medication for severe, persistent hyperglycemia.
- The treatment-regimen estimand aims at reflecting how participants with T2D are treated in clinical practice and considers both tolerability and efficacy.

Primary estimand

The efficacy estimand answers the following question of interest for the primary objective: What is the treatment difference in hemoglobin A1c (HbA1c) change from baseline to week 40 after randomization in participants with T2DM assuming that participants had stayed on treatment and not taken antihyperglycemic rescue medication?

The efficacy estimand is described by the following attributes.

- *Population*: Participants with T2DM inadequately controlled on insulin glargine alone or in combination with metformin with or without SGLT-2i.
- *Endpoint*: HbA1c change from baseline to Week 40 after randomization.
- *Treatment condition:* The randomized treatment.
- *Intercurrent event*: "Permanent discontinuation from treatment for any reason" and "Initiation of antihyperglycemic rescue treatment" will be addressed using the following strategies:
    - Had participants stayed on treatment (hypothetical strategy).
    - Had participants not taken antihyperglycemic rescue medication (hypothetical strategy).
- *Population-level summary:* Difference in mean changes between treatment conditions.

Secondary estimand

The treatment-regimen estimand answers the following question of interest for the primary objective: What is the treatment difference in change from baseline in HbA1c after 40 weeks of treatment in participants with T2DM regardless of treatment discontinuation for any reason and regardless of initiation of antihyperglycemic rescue?

The treatment-regimen estimand is described by the following attributes.

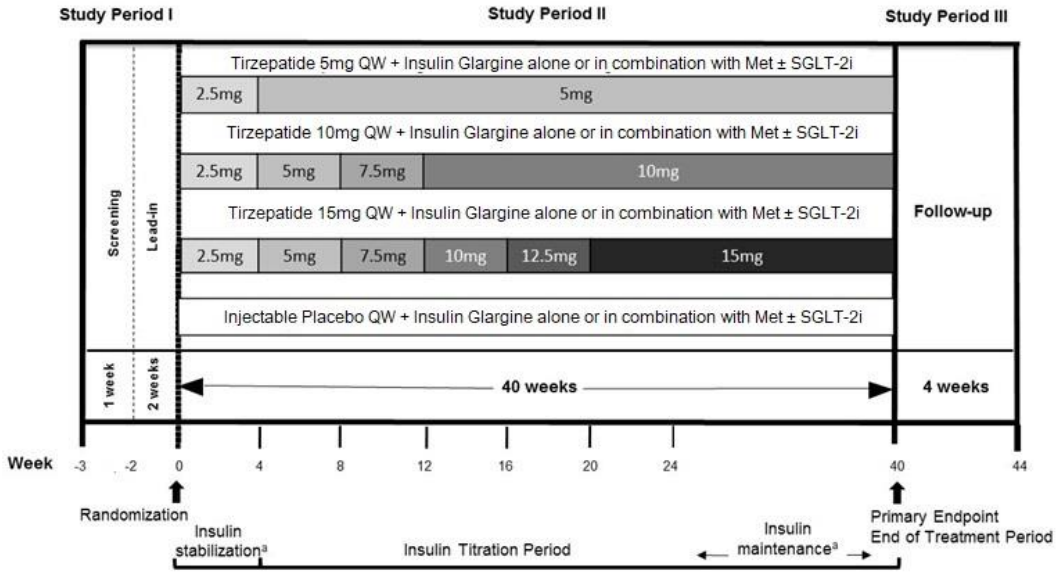
- Population: Participants with T2DM inadequately controlled on insulin glargine alone or in combination with metformin with or without SGLT-2i.
- *Endpoint*: Change from baseline to week 40 after randomization in HbA1c.
- *Treatment condition*: The randomized treatment regardless of adherence to treatment with or without antihyperglycemic rescue medication.
- *Intercurrent events:* "Permanent discontinuation from treatment for any reason" and "initiation of rescue antihyperglycemic treatment" are addressed by the treatment condition.
- *Population-level summary*: Difference in mean changes between treatment conditions.

Both the efficacy and treatment-regimen estimands will be evaluated for the primary and all key secondary objectives. The population, treatment condition, and intercurrent events specified above for each estimand for the primary objective will also apply to the key secondary objectives. The endpoint for each key secondary objective is defined in the table above.

1.2. Study Design

Study GPIM is a multicenter, randomized, double-blind, parallel, placebo-controlled Phase 3 study which will assess the efficacy and safety of the addition of 5 mg, 10 mg, or 15 mg tirzepatide, as compared with placebo in participants with T2DM as an add-on to titrated basal insulin alone or in combination with metformin with or without SGLT-2i over a 40-week treatment. The primary endpoint will be the mean change in HbA1c from baseline to 40 weeks.

This study includes an approximate 3-week screening/lead in period, a 40-week treatment period, and a 4-week safety follow-up (Figure GPIM.1.1).



Abbreviations: Met = metformin; QW = once-weekly; SGLT-2i = sodium-glucose transport protein 2 inhibitor.

[a] Stabilization Period = first 4 weeks after randomization, with restricted insulin dose adjustments.

Basal insulin Titration Period = Weeks 4 to 40 (end of treatment), with unrestricted insulin dose adjustments.

Maintenance Period = Weeks 24 to 40 (end of treatment), the period when basal insulin dose is expected to be stable.

Figure GPIM.1.1. Illustration of study design for clinical protocol I8F-MC-GPIM.

Study Period I (Screening and Lead-in)

Screening (Visit 1)

The purpose of screening procedures at Visit 1 is to establish initial eligibility and to obtain blood samples for laboratory assessments needed to confirm eligibility at Visit 2. Participants who meet all applicable inclusion criteria and none of the applicable exclusion criteria at Visit 1 will continue their pre-study therapy doses between Visits 1 and 2.

Lead-in (Visit 2 to Visit 3)

At Visit 2, the screening laboratory results will be reviewed. For those participants meeting all other eligibility requirements, a dilated fundoscopic examination performed by an ophthalmologist or optometrist, must be completed between Visit 2 and Visit 3 to ensure that participants with proliferative diabetic retinopathy, diabetic macular edema, or nonproliferative diabetic retinopathy who require acute treatment, are identified and not enrolled, unless there is a contraindication to dilate the eye (ie. closed angle glaucoma).

Study Period II (40-Week Treatment Period)

Randomization (Visit 3)

At Visit 3, eligible participants will perform all required baseline study procedures (including the collection of all baseline laboratory measures and ECG) prior to randomization and prior to taking the first dose of study drug.

Treatment period: General Considerations

The treatment period will last 40 weeks, starting with a 4-week stabilization period immediately after randomization and followed by a 36-week insulin glargine titration period. The maintenance period is defined as a part of the titration period when insulin glargine dose is expected to be stable and optimized (Weeks 24 to 40 [Visits 17 to 19]).

The starting dose of tirzepatide will be 2.5 mg QW for 4 weeks, followed by an increase to 5 mg QW, for the duration of the study in the 5 mg group. For the 10 mg group, the starting dose of tirzepatide will be 2.5 mg QW for 4 weeks, then the dose will be increased by 2.5 mg every 4 weeks (2.5 to 5 to 7.5 to 10 mg) until the 10 mg dose is reached and maintained for the duration of the study. For the 15 mg group, the starting dose of tirzepatide will be 2.5 mg QW for 4 weeks, then the dose will be increased by 2.5 mg every 4 weeks (2.5 to 5 to 7.5 to 10 to 12.5 to 15 mg) until the 15 mg dose is reached and maintained for the duration of the study. For the placebo group, patients will inject matched QW placebo for the duration of the study.

Post randomization period (end of Visit 3 to Visit 19):

<u>Stabilization Period (End of Visit 3 through Visit 7 [Weeks 0 through 4])</u>

The main purposes of this period are to introduce randomized study drugs (QW tirzepatide or QW placebo) in a safe manner, to assure regular and correct use of the self-monitoring and insulin dose adjustment procedures, and to assure study diaries are completed correctly. Participants will be encouraged to perform 4-point SMBG profiles twice weekly (generally 3 days apart) during this period. Participants will be required to perform insulin dose assessment per the TTT algorithm twice weekly during this period. In an effort to allow appropriate time for

tirzepatide to reach steady state, insulin glargine dose adjustments during the 4-week stabilization period should be restricted to those needed in the case of significant safety risks due to inadequate insulin dose (hypoglycemia or severe hyperglycemia) in which case patients should be instructed to contact the sites to adjust the insulin glargine dose per the treat-to-target (TTT) algorithm. In addition, for participants with baseline HbA1c ≤8.0%, the insulin glargine dose will be decreased by 20% immediately after randomization, not later than 7 days after the first dose of study drug and will then remain unchanged during the stabilization period to decrease the risk of hypoglycemia. The insulin glargine dose will remain unchanged if baseline HbA1c is >8.0%.

Titration Period (End of Visit 7 through Visit 19 [Weeks 5 through 40])

At the beginning of the titration period, the participant will be instructed to start using the TTT algorithm without restrictions in order to reach the optimal dose of insulin glargine as soon as possible. The participant will be requested to perform insulin dose assessment once weekly during this period.

Visit 99

Visit 99 is only applicable to participants who discontinue the study treatment prematurely before Week 40 and decline to complete the remaining study visits but are willing to return for Visit 99 at Week 40 after randomization. This visit is critical to ensure complete data collection for the primary and key secondary endpoints.

Study Period III (Safety Follow-up Period)

Safety follow-up (Visit 801) visits:

All participants who complete the treatment period are required to complete Visit 801, a safety follow-up visit, approximately 4 weeks after their last visit. Participants discontinuing the study early and performing an early termination (ET) visit will also be asked to perform the safety follow-up visit, so that the safety follow-up visit will be their final visit. During the safety follow-up period, participants will not receive study drug. Initiation of new antihyperglycemic therapy for the safety follow-up period will not be classified as “rescue therapy.”

2. Statistical Hypotheses

The alternative hypotheses for the primary objective are the following:

- $H_{15,1}$: QW tirzepatide 15 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in HbA1c from baseline at 40 weeks.

and/or

- $H_{10,1}$: QW tirzepatide 10 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in HbA1c from baseline at 40 weeks.

The above two hypotheses will be tested in parallel, each at a 2-sided significance level of 0.025.

The alternative hypotheses for the key secondary objective controlling for type 1 error rate are the following:

- $H_{5,1}$: QW tirzepatide 5 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in HbA1c from baseline at 40 weeks.
- $H_{15,2}$: QW tirzepatide 15 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in body weight from baseline at 40 weeks.
- $H_{10,2}$: QW tirzepatide 10 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in body weight from baseline at 40 weeks.
- $H_{5,2}$: QW tirzepatide 5 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in body weight from baseline at 40 weeks.
- $H_{15,3}$: QW tirzepatide 15 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c <7% (53 mmol/mol) at 40 weeks.
- $H_{10,3}$: QW tirzepatide 10 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c <7% (53 mmol/mol) at 40 weeks.
- $H_{5,3}$: QW tirzepatide 5 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c <7% (53 mmol/mol) at 40 weeks.
- $H_{15,4}$: QW tirzepatide 15 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c ≤6.5% (48 mmol/mol) at 40 weeks.
- $H_{10,4}$: QW tirzepatide 10 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c ≤6.5% (48 mmol/mol) at 40 weeks.

Approved on 08 Jan 2025 GMT

- $H_{5,4}$: QW tirzepatide 5 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c ≤6.5% (48 mmol/mol) at 40 weeks.
- $H_{15,5}$: QW tirzepatide 15 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in fasting serum glucose (FSG) from baseline at 40 weeks.
- $H_{10,5}$: QW tirzepatide 10 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in FSG from baseline at 40 weeks.
- $H_{5,5}$: QW tirzepatide 5 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for mean change in FSG from baseline at 40 weeks.
- $H_{15,6}$: QW tirzepatide 15 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c <5.7% (39 mmol/mol) at 40 weeks.
- $H_{10,6}$: QW tirzepatide 10 mg is superior to placebo when added to titrated basal insulin glargine alone or in combination with metformin with or without SGLT-2i for the proportion of participants with HbA1c <5.7% (39 mmol/mol) at 40 weeks.

Operationally the hypotheses will be evaluated by 2-sided tests.

2.1. Multiplicity Adjustment

Since they are intended for different purposes, no type I error rate adjustments will be made for conducting analyses relative to “efficacy” and “treatment-regimen” estimands. For analyses within each estimand, the type I error control strategy for evaluation of primary and key secondary objectives is illustrated in Figure GPIM.2.1.

1. $H_{15,1}$, $H_{15,2,}$ are evaluated hierarchically each at a 2-sided 0.025 significance level conditioned on successfully achieving the preceding objective. In parallel,

2. $H_{10,1}$, $H_{10,2}$ are evaluated hierarchically each at two-sided 0.025 significance level conditioned on successfully achieving the preceding objective.

3.

a) If all objectives in #1 and #2 above are successfully established, $H_{5,1}$, $H_{5,2}$ are evaluated hierarchically, each at a 2-sided 0.05 significance level.
b) If all objectives in only #1 or only #2 above are successfully established, $H_{5,1}$, $H_{5,2}$ are evaluated hierarchically, each at a 2-sided 0.025 significance level.

4. If all objectives: $H_{5,1}$, $H_{5,2}$ are successfully established and

a) if all objectives in #1 and #2 above are successfully established, then $H_{15,3}$, $H_{10,3}$, $H_{5,3}$, $H_{15,4}$, $H_{10,4}$, $H_{5,4}$, $H_{15,5}$, $H_{10,5}$, $H_{5,5}$, $H_{15,6}$, $H_{10,6}$ and $H_{5,6}$ will be evaluated hierarchically each at two-sided 0.05 significance level conditioned on the successfully achieving the preceding objective.
b) If all objectives in only #1 or only #2 above are successfully established, then $H_{15,3}$, $H_{10,3}$, $H_{5,3}$, $H_{15,4}$, $H_{10,4}$, $H_{5,4}$, $H_{15,5}$, $H_{10,5}$, $H_{5,5}$, $H_{15,6}$, $H_{10,6}$ and $H_{5,6}$ will be

c) evaluated hierarchically each at a 2-sided 0.025 significance level conditioned on successfully achieving the preceding objective.

5. If all objectives in #3 and #4 above are successfully established, and at least 1 objective from #1 or #2 above is not successfully established, recycle 100% of the unused alpha back to #1 or #2 above.

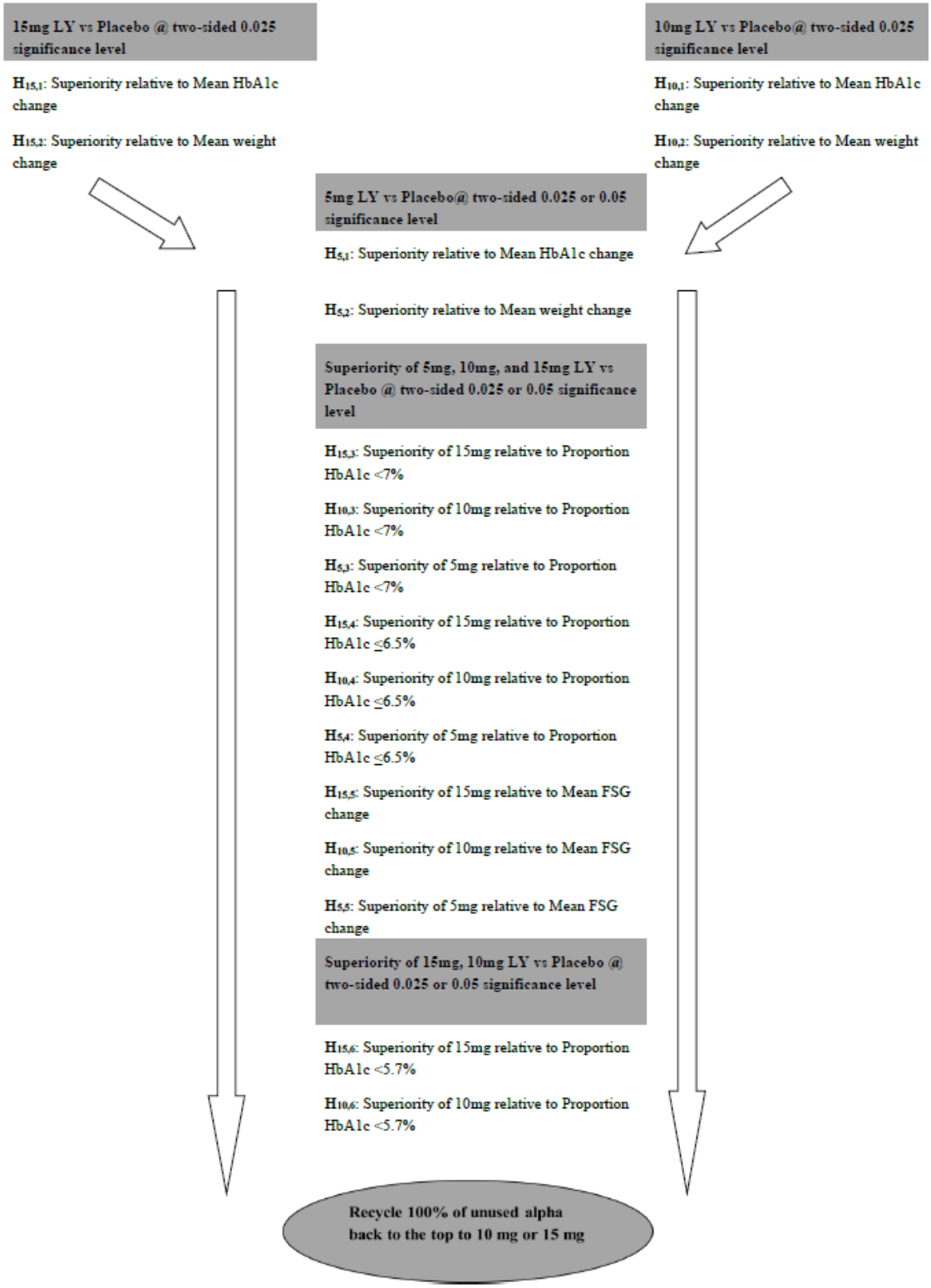


Figure GPIM.2.1 **Type I Error control strategy for primary and key secondary efficacy endpoints.**

3. Analysis Sets

Table GPIM.3.1 defines the analysis population and datasets for the purposes of analysis based on the estimands defined in Section 1.1.

Table GPIM.3.1. Analysis Populations/Data Sets

Population/Analysis Set	Description
Screened population	All participants who signed informed consent.
Randomized population	All participants who are randomly assigned to a treatment group.
Modified intent-to-treat population (mITT)	All randomly assigned participants who are exposed to at least 1 dose of study intervention. Participants will be analyzed according to the treatment they were randomly assigned to regardless of the treatment actually received.
Efficacy Analysis Set: This analysis set will be used to estimate the efficacy estimand for the primary and key secondary objectives	Data obtained during the Treatment Period from the mITT population, excluding participants discontinuing study drug due to inadvertent enrollment, excluding data after permanent discontinuation of treatment or initiation of antihyperglycemic rescue medication.
Full Analysis Set: This analysis set will be used to estimate the treatment-regimen estimand for the primary and key secondary objectives	Data obtained during the treatment period from the mITT population, excluding participants discontinuing study drug due to inadvertent enrollment, regardless of adherence to treatment or initiation of antihyperglycemic rescue medication.
Safety Analysis Set (SS): This analysis will be used to assess the safety of study treatment	Data obtained during the Treatment Period plus Safety Follow-Up from the mITT population, regardless of adherence to treatment or initiation of antihyperglycemic rescue antihyperglycemic medication.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee. The statistical analyses will be performed using SAS® (Version 9.4 or higher) and R® (Version 4.3.2 or higher).

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the statistical analysis plan or clinical study report. Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (for example, few events to justify conducting an analysis). Listings of events will be provided in such situations. The primary analysis will be performed at database release after Last Participant Last Visit (LPLV). Additional analyses of the data may be conducted as deemed appropriate without further changes made to the protocol or SAP, even after the primary or final database locks.

Additionally, to avoid potential selection biases, unless stated otherwise, statistical summaries and analyses will be conducted based on randomized maintenance dose regardless of the actual treatment received by the patient.

Unless specified otherwise, the last measurement during Visit 1 to Visit 3 (including unscheduled visits) collected prior to or on the first dose day will serve as baseline.

- For lab and ECG, baseline needs to be prior to or within one hour after the first dose time.
- For patient-reported outcomes, data collected at Visit 3, regardless of the timing relative to first dose, will serve as baseline.

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated, and all confidence intervals will be given at a 2-sided [95%] level. To avoid potential selection biases, unless stated otherwise, in statistical summaries and analyses, all data will be analyzed by randomized treatment assignment.

There will be 2 estimands of interest in evaluating the primary and secondary efficacy objectives. The first estimand, the “efficacy” estimand, represents efficacy prior to discontinuation of study drug without the confounding effects of rescue therapy for severe persistent hyperglycemia. Analyses relative to the “efficacy” estimand will be conducted using the efficacy analysis set (EAS). The second estimand, the “treatment-regimen” estimand, represents the efficacy irrespective of adherence to investigational product or introduction of rescue therapy for severe persistent hyperglycemia. Analyses relative to the “treatment-regimen” estimand will be conducted using the full analysis set (FAS).

Safety will be assessed using Safety Analysis Set. Selected safety analyses may be conducted after excluding data on rescue therapy or data after starting another antihyperglycemic medication.

A listing of patients randomized but not included in efficacy analyses (not treated or discontinued treatment due to inadvertent enrollment) will be provided.

End of study participation for a patient will be the earliest of date of death, date of withdrawal from further participation in the study, or date of safety follow up visit (Visit 801). For patients considered to be lost to follow-up, end of study participation will be the date of lost to follow-up reported by the investigator. Patient data included in the database after the last date of study participation (date of death, date of early discontinuation or date of safety follow-up) will be excluded from statistical analysis. Listings of such data may be provided.

Summary statistics for continuous measures may include sample size, mean, standard deviation, median, minimum, and maximum.

The Kaplan-Meier method may be used for estimation of cumulative event-free survival rates over time, and Cox proportional hazards regression analysis will be used to compare hazard rates among treatments.

Summary statistics for categorical measures, including categorized continuous measures, will include sample size, frequency, and percentages. Fisher's exact test will be used to examine the treatment difference in categorical outcomes. Logistic regression may be used to examine the treatment difference in binary efficacy outcomes. The negative binomial regression model will be used for the treatment comparison of discrete count measures if deemed appropriate.

Statistical treatment comparisons will only be performed between tirzepatide doses and placebo. Comparisons among tirzepatide arms will not be performed unless specified otherwise.

Statistical summaries and results of statistical analyses will be displayed in the following treatment order: tirzepatide 5 mg, tirzepatide 10 mg, tirzepatide 15 mg , placebo.

4.2. Participant Dispositions

Reasons for screen failures as reported by investigators will be summarized.

A listing of final study disposition and a listing of randomized treatment assignment (planned treatment) for all randomized patients will be provided. The number and percentage of participants prematurely discontinuing study treatment and study will be provided by study treatment. Reasons for prematurely discontinuing study treatment and study will be provided by study treatment.

4.3. Primary Endpoint(s) Analysis

4.3.1. Definition of endpoint(s)

The primary endpoint for this study is mean HbA1c change from baseline at 40 weeks This endpoint will be used to evaluate the primary objective of the study for both the treatment-regimen and the efficacy estimands (Section 1.1).

The primary efficacy measure will be change in HbA1c (% and mmol/mol) from baseline (postbaseline – baseline). Both HbA1c values as well as change from baseline in HbA1c will be summarized by treatment and nominal visit (week). If scheduled HbA1c data at the primary endpoint visit is not available, unscheduled HbA1c data collected for the primary endpoint visit will be included in the analysis.

Approved on 08 Jan 2025 GMT

4.3.2. Primary Analysis Relative to the Efficacy Estimand

The primary analysis relative to the "efficacy" estimand will be conducted using HbA1c data in the EAS from baseline through the 40 -week visit with the aid of a mixed-model repeated measure (MMRM). Restricted maximum likelihood (REML) will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate denominator degrees of freedom. The response variable of the MMRM model will be change in HbA1c values from baseline obtained at each scheduled postbaseline visit. The independent variables of the MMRM model are treatment, visit, treatment-by-visit interaction, SGLT-2i use (Yes or No) as fixed effects, and baseline HbA1c as covariates. Missing data will be addressed by the MMRM model. No explicit imputation methods for missing data will be employed. An unstructured covariance matrix will be used to model the within-patient errors. If this model fails to converge, the following covariance structures will be tested in the following order:

1. Heterogeneous Toeplitz
2. Heterogeneous First Order Autoregressive
3. Heterogeneous Compound Symmetry
4. Toeplitz
5. First Order Autoregressive, and
6. Compound Symmetry.

The first covariance structure that converges will be used. The resulting least squares mean (LSM) estimate of mean change from baseline in HbA1c will be summarized by visit and by study treatment.

With the aid of the MMRM analysis, p-values, and 2-sided 95% confidence intervals (CIs) for mean change in HbA1c from baseline to the 40-week visit will be derived and summarized for the 5 mg, 10 mg, and 15 mg doses of tirzepatide compared to placebo.

4.3.3. Primary Analysis Relative to the Treatment-Regimen Estimand

The primary analysis relative to the treatment-regimen estimand will be conducted utilizing HbA1c data in the FAS at baseline and at the 40-week visit with the aid of an ANCOVA model. The response variable will be the primary measure and model terms will include treatment, SGLT-2i use (Yes or No) as fixed effects and baseline HbA1c as a covariate. The ANCOVA analysis will be conducted using multiple imputation of missing primary measures (see Section 4.3.4 for details) and statistical inference over multiple imputation of missing data guided by Rubin (1987). With the aid of the ANCOVA analysis, p-values and 2-sided 95% CIs for mean change in HbA1c from baseline to the 40-week visit will be derived and summarized for the 5 mg, 10 mg, and 15 mg doses of tirzepatide compared to placebo.

4.3.4. Methods for Multiple Imputations

For efficacy analyses relative to the "treatment-regimen" estimand, missing HbA1c data at the 40-week visit will be imputed based on "retrieved dropouts," defined as patients who had their HbA1c value measured at the 40-week visit in the same treatment arm who prematurely

discontinued study drug. A pseudo-SAS program for implementing multiple imputations using data from retrieved dropouts is included in Section 7.1 (Appendix 1). In cases where there are not enough retrieved dropouts (i.e. for any group with ≤4 discontinued patients who have data recorded at the 40-week visit) to provide a reliable imputation model, an alternative multiple imputation method with reference to the placebo group (placebo multiple imputation) will be used as the primary analysis relative to the treatment-regimen estimand. Analyses will be conducted with multiple imputations, and statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987). If the primary analysis using "retrieved dropouts" for imputation converges, the analysis using "placebo multiple imputation" to impute the missing data will be conducted as a sensitivity analysis. If value of the imputed HbA1c change from baseline is <-6.0% or >6.0%, that value will be set to -6.0% or 6.0%, respectively, to avoid unrealistic imputed values.

4.4. Secondary Endpoint(s) Analysis

4.4.1. Key Secondary Endpoints

The endpoints corresponding to secondary study objectives subject to type 1 error rate control are specified in Section 1.1 under "Key Secondary (Controlled for Type 1 error) endpoints".

The alternative hypotheses corresponding to the key secondary objectives can be found in Section 2.

Key secondary objectives will be evaluated based on the treatment-regimen and the efficacy estimands (Section 1.1), similar to the primary objective.

4.4.1.1. Mean change in HbA1c from Baseline at the 40-week visit

Assessment of superiority of mean changes in HbA1c of tirzepatide 5 mg compared with versus placebo at Week 40 will be conducted using the same statistical models as those used for evaluating the primary objective in Section 4.3.2 (efficacy estimand) and Section 4.3.3 (treatment-regimen estimand). Two-sided p-values will be reported for comparisons between tirzepatide and placebo. Inferences will be made based on the multiplicity testing framework outlined in Section 2.1.

4.4.1.2. Mean Change in Body Weight from Baseline at the 40-week Visit

The analysis for change in body weight from baseline (postbaseline - baseline) at Week 40 will be conducted in a manner similar to the primary analysis in Section 4.3.2 (efficacy estimand) and Section 4.3.3 (treatment-regimen estimand) with the following differences:

- For the efficacy estimand, the MMRM model (Section 4.3.2) will be updated by adding the baseline category for HbA1c (≤8.0%,>8.0%) as a fixed factor in place of baseline HbA1c, and baseline weight as an additional covariate.
- For the treatment-regimen estimand, the ANCOVA model (Section 4.3.3), baseline category for HbA1c (≤8.0%,>8.0%) will be used as a fixed factor in place of baseline HbA1c as a covariate and baseline weight will be used as an additional covariate in the statistical model. Imputation of missing values at either Week 40 will be done in a similar manner as described in Section 4.3.4. If value of the imputed weight change from

baseline is <-50 kg or >50 kg, that value will be set to -50 kg or 50 kg, respectively, to avoid unrealistic imputed values.

4.4.1.3. Mean change in FSG from Baseline at the 40-week Visit

The analysis for change in FSG from baseline (postbaseline - baseline) at Week 40 will be conducted in a manner similar to the primary analysis in Section 4.3.2 (efficacy estimand) and Section 4.3.3 (treatment-regimen estimand) with the following differences:

- For the efficacy estimand, the MMRM model (Section 4.3.2) will be updated by adding the baseline category for HbA1c (≤8.0%,>8.0%) as a fixed factor in place of baseline HbA1c, and baseline FSG as an additional covariate.
- For the treatment-regimen estimand, the ANCOVA model (Section 4.3.3), baseline category for HbA1c (≤8.0%,>8.0%) will be used as a fixed factor in place of baseline HbA1c as a covariate and baseline FSG will be used as an additional covariate in the statistical model. Imputation of missing values at either Week 40 will be done in a similar manner as described in Section 4.3.4. If value of the imputed fasting serum glucose change from baseline is <-20 mmol/L or > 20 mmol/L, that value will be set to -20 mmol/L or 20 mmol/L, respectively, to avoid unrealistic imputed values.

4.4.1.4. Proportion of participants with HbA1c <7.0%, ≤6.5%, <5.7%

The analysis for binary outcomes relative to the "efficacy" estimand will be performed using EAS with missing values imputed by the predicted value from the MMRM model specified in Section 4.3.2 and dichotomized data will then be derived based on continuous imputed values. After dichotomizing, the data is analyzed using a logistic regression model which include treatment, SGLT-2i use (Yes or No) as fixed effects, and baseline HbA1c as covariates. In addition, analysis will be conducted utilizing data using EAS from baseline through the 40-week visit with the aid of a longitudinal logistic regression with repeated measurements with treatment, visit, treatment-by-visit interaction and SGLT-2i use (Yes/No) as fixed effects and baseline HbA1c as a covariate. In case the longitudinal logistic model does not converge due to small number of events, logistic regression will be utilized to analyze proportion of patients achieving HbA1c targets at nominal visits.

The analysis for binary outcomes relative to the "treatment-regimen" estimand will be performed using FAS at baseline and at the 40-week visits with the aid of a logistic regression with multiple imputation of missing data at the 40-week visit (see Section 4.3.4 for details) and dichotomized data will then be derived based on continuous imputed values. After dichotomizing, the data is analyzed using a logistic regression model which include treatment, SGLT-2i use (Yes or No) as fixed effects, and baseline HbA1c as covariates. Statistical inference over multiple imputations will be guided by Rubin (1987).

4.4.2. Additional Secondary Endpoints

Additional secondary endpoints specified below will use the efficacy estimand and will be summarized by treatment and nominal visit.

- Proportion of participants with HbA1c <5.7% (39 mmol/mol)

- Mean change in daily average 7-point self-monitored blood glucose profiles from baseline
- Proportion of participants who achieved weight loss of ≥5%, ≥10%, and ≥15% from baseline
- Percentage change from baseline in daily mean insulin glargine dose
- Proportion of participants with HbA1c <7.0%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.0 mmol/L [54 mg/dL] or severe hypoglycemia)
- Proportion of participants with HbA1c ≤6.5%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.0 mmol/L [54 mg/dL] or severe hypoglycemia)
- Proportion of participants with HbA1c <7.0%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.9 mmol/L [70 mg/dL])
- Proportion of participants with HbA1c ≤6.5%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.9 mmol/L [70 mg/dL])
- Proportion of participants with HbA1c <5.7%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.0 mmol/L [54 mg/dL])
- Proportion of participants with HbA1c <5.7%, without weight gain (<0.1 kg) and without hypoglycemia (blood glucose <3.9 mmol/L [70 mg/dL])

4.4.2.1. Analysis for Continuous Outcomes

The analysis to make comparisons between treatment groups for additional secondary continuous outcomes relative to the "efficacy" estimand will be conducted similar to the primary efficacy analysis in Section 4.3.2. The MMRM model will be updated by adding the baseline category for HbA1c (≤8.0%, >8.0%) as a fixed factor in place of baseline HbA1c, and corresponding baseline values as an additional covariate. Missing data will be addressed by the MMRM model and no explicit imputation methods will be used. Daily mean insulin glargine dose may be transformed.

4.4.2.2. Analysis for Binary Outcomes

The analysis for binary outcomes relative to the "efficacy" estimand will be performed using EAS with missing values imputed by the predicted value from the MMRM model specified in Section 4.3.2 and dichotomized data will then be derived based on continuous imputed values. After dichotomizing, the data is analyzed using a logistic regression model.

In addition, analysis will be conducted utilizing data using EAS from baseline through the 40-week visit with the aid of a logistic regression with repeated measurements with treatment, visit, treatment-by-visit interaction, SGLT-2i use (Yes or No) and baseline category for HbA1c (≤8.0%,>8.0%) as fixed effect, and baseline of the corresponding variable and age as covariates.

- For analysis of HbA1c binary endpoints, the logistic model will include treatment, SGLT-2i use (Yes or No) as fixed effects, and baseline HbA1c as covariates.
- For analysis of weight related binary endpoints, the logistic model will include treatment, SGLT-2i use (Yes or No), and baseline category for HbA1c (≤8.0%,>8.0%) as fixed effects, and baseline weight as covariates.
- For analysis of composite binary endpoints, the longitudinal logistic regression model will include treatment and SGLT-2i use (Yes or No) as fixed effects, and baseline HbA1c, baseline weight as covariates

4.5. Exploratory Endpoints

Other exploratory (prespecified/non-prespecified in the protocol) efficacy measures will use the efficacy estimand and will be summarized by treatment and nominal visit. Unless otherwise specified, missing data will not be explicitly imputed, and assessments are not subject to type 1 error rate control. Some parameters may be log transformed, if necessary.

Table GPIM.4.1. Exploratory Efficacy Analyses Not Controlled for Type I Error

Objective	Relative to the Efficacy Measure	Analysis Conducted in a Manner Similar to Section	Additional Information
Exploratory Objectives			
To compare QW tirzepatide 5 mg, 10 mg, and 15 mg to placebo at 40 weeks for:	Percentage change in lipids (total cholesterol, HDL, LDL, VLDL, non-HDL and TG) from baseline	MMRM model in Section 4.4.2.1	Use corresponding baseline parameter as a covariate. Model based on log transformation.
	Percentage change in UACR from baseline	ANCOVA model in Section 4.3.3	Use baseline UACR as a covariate. Model based on log transformation.
	Change in waist circumference from baseline	MMRM model in Section 4.4.2.1	Use baseline waist circumference as a covariate.
	Change in BMI from baseline	MMRM model in Section 4.4.2.1	Use baseline BMI as a covariate.
	Change in ability to Perform Physical Activities of Daily Living (APPADL) from baseline	ANCOVA model in Section 4.3.3	Use baseline Physical Activities of Daily Living (APPADL) score as a covariate.
	Change in impact of Weight on Self Perception (IW-SP)	ANCOVA model in Section 4.3.3	Use baseline impact of Weight on Self Perception (IW-SP) score as a covariate.
	Diabetes Treatment Satisfaction Questionnaire change (DTSQc)	ANCOVA model in Section 4.3.3	Use DTSQs as a covariate.
	Change in EQ-5D-5L	ANCOVA model in Section 4.3.3	Use baseline EQ-5D-5L score as a covariate.

Abbreviations: APPADL = Ability to Perform Physical Activities of Daily Living; BMI = body mass index; FSG = fasting serum glucose; HbA1c = hemoglobin A1c; HDL-C = high-density lipoprotein-cholesterol; IW-SP = Impact of Weight on Self-Perceptions Questionnaire; LDL-C = low-density lipoprotein cholesterol; QW = once-weekly; TG = triglyceride; VLDL-C = very-low density lipoprotein-cholesterol; EQ-5D-5L= European Quality of Life5 Dimensions5 level; DTSQs = Diabetes Treatment Satisfaction Questionnaire Status; UACR=urinary albumin to creatinine ratio

Note: Exploratory Efficacy Analyses Not Controlled for Type I Error.

4.5.1. Exploratory Analysis not Specified in the Protocol

Table GPIM.4.2. Non-protocol Specified Exploratory Efficacy Analyses Not Controlled for Type I Error

Objective	Relative to the Efficacy Measure	Analysis Conducted in a Manner Similar to Section	Additional Information
To compare QW tirzepatide 5 mg, 10 mg, and 15 mg to placebo at 40 weeks for:	Change in self-monitored blood glucose profiles • Morning premeal (fasting) • Morning 2-hour postmeal • Midday premeal • Midday 2-hour postmeal • Evening premeal • Evening 2-hour postmeal • Bedtime • Morning premeal to 2-hour postmeal excursion • Midday premeal to 2-hour postmeal excursion • Evening premeal to 2-hour postmeal excursion • the mean of all meals 2-hour excursion • the mean of all 7 point measurements • the mean of all pre-meal measurements • the mean of all 2-hour postprandial measurements	MMRM model in Section 4.4.2.1	Use corresponding baseline parameter as a covariate.
	• Change in eGFR from baseline	MMRM model in Section 4.4.2.1	Use baseline eGFR as a covariate.
	• Percentage of participants who achieved the following composite endpoints where No-Hypo is defined as BG <54 mg/dL (<3.0 mmol/L) and/or severe hypoglycemia: ○ HbA1c<5.7% & Weight Loss ≥5% & No-Hypo ○ HbA1c <5.7% & Weight Loss ≥10% & No-Hypo	Longitudinal logistic model in Section 4.4.2.2	None

Approved on 08 Jan 2025 GMT

Objective	Relative to the Efficacy Measure	Analysis Conducted in a Manner Similar to Section	Additional Information
	○ HbA1c <5.7% & Weight Loss ≥15% & No-Hypo ○ HbA1c <5.7% & Weight Gain <0.1kg & No-Hypo ○ HbA1c ≤6.5% & Weight Loss ≥5% & No-Hypo ○ HbA1c ≤6.5% & Weight Loss ≥10% & No-Hypo ○ HbA1c ≤6.5% & Weight Loss ≥15% & No-Hypo ○ HbA1c <7.0% & Weight Loss ≥5% & No-Hypo ○ HbA1c <7.0% & Weight Loss ≥10% & No-Hypo ○ HbA1c <7.0% & Weight Loss ≥15% & No-Hypo		

4.6. Safety Analyses

Unless specified otherwise, safety assessments will be conducted based on the Safety Analysis Set (see Section 3) irrespective of adherence to study intervention or initiation of rescue antihyperglycemic therapy. Selected safety analyses may be conducted after excluding data on rescue therapy or data after starting another antihyperglycemic medication. For rare events (<10 patients have the events), summary tables may not be generated, and individual patient level data will be listed.

4.6.1. Extent of Exposure

A listing of patients randomized but not receiving study treatment will be provided, if applicable. The listing will include patient identification, randomized treatment arm, and the reason for not receiving study treatment (if the data is available). A summary of duration to follow-up (defined as time in days from date of randomization to date of safety follow-up, date of early study discontinuation or date of death) and duration on study treatment (defined as time in days from date of first dose of study treatment to date of last dose of study treatment plus 7 days) will be provided by study treatment.

Time-to-event analysis of premature study treatment discontinuation will be conducted.

A listing of patients who re-initiate tirzepatide/placebo due to missing ≥3 consecutive doses may be produced if data warrants.

4.6.2. Adverse Events

Adverse events will be coded from the actual term using the Medical Dictionary for Regulatory Activities and reported with preferred terms and system organ class.

A TEAE is defined as medical occurrence that emerges during a defined treatment period, having been absent pre-treatment, or worsens relative to the pre-treatment state, and does not necessarily have to have a causal relationship with this treatment. The maximum severity for each low level term (LLT) during the baseline period including ongoing medical history will be used as baseline severity. For events with a missing severity during the baseline period, it will be treated as 'mild' in severity for determining treatment-emergence. Events with a missing severity during the postbaseline period will be treated as 'severe' and treatment-emergence will be determined by comparing to baseline severity.

The percentages of patients with TEAEs will be summarized by treatment using MedDRA preferred term (PT) nested within system organ class (SOC) and Fisher's exact test will be used to compare the treatment groups at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC. For events that are sex-specific, the denominator and computation of the percentage will include only patients from the given sex.

Overview of the number and percentage of patient who experienced a TEAE, serious adverse event (SAE), death, discontinued from study treatment, or study due to an AE, and relationship to study drug, will be summarized by treatment.

The percentages of patients with TEAEs, overall and common (common TEAEs occurred in ≥5% of patients before rounding), will be summarized by treatment using MedDRA PT. Events will be ordered by decreasing frequency.

The percentages of patients with TEAEs by maximum severity will be summarized by treatment using MedDRA PT. For each patient 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. If all severities are missing for the defined postbaseline period of interest, it will show as missing in the table. Only counts and percentages will be included for the TEAEs by maximum severity.

Patient narratives will be provided for all patients who experience any of the following "notable" events:

- Deaths
- SAEs
- Permanent Discontinuations of study treatment due to AEs, or
- Severe adverse events of special interest

4.6.2.1. Deaths

A listing of all deaths will be provided. The listing will include patient identification including the treatment, site number, date of death, age at the time of enrollment, gender, MedDRA PT of associated AE, time from first dose of study drug to death, time from last dose of study drug to death (if patient had discontinued study drug), cause of death as reported by investigator, cause of death as adjudicated by Clinical Endpoint Committee (CEC).

Approved on 08 Jan 2025 GMT

4.6.2.2. Serious Adverse Events

The number and percentage of patients who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) during the study including the follow-up period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all SAEs will be provided. Listing will include but not limited to treatment, patient identification including the site number, treatment group, date of event, age at the time of enrollment, gender, MedDRA SOC and PT, severity, action taken, outcome, relationship to study drug, time from first dose of study drug to the event, and event duration.

4.6.2.3. Discontinuation from Study Due to Adverse Event

The number and percentage of patients who prematurely discontinue the study due to an AE will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all the discontinuation from Study due to Adverse Event will be provided.

4.6.2.4. Discontinuation from Study Treatment Due to Adverse Event

The number and percentage of patients who prematurely discontinue study drug due to an AE will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all the discontinuation from Study Treatment due to Adverse Event will be provided. A time-to-event analysis will be conducted by treatment on time to study drug discontinuation as well as on time to study drug discontinuation due to an AE.

4.6.2.5. Treatment of Overdose

Study drug overdose (more than the specified number of injections) will be reported as an AE. A listing of patients reporting over-dosing of tirzepatide will be provided as a protocol deviation.

4.6.3. Special Safety Topics

4.6.3.1. Hypoglycemic Events

Definitions of different categories of hypoglycemic events are mentioned below-

Level 1

Glucose <70 mg/dL (<3.9 mmol/L) and ≥54 mg/dL (≥3.0 mmol/L)

Level 1 hypoglycemia can alert a person to take action such as treatment with fast-acting carbohydrates. Providers should continue to counsel participants to treat hypoglycemia at this glucose alert value.

Level 2

Glucose <54 mg/dL (<3.0 mmol/L)

Level 2 hypoglycemia is also referred to as documented or blood glucose-confirmed hypoglycemia with glucose <54 mg/dL (<3.0 mmol/L). This glucose threshold is clinically relevant regardless of the presence or absence of symptoms of hypoglycemia.

Level 3 Severe

A severe hypoglycemic event is characterized by altered mental or physical status requiring assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions for the treatment of hypoglycemia.

The determination of 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.

Examples of severe hypoglycemia in adults are-

- altered mental status and the inability to assist in their own care
- semiconscious or unconscious, or
- coma with or without seizures.

Nocturnal hypoglycemia: Nocturnal hypoglycemia is a hypoglycemia event, including severe hypoglycemia, which **occurs at night** and presumably during sleep.

To avoid duplicate reporting, all consecutive blood glucose values <70 mg/dL (3.9 mmol/L) occurring within a 1-hour period may be considered to be a single hypoglycemic event and the one with the lowest blood glucose value can be selected to be the representative. If two records with the same lowest values within an hour, the earlier occurrence will be selected.

A listing of level 2 and level 3 hypoglycemic events will be provided. Statistical summaries and analyses will exclude hypoglycemic events occurring after initiation of a new antihyperglycemic therapy. For severe hypoglycemia, level 1 and level 2 hypoglycemia, incidence as well as rate per patient year of exposure will be provided by treatment at specified time intervals. If data warrants, additional statistical analyses will be conducted on hypoglycemic incidence and rate.

The incidence of hypoglycemic event will be analyzed using logistic regression with treatment, SGLT-2i use (Yes or No), baseline HbA1C category (≤8.0%,>8.0%) as fixed effects. The rate of hypoglycemic episodes per patient year may be analyzed using a generalized linear mixed-effects model assuming the number of hypoglycemic episodes follows a negative binomial distribution with the mean modeled using baseline HbA1C category (≤8.0%,>8.0%), SGLT-2i use (Yes or No), and treatment as fixed effects if data warrants. The logarithm of years in specified time interval will be adjusted as an offset to account for possible unequal treatment duration in the specified time interval between patients. When the number of hypoglycemic events is less than 10, the listing of hypoglycemic events will be provided instead.

4.6.3.2. Severe, Persistent Hyperglycemia

A summary statistic of initiation of rescue therapy in response to severe, persistent hyperglycemia will be provided by treatment. A listing of patients who initiated rescue therapy will be provided.

4.6.3.3. Pancreatitis

If data warrants, summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Determination of investigator-reported events will be through the pre-defined SMQ search for acute pancreatitis and MedDRA PT of pancreatitis chronic. Detailed searching criteria can be found in Appendix 2 (Section 7.2).

4.6.3.3.1. Pancreatic Hyperenzymeia

Observed pancreatic enzyme data (p-amylase and lipase) will be summarized and listed by treatment and nominal visit.

Additionally, the number and proportion of patients with maximum postbaseline pancreatic enzyme values exceeding the following thresholds will be provided by maximum baseline pancreatic enzyme value (≤1×upper limit of normal (ULN), >1×ULN), and treatment: ≤1×ULN, (>1 to ≤3) × ULN, (>3 to ≤5) × ULN, (>5 to ≤10) × ULN, >10 × ULN.

An MMRM analysis will be used to analyze each pancreatic enzyme with a log transformed (postbaseline measure/baseline measure) response variable and SGLT-2i use (Yes/No), baseline HbA1c category (≤8.0%, >8.0% [≤64, > 64 mmol/mol]), treatment, visit, and treatment-by-nominal visit interaction as fixed effects, and baseline value as a covariate.

4.6.3.4. Thyroid Malignancies and C-Cell Neoplasms

Treatment-emergent thyroid malignancies and C-Cell Neoplasms will be identified using pre-defined MedDRA High Level Terms (HLTs) of thyroid neoplasms malignant and PT of thyroid C-cell hyperplasia. Detailed searching criteria can be found in Appendix 2 (Section 7.2). A summary and a listing by treatment and PT will be provided.

4.6.3.4.1. Calcitonin

Observed calcitonin data will be summarized by treatment and nominal visit. Additionally, the number and proportion of patients with a maximum postbaseline calcitonin value exceeding the following thresholds will be provided by treatment and maximum baseline calcitonin value (≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L): ≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L to ≤50 ng/L, >50 ng/L to ≤100 ng/L, > 100 ng/L.

4.6.3.5. Malignancies

The AE database will be searched using pre-defined SMQs to identify events consistent with malignancy. Detailed searching criteria can be found in Appendix 2 (Section 7.2). A listing of TEAEs will be provided. Malignancy will be considered as an AESI.

4.6.3.6. Major Adverse Cardiovascular Events

Deaths and nonfatal Cardiovascular Events AEs will be adjudicated by a committee of physicians external to Lilly with cardiology expertise. The nonfatal Cardiovascular Events AEs to be adjudicated include the following:

- Myocardial infarction
- Hospitalization for unstable angina
- Hospitalization for heart failure
- Coronary interventions, such as coronary artery bypass graft or percutaneous coronary intervention, and
- Cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack.

The number and proportion of patients with Major Adverse Cardiovascular Events will be reported by treatment. Listing of deaths, myocardial infarctions, strokes, and hospitalization for unstable angina confirmed by an independent clinical endpoint committee will be provided. The dates of randomization, event, first dose and last dose of study intervention, and time from randomization to event will be listed.

4.6.3.7. Arrhythmias and Cardiac Conduction Disorders

The AE database will be searched using pre-defined standardized MedDRA query (SMQ) or MeDRA HLT to identify events consistent with supraventricular arrhythmias and cardiac conduction disorders. Detailed searching criteria can be found in Appendix 2 (Section 7.2). Incidence of the resulting TEAEs will be summarized by treatment and PT within SMQ and HLT.

4.6.3.8. Hypersensitivity Events

Hypersensitivity reactions and related information will be summarized by treatment. Two main analyses are performed:

- **Potential Immediate Hypersensitivity:** Analysis of TEAEs occurring from the start of study drug administration up to 24 hours after the end of study drug administration. For events without the hypersensitivity electronic case report form (eCRF), only date (no time) information are collected, the events occurred on the same date as the study drug injection date will be included.
- **Potential Non-Immediate Hypersensitivity:** Analysis of TEAEs occurring more than 24 hours after the end of study drug administration, but prior to subsequent study drug administration.

Summaries of all potential hypersensitivity reactions will be generated by PT with decreasing frequency by treatment. Detailed searching criteria can be found in Appendix 2 (Section 7.2).

4.6.3.9. Injection Site Reactions

Injection site reactions, incidence, and related information reported via the "Injection Site Reactions" eCRF will be summarized by treatment. Information to be summarized includes the timing of the reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, pruritis, and edema. Patient based and event-based summaries will be created.

Additionally, potential injection site reactions will be searched by pre-defined MedDRA HLTs of injection site reactions, administration site reactions, and infusion related reactions. Detailed searching criteria for injection site reaction events can be found in Appendix 2 (Section 7.2). The PT will be used for summary by treatment within each HLT category.

4.6.3.10. Diabetic Retinopathy Complications

Results of the baseline dilated fundoscopic exam will be summarized by treatment and by visit at Baseline. A listing of unscheduled visits only will be provided. Any TEAE suspected of worsening retinopathy triggers a follow-up dilated fundoscopic exam. A summary of TEAEs suspected of worsening retinopathy and a summary of the results of the follow-up dilated

fundoscopic exam will be summarized by treatment and PT. The cases with repeated fundoscopy during the course of the trial, based on clinical suspicion of worsening retinopathy that have either findings of de novo retinopathy or progression of retinopathy, and severe/serious adverse events from the PTs defined in searching criteria in Appendix 2 (Section 7.2) will be summarized.

4.6.3.11. Hepatic Safety

4.6.3.11.1. Hepatobiliary Disorders

The AE database will be searched using standardized MedDRA query SMQs to identify events consistent with hepatobiliary disorders. Detailed searching criteria can be found in Appendix 2 (Section 7.2). A summary by treatment and PT within SMQ will be provided.

4.6.3.11.2. Acute Gallbladder Disease

The AE database will be searched using pre-defined SMQs to identify events consistent with acute gallbladder diseases. Detailed searching criteria for these AEs can be found in Appendix 2 (Section 7.2). A summary by treatment and PT within SMQ will be provided.

4.6.3.11.3. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 4.6.4. This section describes additional analyses of liver enzymes. In addition, the following will be provided by treatment group:

- A shift table of maximum to maximum alanine aminotransferase (ALT) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN.
- A shift table of maximum to maximum aspartate transaminase (AST) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN.
- Shift tables of maximum to maximum total bilirubin and direct bilirubin from baseline to postbaseline with the following categories: ≤1 × ULN, > 1 to <2 × ULN, ≥2 × ULN.
- Shift tables of serum alkaline phosphatase from baseline to postbaseline with the following categories: ≤1 × ULN, >1 to <2 × ULN, ≥2 × ULN.

Maximum baseline will be the maximum non-missing observation in the baseline period. The maximum postbaseline value will be the maximum non-missing value from the postbaseline period. Planned and unplanned measurements will be included.

4.6.3.12. Gastrointestinal Safety

The time courses of prevalence and incidence (newly-occurring episodes) of nausea, vomiting, diarrhea, and combined will be plotted by treatment and maximum severity.

The maximum severity and duration of treatment-emergent nausea, vomiting, diarrhea, and combined through the end of the study will be summarized by treatment.

The PTs in the gastrointestinal SOC will be used to identify gastrointestinal AEs. The incidence of the resulting TEAEs will be summarized by treatment and PT.

4.6.3.13. Acute Renal Events

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 4.6.3.13.

Two shift tables examining renal function will be created. A min-to-min shift table of eGFR estimated by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation with units mL/min/1.73m^2, using categories (<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73m^2). A max-to-max shift table of urine albumin-to-creatinine ratio (UACR), using the categories UACR <30 mg/g, 30 mg/g ≤ UACR ≤300 mg/g, UACR >300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

The AE database will be searched using SMQs of acute renal failure and chronic kidney disease to identify events consistent with acute renal events. The incidence of the resulting TEAEs will be summarized by treatment and PT. Detailed searching criteria can be found in Appendix 2 (Section 7.2).

4.6.3.14. Dehydration

The AE database will be searched using SMQ of dehydration to identify events consistent with dehydration. Detailed searching criteria can be found in Appendix 2 (Section 7.2).

4.6.3.15. Metabolic Acidosis, Including Diabetic Ketoacidosis

The AE database will be searched using MedDRA PT to identify events consistent with metabolic acidosis, including diabetic ketoacidosis. Detailed searching criteria can be found in Appendix 2 (Section 7.2). The incidence of the resulting TEAEs will be summarized by treatment and PT.

4.6.3.16. Amputation/Peripheral Revascularization

The AE database will be searched using MedDRA PT to identify events consistent with amputation or peripheral revascularization. The incidence of the resulting TEAEs will be summarized by treatment and PT.

4.6.3.17. Major Depressive Disorder/Suicidal Ideation

The AE database will be searched using SMQs to identify events consistent with major depressive disorder or suicidal ideation. Detailed searching criteria can be found in Appendix 2 (Section 7.2). The incidence of the resulting TEAEs will be summarized by treatment and PT.

4.6.4. Vital Signs

Descriptive summaries by treatment and by nominal visit will be provided for baseline and postbaseline values as well as change from baseline values. If 2 records are taken at the same visit, they will be averaged prior to being used for data summaries and analyses.

An MMRM using REML model will be used to analyze the changes from baseline in vital signs at all scheduled postbaseline visits. The model will include baseline HbA1c (≤8.0%, >8.0%), treatment group, SGLT-2i use (Yes or No), visit, and treatment-by-visit interaction as fixed effects, and baseline value of the dependent variable as a covariate.

Counts and percentages of patients with treatment-emergent abnormal sitting systolic blood pressure (BP), sitting diastolic BP, and pulse will be presented by treatment. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table GPIM.4.3.

Table GPIM.4.3. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurements

Parameter	Low	High
Systolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤90 and decrease from baseline ≥20	≥140 and increase from baseline ≥20
Diastolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤50 and decrease from baseline ≥10	≥90 and increase from baseline ≥10
Pulse (bpm) (Supine or sitting)	<50 and decrease from baseline ≥15	>100 and increase from baseline ≥15

Abbreviations: BP = blood pressure; bpm = beats per minute.

4.6.5. Electrocardiogram

Summary statistics by treatment and by nominal visit will be provided for electrocardiogram (ECG) parameters (heart rate, PR, QRS, QT, and QT corrected using Fridericia's correction factor [QTcF]). When the QRS is prolonged (for example, a complete bundle branch block), QT and QTc should not be used to assess ventricular repolarization. Thus, for a particular ECG, the following will be set to missing (for analysis purposes) when QRS is ≥120 msec: QT and QTcF.

The criteria for identifying patients with treatment-emergent quantitative ECG abnormalities is outlined in Table GPIM.4.4.

In addition, the percentage of patients with QT greater than 500 msec will be summarized, and the percentage of patients with QTcF greater than 500 msec will also be summarized (refer to PSAP).

The percentage of patients who experienced a treatment-emergent increase from baseline in QTcF interval of greater than 30 msec, 60 msec, or 75 msec at any time will be summarized. The maximum value during the study follow-up will be analyzed. Planned and unplanned measurements will be included.

Table GPIM.4.4. Selected Categorical Limits for ECG Data

	Low		High	
Parameter	**Males**	**Females**	**Males**	**Females**
Heart Rate (bpm)	<50 and decrease ≥15	<50 and decrease ≥15	>100 and increase ≥15	>100 and increase ≥15
PR Interval (msec)	<120	<120	≥220	≥220
QRS Interval (msec)	<60	<60	≥120	≥120
QTcF (msec)	<330	<340	>450	>470

Abbreviations: bpm = beats per minute; ECG = electrocardiogram; QTCF = Fridericia's corrected QT interval.

4.6.6. Clinical Laboratory Evaluation

All laboratory data will be reported in the International System of Units and Conventional Units. Values that are outside of reference ranges will be flagged as high (H) or low (L) in the listings. Descriptive summaries by treatment and by nominal visit will be provided for the baseline and postbaseline values as well as the change from baseline values for selected measurements.

Observed and change from baseline values for selected measurements for each visit will be displayed graphically for patients who have both a baseline and a postbaseline planned measurement. Unplanned measurements will be excluded from graphs.

Shift tables will be produced for selected measurements. A shift table will include unplanned measurements. The shift table will include the number and percentage of patients within each baseline category (low, normal, high, or missing) versus each postbaseline category (low, normal, high, or missing) by treatment. The proportion of patients shifted will be compared between treatments.

A listing of abnormal findings will be created for laboratory analyte measurements. The listing will include patient ID, treatment group, laboratory collection date, study day, analyte name, and analyte finding.

4.6.7. Device Product Complaints

A listing or a summary of all device product complaints, inclusive of device product complaints that lead to an AE or that could have led to an SAE had intervention not been taken will be provided. Additional summaries will be provided as deemed appropriate.

4.7. Other Analyses

4.7.1. Health Outcomes

The patient-reported outcome questionnaires will be completed by the patients at baseline and at 40 weeks (or early discontinuation visit prior to 40 weeks). Using ANCOVA analyses on EAS, p-values and 2-sided 95% CIs of health outcome measures described below will be derived and summarized for the 5 mg, 10 mg, and 15 mg doses of tirzepatide compared to placebo. No multiplicity adjustment will be made in the evaluation of health outcome measures. Item level missingness is dealt with as per the instrument developer's instruction.

4.7.1.1. Ability to Perform Physical Activities of Daily Living

A descriptive frequency table of individual items in APPADL questionnaire will be presented at baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation). The changes from baseline to Week 40, with and without LOCF, of the of the raw and transformed total APPADL scores will be analyzed using an ANCOVA model with model terms of treatment, baseline HbA1c (≤8.0%, >8.0% [≤64, >64 mmol/mol]) and SGLT-2i use (Yes or No) as fixed effects, and corresponding baseline APPADL score as a covariate.

4.7.1.2. Impact of Weight on Self-Perceptions Questionnaire

A descriptive frequency table of individual items in IW-SP questionnaire will be presented at baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation). The changes from baseline to Week 40, with and without LOCF, of the raw and transformed total IW-SP scores will be analyzed using an ANCOVA model with model terms of treatment, , baseline HbA1c (≤8.0%, >8.0% [≤64, >64 mmol/mol]) and SGLT-2i use (Yes or No) as fixed effects, and corresponding baseline IW-SP score as a covariate.

4.7.1.3. Diabetes Treatment Satisfaction Questionnaire

Descriptive summaries will be provided at baseline (DTSQs only),observed endpoint, and endpoint including last observation carried forward excluding baseline observations (DTSQc only) for the perceived hyperglycemia item, perceived hypoglycemia item, and 6-item overall satisfaction score. The DTSQc score at Week 40, with and without LOCF, will be analyzed using an ANCOVA model with model terms of treatment, baseline HbA1c (≤8.0%, >8.0% [≤64, >64 mmol/mol]), and SGLT-2i use (Yes or No) as fixed effects, and baseline DTSQs score as a covariate. The analyses will be conducted for the perceived hyperglycemia item, perceived hypoglycemia item, and 6-item overall satisfaction score

4.7.1.4. EQ-5D-5L

A descriptive frequency table of individual items in EQ-5D-5L questionnaire will present baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation) values as separate summaries. The changes from baseline to Week 40, with and without LOCF of the index and visual analog scale (VAS) scores will be analyzed using an ANCOVA model with model terms of treatment, baseline HbA1c (≤8.0%, >8.0% [≤64, >64 mmol/mol]) and SGLT-2i use (Yes or No) as fixed effects, and corresponding baseline EQ-5D-5L score as a covariate. The EQ-5D-5L health states, defined by the EQ-5D-5L descriptive system, may be converted into a single summary index by applying a formula (UK or China) that essentially attaches values (also called weights) to each of the levels in each dimension (EuroQol Group 2019).

4.7.2. Subgroup Analyses

Subgroup analyses of the primary endpoint, change from baseline in HbA1c, at week 40, will be made based on the efficacy estimand using the EAS to assess consistency of the intervention effect across the following subgroups:

- Age group: <65 years vs ≥ 65 years
- Sex: female vs male
- Duration of diabetes: (≤5,>5 to ≤10, >10 years)
- Duration of diabetes (<median, ≥median)
- Baseline BMI (<24, ≥24 to <28, ≥28/kg/m^2),
- Concomitant OAM (no OAM, only metformin, metformin+SGLT-2i)
- Baseline HbA1c (≤8.0%, >8.0%)
- eGFR (<90, ≥90 mL/min/1.73m^2)

4.7.2.1. Subgroup Analysis of Weight Change at Week 40

Subgroup analyses by the following baseline characteristics will be provided for weight change from baseline at Week 40 based on the efficacy estimand using EAS:

- Age group: <65 years vs ≥ 65 years
- Sex: female vs male
- Duration of diabetes (<median, ≥ median)
- Duration of diabetes: (≤5, >5 to ≤10, >10 years)
- Baseline BMI (<24, ≥24 to <28, ≥28kg/m^2)
- Concomitant OAM (no OAM, only metformin, metformin+SGLT-2i)
- eGFR (<90, ≥90 mL/min/1.73m^2)
- Baseline HbA1c (≤8.0%, >8.0%)

4.8. Interim Analyses

No interim analyses are planned for this study.

4.9. Important Protocol Deviations

Important protocol deviations are specified in the Trial Issues Management Plan (TIMP). A listing and a summary of important protocol deviations by treatment will be provided.

4.10. Changes to Protocol-Planned Analyses

NA

5. Sample Size Determination

Participants will be randomized in a 1:1:1:1 ratio to tirzepatide 5 mg, 10 mg, 15 mg, or placebo.

The trial is powered to assess superiority of tirzepatide 10 mg and 15 mg, each tested in parallel, against placebo at a 2-sided significance level of 0.025, relative to the primary endpoint (mean change in HbA1c from baseline to 40 weeks), under the following assumptions: use of a 2-sample t-test utilizing HbA1c data collected before initiation of any rescue medication or premature treatment discontinuation with no more than 25% of participants initiating rescue medication or prematurely discontinuing treatment in each treatment group; 0.95% greater mean reduction in HbA1c from baseline for 10 and 15 mg tirzepatide compared with placebo; 1:1:1:1 randomization; and a common standard deviation (SD) of 1.3%. On the basis of these assumptions, a sample size of 256 participants is required to ensure at least 90% power to demonstrate that tirzepatide 10 mg and/or 15 mg are superior to placebo relative to the primary endpoint.

Assignment to treatment arms will be determined by a computer-generated random sequence using an interactive web response system (IWRS). Patients will be randomized in a 1:1:1:1 ratio to receive 5 mg tirzepatide, 10 mg tirzepatide, 15 mg tirzepatide, or placebo. The randomization will be stratified by baseline HbA1c concentration (≤8.0%, >8.0% [≤64, >64 mmol/mol]), and SGLT-2i use (Yes/No).

6. Supporting Documentation

6.1. Demographic and Baseline Characteristics

A listing of patient demographics will be provided. All demographic and baseline clinical characteristics will be summarized by study treatment for the patients in the modified intent-to-treat (mITT) population. Baseline demographic and clinical characteristics of special interest include but not limited to: age, gender, weight, BMI, HbA1c, fasting glucose, duration of T2DM, and estimated glomerular filtration rate (eGFR).

6.2. Treatment Compliance

Overall treatment compliance will be defined as taking at least 75% of the scheduled study drug or insulin glargine injections. Study drug compliance will be calculated by taking the number of injections (regardless of the actual dose administered) divided by the total number of injections expected to be administered × 100. Insulin glargine compliance will be calculated by taking the total number of injections expected minus the total number of injections missed (regardless of the actual dose administered) divided by the total number of injections expected to be administered × 100. Overall treatment compliance will be summarized descriptively by treatment using the mITT population.

6.3. Concomitant Therapy

The prespecified concomitant medications of interest will be summarized by treatment at randomization. Additionally, medications of interest initiated after randomization and change to medications of interest used at randomization will be summarized. The concomitant therapies will be mapped using the World Health Organization Drug dictionary in the clinical trial database.

The concomitant medications of interest include the following groups of medication:

- baseline antihyperglycemic medication
- baseline antihypertensive therapy
- baseline lipid lowering therapy
- changes to baseline medication in Study Period
    - antihypertensive therapy
    - lipid lowering therapy
- rescue therapy due to severe persistent hyperglycemia
- initiation of antihyperglycemic medication after study treatment discontinuation
- initiation of the following medications in Study Period:
    - antidiarrheal medication
    - antiemetic medication

6.4. Clinical Trial Registry

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

Analyses provided for the CTR requirements include the following:

- Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Non-Serious Adverse Events are summarized: by treatment group, by MedDRA preferred term.
    - An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE).
    - An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
        - the number of participants at risk of an event
        - the number of participants who experienced each event term
        - 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 (12-17 years), adults (18-64 years), 65-85 years, and 85 years and over.

7. Appendices

7.1. Appendix 1. SAS Code for Multiple Imputation

```
*retrieved dropout imputation*;
*ret_fl is a flag variable for subjects that are retrieved dropouts*;
proc mi data=&dataset_in. seed=&seed. out=&dataset_out. nimpute=&n_impute.;
by trtpn;
class ret_fl;
monotone reg(chg/details);
mnar model (chg/modelobs=(ret_fl='Y'));
var base chg;
run;

*placebo imputation*;
*replace 'placebo' with the assigned value for the placebo group*;
proc mi data=&dataset_in. seed=&seed. out=&dataset_out_2. nimpute=&n_impute.;
class trtpn;
monotone reg(chg/details);
mnar model (chg/modelobs=(trtpn='placebo'));
var base chg;
run;
```

7.2. Appendix 2. Searching Criteria for Adverse Events of Special Interest

The adverse events of special interest (AESI) analyses are detailed in Section 4.6.3. The search criteria for each AESI are stored in CLUWE:

\\statsclstr\lillyce\prd\diabetes\incretin\aesi\data\aesi_lookup.xlsx.

8. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. https://doi.org/10.1002/sim.5974

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

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

Dahl D, Onishi Y, Norwood P, et al. Effect of subcutaneous tirzepatide vs placebo added to titrated insulin glargine on glycemic control in patients with type 2 diabetes: The SURPASS-5 randomized clinical trial. *JAMA*. 2022;327(6):534–545. https://doi.org/10.1001/jama.2022.0078

Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons Inc.; 1987.

Signature Page for VV-CLIN-159774 v2.0

Approval	PPD n 08-Jan-2025 02:12:38 GMT+0000

Signature Page for VV-CLIN-159774 v2.0