I8F-MC-GPHD Statistical Analysis Plan Version 2

# A Randomized, Phase 3, Open-label Trial Comparing the Effect of the Addition of Tirzepatide Once Weekly Versus Insulin Lispro (U100) Three Times Daily in Participants With Type 2 Diabetes Inadequately Controlled on Insulin Glargine (U100) With or Without Metformin (SURPASS-6)

NCT04537923

Approval Date: 31-Oct-2022

# 1. Statistical Analysis Plan: I8F-MC-GPHD: A Randomized, Phase 3, Open-label Trial Comparing the Effect of the Addition of Tirzepatide Once Weekly versus Insulin Lispro (U100) Three Times Daily in Participants with Type 2 Diabetes Inadequately Controlled on Insulin Glargine (U100) with or without Metformin (SURPASS-6)


## Confidential Information

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

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


Tirzepatide (LY3298176) Type 2 Diabetes Mellitus

A phase 3, randomized, open-label trial comparing the effect of the addition of tirzepatide once weekly versus Insulin Lispro (U100) three times daily in participants with type 2 diabetes inadequately controlled on Insulin Glargine (U100) with or without Metformin.


Eli Lilly and Company
Indianapolis, Indiana USA 46285


Protocol I8F-MC-GPHD
Phase 3

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

# 2. Table of Contents

**Section** **Page**

## Table of Contents

**Table** **Page**

## Table of Contents

**Figure** **Page**

## Table of Contents

**Appendix** **Page**

# 3. Revision History

The first version of the statistical analysis plan (SAP) was approved prior to First Patient Visit (FPV).

The second version is approved before the final database lock. The following represent major changes made for the second version:

1. Updated definition of analysis set. For consistency with SURPASS 1-5 trials, exclude patients discontinuing study drug due to inadvertent enrollment from efficacy analyses.
2. Updated baseline definition as time is not collected in GPHD eCRF.
3. Missing data imputation: Modified the definition of "retrieved dropouts".
4. Updated analysis used for hemoglobin A1c (HbA1c) and weight loss threshold analyses.
5. Change in alpha recycling for type 1 error rate control.
6. Removed 'The patient recovered/ resolved with sequelae" from Severe hypoglycemia requiring hospitalization, documented medical help, or is life threatening as this criteria alone will not be sufficient to meet additional criteria.
7. Removed section on Anti-Drug Antibodies as it was not collected for all patients.
8. Added section to assess SARS-CoV-2 (COVID-19) impact.

# 4. Study Objectives

## 4.1. Primary Objectives

The primary objective of the study is to demonstrate noninferiority of tirzepatide pooled cohort of 5 mg, 10 mg, and 15 mg once weekly (QW) versus insulin lispro (U100) three times a day (TID), when added to insulin glargine (U100), with or without metformin, with respect to mean change in hemoglobin A1c (HbA1c) from baseline at 52 weeks.

## 4.2. Key Secondary Objectives Subject to Strong Type 1 Error Rate Control

Together with the primary objective, the following secondary objectives are subject to strong control of the type 1 error rate (see Section 6.12.3):

- To demonstrate superiority of tirzepatide pooled cohort of 5 mg, 10 mg, and 15 mg QW versus insulin lispro (U100) TID, when added to insulin glargine (U100), with or without metformin at 52 weeks for
    - mean change in HbA1c from baseline
    - mean change in body weight from baseline, and
    - proportion of participants with HbA1c target values of <7.0% (53 mmol/mol).
- To demonstrate noninferiority of tirzepatide 5 mg, 10 mg, and/or 15 mg QW versus insulin lispro (U100) TID, when added to insulin glargine (U100), with or without metformin at 52 weeks for
    - mean change in HbA1c from baseline.
- To demonstrate superiority of tirzepatide 5 mg, 10 mg, and/or 15 mg QW versus insulin lispro (U100) TID, when added to insulin glargine (U100), with or without metformin at 52 weeks for
    - mean change in HbA1c from baseline and
    - mean change in body weight from baseline.

## 4.3. Other Secondary and Exploratory Efficacy Objectives Not Subject to Type 1 Error Rate Control

The following secondary efficacy objectives are not subject to strong control of the type 1 error rate:

- To demonstrate that tirzepatide (pooled cohort of 5 mg, 10 mg, and 15 mg) QW is superior to insulin lispro (U100) TID, when added to insulin glargine (U100), with or without metformin at 52 weeks for:
    - proportion of participants achieving HbA1c target ≤6.5% (48 mmol/mol)

- mean change in fasting serum glucose (central laboratory) from baseline
- mean change in daily average 7-point self-monitored blood glucose profiles from baseline
- proportion of participants who achieved HbA1c target value of <7.0% (53 mmol/mol) without hypoglycemia (confirmed glucose <54 mg/dL [3.0 mmol/L] or report of severe hypoglycemia)
- proportion of participants who achieved weight loss of ≥5% from baseline. and
- mean change from randomization in Short Form-36 Health Survey (SF-36v2) acute form
  - Physical Component Summary score
  - Mental Component Summary score
  - Physical Functioning domain score
  - General Health domain score
  - Vitality domain score
  - Role-Physical domain score
  - Bodily Pain domain score
  - Social Functioning domain score
  - Role-Emotional domain score
  - Mental Health domain score.

- To demonstrate that tirzepatide 5 mg, 10 mg, and/or 15 mg QW is superior to insulin lispro (U100) TID, when added to insulin glargine (U100), with or without metformin at 52 weeks for:
  - proportion of participants achieving HbA1c target <7.0% (53 mmol/mol)
  - proportion of participants achieving HbA1c target ≤6.5% (48 mmol/mol)
  - mean change in fasting serum glucose (central laboratory) from baseline
  - mean change in daily average 7-point self-monitored blood glucose profiles from baseline
  - proportion of participants who achieved HbA1c target value of <7.0% (53 mmol/mol) without hypoglycemia (confirmed glucose <54 mg/dL [3.0 mmol/L] or report of severe hypoglycemia)
  - proportion of participants who achieved weight loss of ≥5% from baseline. and
  - mean change from randomization in SF-36v2 acute form

- Physical Component Summary score
- Mental Component Summary score
- Physical Functioning domain score
- General Health domain score
- Vitality domain score
- Role-Physical domain score
- Bodily Pain domain score
- Social Functioning domain score
- Role-Emotional domain score
- Mental Health domain score.

The following tertiary/exploratory objectives are also not subject to strong control of the type 1 error rate:

- To compare tirzepatide (pooled cohort of 5 mg, 10 mg, and 15 mg) to insulin lispro (U100) TID when added to insulin glargine (U100), with or without metformin at 52 weeks for:
  - mean change in lipids (total cholesterol, high-density lipoprotein [HDL-C], low-density lipoprotein [LDL-C], very low-density lipoprotein [VLDL-C], and triglycerides [TG]) from baseline
  - mean change in waist circumference from baseline
  - mean change in body mass index (BMI) from baseline, and
  - patient-reported outcomes
    - EQ-5D-5L
    - Ability to Perform Physical Activities of Daily Living (APPADL)
    - Impact of Weight on Self-Perception (IW-SP).

## 4.4. Safety Objectives

To evaluate the safety of tirzepatide (pooled cohort of 5 mg, 10 mg, and 15 mg) QW to insulin lispro TID, when added to insulin glargine (U100), with or without metformin at 52 weeks, and at the end of the safety follow-up period, with respect to the following outcomes:

- treatment-emergent adverse events (TEAEs)
- serious adverse events (SAEs)
- early discontinuation of study intervention (tirzepatide or insulin lispro [U100]) due to adverse events (AEs)

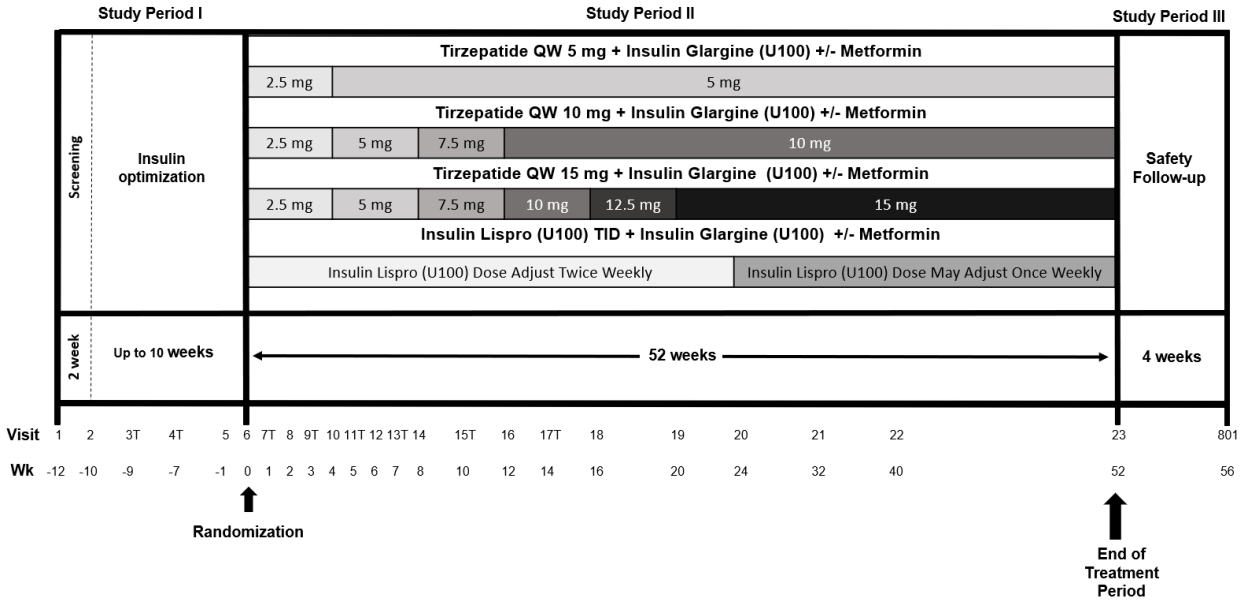
- adjudicated pancreatitis
- serum calcitonin
- incidence of treatment-emergent antidrug antibodies (ADAs) and systemic hypersensitivity reactions
- mean change in systolic and diastolic blood pressure and heart rate from baseline
- incidence of initiation of rescue therapy for severe, persistent hyperglycemia, and
- occurrence of hypoglycemic episodes.

# 5. Study Design

## 5.1. Summary of Study Design

Study I8F-MC-GPHD (GPHD) is a Phase 3b, multicenter, multinational, open-label, parallel-arm, randomized study, which will assess the safety and efficacy of the addition of tirzepatide (5 mg, 10 mg, or 15 mg QW) versus insulin lispro (U100) for change from baseline in HbA1c in participants with type 2 diabetes mellitus (T2DM) inadequately controlled on insulin glargine (U100) with or without metformin over a 52-week treatment.

Figure GPHD.5.1 illustrates the study design.




Abbreviations: QW = once weekly; TID = three times a day.

**Figure GPHD.5.1. Illustration of study design for clinical protocol I8F-MC-GPHD.**

### Study Period I (Screening and Insulin Optimization)

*Visit 1 (Screening)*

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 on their prestudy therapy doses between Visit 1 and Visit 2.

*Visit 2 to Visit 6 (Insulin Glargine Optimization)*

At Visit 2, the screening laboratory results from Visit 1 will be reviewed. A dilated fundoscopic exam will be performed between Visit 2 and Visit 3 as results are required to confirm eligibility. If calcitonin results are unavailable, then Visit 2 can proceed and calcitonin results can be reviewed later before randomization occurs.

At Visit 2, all participants should switch their pre-trial basal insulin therapy to Lilly provided insulin glargine (U100) and discontinue other oral glucose lowering agents except metformin. Between Visit 2 up to Visit 6, participants should continue taking the same dose and formulation of metformin ≥1500 mg/day, if used, unless any contraindication or clinical condition develops that requires adjustment of the dose. Participants should receive optimal therapy or clinical care for diabetes related comorbidities.

Figure GPHD.5.2 illustrates how participants who are eligible would proceed for insulin glargine optimization.

For all participants who are eligible until completion of Visit 2, the investigator will determine whether the participant belongs to **Group 1** or **Group 2** at Visit 2 and follow insulin optimization plan accordingly. **Group 1** consists of participants who are on pre-trial basal insulin regimen of insulin glargine (U100) daily at bedtime ± metformin (≥1500 mg/day). **Group 2** consists of participants who are on a pre-trial basal insulin regimen **other than** insulin glargine (U100) daily at bedtime ± metformin (≥1500 mg/day) or require discontinuation of sulfonylureas (SU) and/or dipeptidyl peptidase-4 inhibitors (DPP4i). Participants from **Group 1** will continue on their current dose of insulin glargine (U100) ± metformin (≥1500 mg/day) between Visit 2 and Visit 3, and measure fasting blood glucose (FBG) everyday between Visit 2 and Visit 3. Participants from **Group 2** will initiate optimization of insulin glargine (U100) for 10 weeks, starting at Visit 2 and up to Visit 6.

At Visit 3, participants from **Group 1** who do not require further insulin glargine (U100) optimization **(Group 1A)** (ie, if median of last 3 FBG is ≤125 mg/dL [≤6.9 mmol/L] at Visit 3) should have their randomization visit (Visit 6) within 2 weeks after Visit 3. These participants will skip Visits 4 and 5. Participants from **Group 1** who require further insulin glargine (U100) optimization **(Group 1B)** (ie, if median of last 3 FBG is >125 mg/dL [>6.9 mmol/L] at Visit 3) will initiate optimization of insulin glargine (U100) for 9 weeks, starting at Visit 3 and up to Visit 6.

For **Group 1B** and **Group 2** participants , the insulin glargine (U100) dose optimization and self-monitoring blood glucose (SMBG) plan should be targeted at a FBG of 100 to 125 mg/dL (5.6 to 6.9 mmol/L, inclusive) and determined by the investigator per clinical judgement. An additional blood sample will be collected for **Group 1B** and **Group 2** participants for HbA1c assessment at Visit 5. Only participants with Visit 5 HbA1c ≥7.5% (58 mmol/mol) to ≤11% (97 mmol/mol) will be eligible for randomization.

Those participants who discontinue the study prior to Visit 6 do not need to perform an early termination visit and safety follow-up visit.

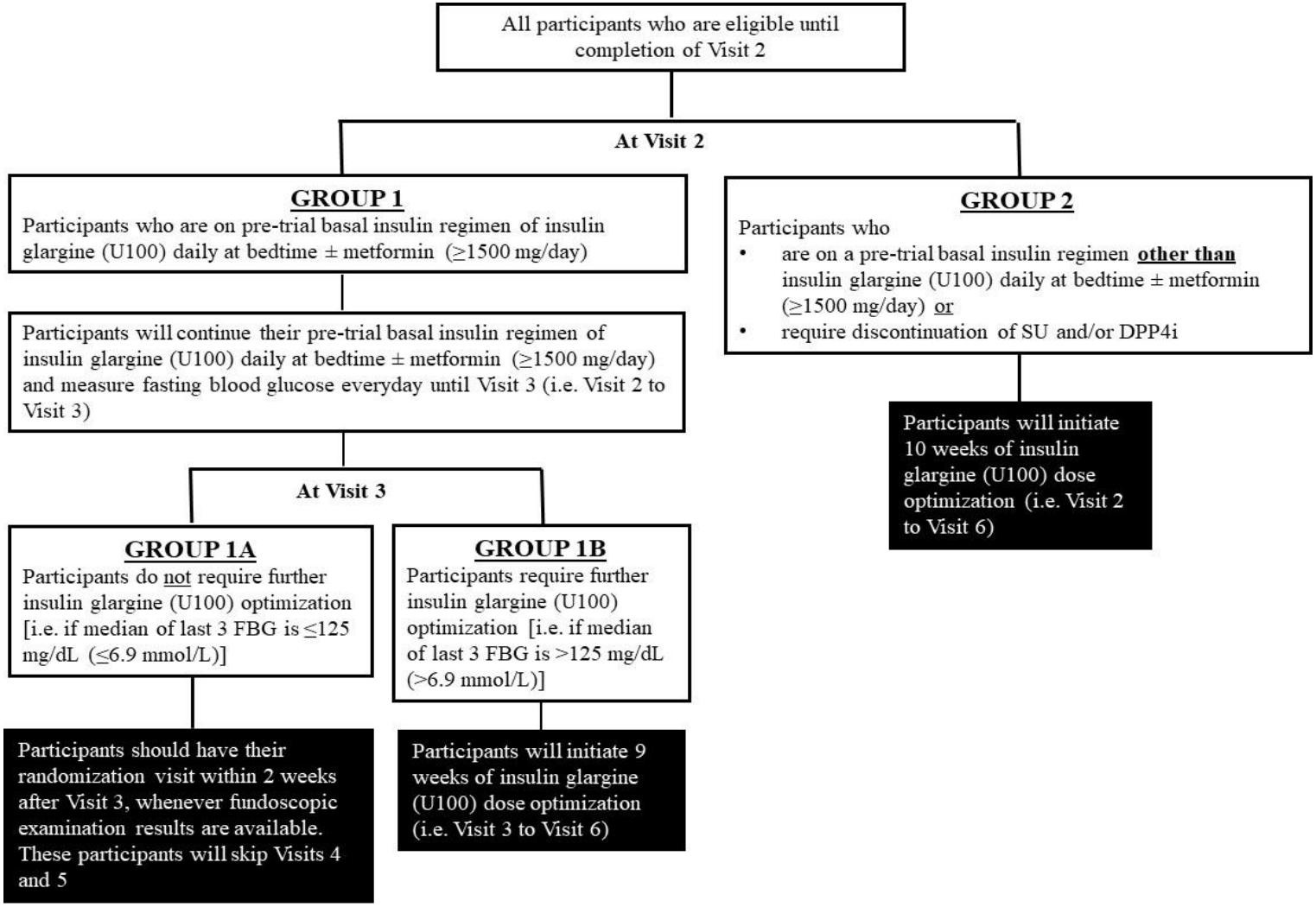



**Figure GPHD.5.2. Flowchart to ascertain need for insulin glargine (U100) optimization.**

**Study Period II (52-week treatment period)**

*Visit 6 (Randomization)*

At Visit 6, participants will undergo all required baseline study procedures (including the collection of all baseline laboratory measures, vital signs, electrocardiogram (ECG), and questionnaires) prior to randomization and prior to injecting their first dose of study intervention, tirzepatide or insulin lispro (U100).

All participants, randomized to either tirzepatide or insulin lispro (U100) TID, will **reduce their dose of insulin glargine (U100) by 30% at Visit 6** to reduce the potential risk of hypoglycemia and safely introduce tirzepatide or insulin lispro (U100) and follow titration instructions for insulin glargine (U100) treatment. Participants will follow dose-escalation schemes for tirzepatide or titration instructions for insulin lispro (U100) thereafter.

*End of Visit 6 to Visit 23 (Post-Randomization)*

Insulin Glargine (U100)

For participants in the tirzepatide arms, dose of insulin glargine (U100) will remain unchanged after the initial 30% dose reduction at Visit 6 until 4 weeks after randomization (ie, until Visit 10) except for safety reasons (ie, hypoglycemia/severe hyperglycemia). Participants will initiate insulin glargine (U100) titration after Visit 10 per protocol-defined Insulin Glargine (U100) Titration Algorithm.

For participants in the insulin lispro (U100) arm, participants will initiate titration of insulin glargine (U100) after Visit 6 per protocol-defined Insulin Glargine (U100) Titration Algorithm.

For insulin glargine (U100) titration, all participants should adjust insulin glargine (U100) dose weekly. Participants will be instructed to adjust insulin glargine (U100) doses to a targeted FBG of 100 to 125 mg/dL (5.6 to 6.9 mmol/L). If the participant achieves the target FBG per titration algorithm for 2 consecutive weeks and the dose of insulin glargine (U100) is less than 10 units, investigator may choose to temporarily interrupt insulin glargine (U100) injection.

Tirzepatide

For participants randomized to tirzepatide 5 mg QW arm, 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. For the 10 mg arm, 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 arm, 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.

Insulin Lispro

Participants randomized to insulin lispro (U100) arm will start administrating 4 U of insulin lispro, prior to the 3 most significant meals of the day. Participants should adjust insulin lispro

(U100) dose twice-weekly until Week 24 after randomization (ie, Visit 20). After Visit 20, the investigator may choose to reduce the frequency of insulin lispro (U100) dose adjustment to once weekly per clinical judgement.

Participants will be instructed to adjust insulin lispro (U100) doses to a target pre-meal and bedtime SMBG of 100 to 125 mg/dL (5.6 to 6.9 mmol/L) according to the protocol defined Insulin Lispro (U100) Titration Algorithm. If the participant achieves the glycemic goal per titration algorithm for 2 consecutive weeks and the dose of insulin lispro (U100) for corresponding meal is less than 4 units, investigator may choose to temporarily interrupt insulin lispro (U100) injections related to that mealtime.

**Study Period III (Safety Follow-up Period)**

*Visit 801*

All participants who complete the treatment period are required to complete Visit 801, a safety follow-up visit, approximately 4 weeks after their last treatment 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 intervention (tirzepatide or insulin lispro [U100]). Initiation of new antihyperglycemic therapy for the safety follow-up period will not be classified as "rescue therapy." Insulin glargine (U100) use will be permitted during the period.

## 5.2. Sample Size Determination

Although the primary objective of the trial is to demonstrate that the pooled cohort of tirzepatide (5 mg, 10 mg, and 15 mg) QW is noninferior to insulin lispro (U100) TID, relative to the primary endpoint (mean change in HbA1c from baseline to 52 weeks), sample size selection is guided by establishing noninferiority of each tirzepatide dose, tested against insulin lispro (U100), relative to the primary endpoint (mean change in HbA1c from baseline to 52 weeks). The power is assessed based on the following assumptions:

- the evaluation of noninferiority of each individual tirzepatide dose to insulin lispro (U100) will be conducted at one-sided 0.025 significance level, using a two-sample t-test
- a 0% difference in mean change in HbA1c from baseline to 52 weeks for each tirzepatide arm compared with insulin lispro (U100)
- a common standard deviation (SD) of 1.3% (accounting for increase in SD due to the inclusion of data on rescue medications and after premature treatment discontinuation and imputation of missing data)
- a noninferiority margin of 0.3%, and
- 1:1:1:3 randomization ratio (chosen to optimize power for establishing primary objective and to optimize safety comparison between pooled tirzepatide cohort versus insulin lispro).

On the basis of these assumptions, randomizing 1182 participants in a 1:1:1:3 ratio to tirzepatide 5 mg (197 participants), tirzepatide 10 mg (197 participants), tirzepatide 15 mg (197 participants) and insulin lispro (U100) (591 participants) provides 80% power to demonstrate noninferiority of each tirzepatide dose to insulin lispro (U100). Furthermore, this sample size will ensure 97% power to achieve the primary objective: establishing noninferiority of the pooled cohort of tirzepatide (5 mg, 10 mg, and 15 mg) compared to insulin lispro (U100), using a 2-sample t-test at 1-sided significance level of 0.025, provided a 0.0% difference in mean change in HbA1c from baseline to 52 weeks for the pooled cohort of tirzepatide doses compared with insulin lispro (U100), a common SD of 1.3% and noninferiority margin of 0.3%.

## 5.3. Method of Assignment to Treatment

Approximately 1182 participants who meet all criteria for enrollment will be randomized to one of the study treatment arms at Visit 6. Assignment to treatment arms will be determined by a computer-generated random sequence using an interactive web response system (IWRS). Participants will be randomized in a 1:1:1:3 ratio to receive 5 mg tirzepatide, 10 mg tirzepatide, 15 mg tirzepatide, or insulin lispro (U100). The randomization will be stratified by country, pre-randomization HbA1c concentration (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline metformin use (Yes/No).

# 6. A Priori Statistical Methods

## 6.1. Populations for Analyses

For purposes of analysis, Table GPHD.6.1 defines the following analysis populations/data sets.

**Table GPHD.6.1. Analysis Populations/Data Sets**

| Population/Data Set | Description |
|---|---|
| Screened population | All participants who sign informed consent |
| Enrolled population | All participants who are exposed to at least 1 dose of insulin glargine (U100) after Visit 2 |
| Randomized population | All participants who are randomly assigned a treatment arm |
| Safety Population | All randomly assigned participants who are exposed to at least 1 dose of study drug (tirzepatide or insulin lispro [U100]). |
| Modified intention-to-treat (mITT) population | All randomly assigned participants who are exposed to at least 1 dose of study drug (tirzepatide or insulin lispro [U100]), excluding inadvertently enrolled participants who have discontinued study drug for that reason. |
| Lead-in analysis set (LAS) | Data obtained during Study Period I, after first dose of insulin glargine (U100) after Visit 2 and prior to the first dose of tirzepatide or insulin lispro (U100) (or Visit 6 date if the dose date is missing). |
| Efficacy analysis set (EAS) | Data obtained during Study Period II from the mITT population, excluding data after initiating rescue antihyperglycemic medication or early discontinuation of study drug (tirzepatide: last dose date + 7 days; insulin lispro (U100): last dose date). |
| Full analysis set (FAS) | All available data obtained during Study Period II, regardless of adherence to study drug or initiation of rescue antihyperglycemic medication. |
| Safety analysis set (SS) | All available data obtained during Study Periods II and III regardless of adherence to study drug or initiation of new antihyperglycemic medication. |

## 6.2. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee. All statistical analyses will be conducted with SAS Version 9.4 or higher unless otherwise stated. Any change to the data analysis methods described in the protocol will require an amendment only if it changes a principle 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 clinical study report (CSR). Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (eg, few events to justify conducting an analysis). Listings of events will be provided in such situations. Additional analyses of the data may be conducted as deemed appropriate without further changes made to the protocol or SAP, even after the final database lock (DBL).

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 participants. In the event of a treatment error, participants will be analyzed according to the treatment they were randomized.

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 intervention and 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).

Unless specified otherwise, safety analyses for the treatment and the follow up period will be conducted relative to the "treatment-regimen" estimand using the safety analysis set (SS).

Unless specified otherwise, analysis for the lead-in period will be conducted using the lead-in analysis set (LAS).

Unless specified otherwise, the last available measurement up to Visit 6 (including unscheduled visits) collected prior to or on the first dose day will be used as the baseline for Study Period II. Unless specified otherwise, measurements prior to the first post-visit-2 insulin glargine (U100) dose will be used as the baseline for LAS.

The end of study participation for a participant will be the earliest of date of death, date of withdrawal from further participation in the study, or date of the safety follow-up visit (Visit 801). For participants considered to be lost to follow-up, end of study participation will be the date of lost to follow-up reported by the investigator. Participant data included in the database after the last date of study participation (date of death, date of early termination or date of safety follow-up) will be excluded from statistical analyses. A listing of such data may be provided.

Summary statistics for categorical measures (including categorized continuous measures) will include sample size, frequency, and percentages. Summary statistics for continuous measures will include sample size, mean, SD, median, minimum, and maximum. The summary statistics will be presented by nominal visit.

Statistical treatment comparisons will be performed between the pooled/individual tirzepatide doses and insulin lispro (U100). 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: individual tirzepatide doses (5 mg, 10 mg, 15 mg) and/or pooled cohort of tirzepatide doses and insulin lispro (U100).

## 6.3. Adjustments for Covariates

The study is stratified by country, pre-randomization HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline metformin use (Yes/No). Where necessary to be included as a fixed effect, countries with fewer than 10 randomized participants will be pooled into one category (pooled country). For HbA1c related analyses, country/pooled country and baseline metformin use (Yes/No) will be used as fixed effects and baseline HbA1c as a covariate. For

other efficacy analyses, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline metformin use (Yes/No) will be used as fixed effects and respective baseline value as a covariate.

## 6.4. Handling of Dropouts or Missing Data

For the primary and secondary efficacy endpoint analyses relative to treatment-regimen estimand and subject to type 1 error rate control, data for participants with missing values at the 52-week visit will be imputed based on the method described in Section 6.12.1.3. For analysis relative to efficacy estimand, missing values will be considered missing at random. Unless specified otherwise, imputation of missing data will be limited to primary and key secondary efficacy endpoint analyses. Any other secondary or exploratory efficacy parameter values or safety laboratory values that are missing will not be explicitly imputed.

## 6.5. Multicenter Studies

To investigate potential regional influences, country/pooled country will be used as a fixed effect in statistical analyses.

## 6.6. Multiple Comparisons/Multiplicity

The type 1 error rate control strategy for the primary and key secondary efficacy objectives is illustrated in Section 6.12.3. No multiplicity adjustments will be made for conducting separate analyses relative to the "efficacy" and "treatment-regimen" estimands, evaluating other secondary or exploratory efficacy objectives, or safety assessments.

## 6.7. Patient Disposition

A summary of study disposition during the lead-in period for all enrolled patients will be provided. Reasons for screen failures 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. Summaries of the final study disposition and the study drug disposition for all randomized patients will be provided by planned study treatment.

## 6.8. Patient Characteristics

A listing of patient demographics will be provided for the enrolled population. Demographic and baseline clinical characteristics will be summarized by randomization status (yes/no) for the patients in the enrolled population using the LAS data set. A similar summary will be generated by study treatment for the patients in the modified intent-to-treat (mITT) population. Baseline demographic and clinical characteristics of special interest include: age, gender, race, ethnicity, weight, BMI, country of enrollment, HbA1c, fasting serum glucose (only for the mITT population), estimated glomerular filtration rate (eGFR), insulin glargine daily dose, pre-trial oral glucose lowering medications (SUs, DPP4is and metformin) use and duration of T2DM.

## 6.9. Prior and Concomitant Therapy

Prespecified concomitant medications of interest ongoing at randomization will be summarized by treatment. Medications of interest initiated after randomization and changes to medications of interest used at randomization will be summarized. Additionally, medications of interest ongoing at the beginning of the insulin glargine (U100) optimization period (Visit 2) will be summarized.

Concomitant therapies will be mapped using the World Health Organization (WHO) DRUG dictionary in the clinical trial database and will be further classified using Anatomic-Therapeutic-Chemical (ATC) codes for reporting purposes. Concomitant medication will be summarized and compared between treatment groups.

Following table provides summary of descriptive analysis that will be generated for different concomitant medications.

| Type of medication | Population | Analyses Set | ATC level | Remarks |
|---|---|---|---|---|
| Anti-hyperglycemic medications[a] | Enrolled | LAS | 4 | Use at Visit 2 (pre-trial antihyperglycemic medications: types and doses of basal insulin, metformin, SUs and DPP4is) Summary of dose at Visit 2 will be provided at ATC 4 level |
| | mITT | FAS | 4 | Use at Visit 6 and Overall use anytime during Study Period 2 will be provided Summary of use as rescue therapy for severe persistent hyperglycemia |
| | mITT | SS | 4 | Overall use anytime during Study Period 2 or 3 will be provided |
| Anti-diarrheal medication | mITT | FAS | 4 | Initiated use during Study Period 2 will be provided |
| Anti-emetic medication | mITT | FAS | 4 | Initiated use during Study Period 2 will be provided |
| Other | mITT | FAS | 4 | -- |

Abbreviations: ATC = Anatomic-Therapeutic-Chemical; FAS = full analysis set; LAS = lead-in analysis set; mITT = modified intent-to-treat; SS = safety analysis set; SU = sulfonylureas.

[a] Insulin glargine (U100) use will be included in analysis

## 6.10. Treatment Exposure and Compliance

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 first dose date to last dose date plus 7 days for tirzepatide arms; from first dose date to last dose date for insulin lispro (U100) arm) will be provided by treatment.

## *6.10.1. Exposure and Compliance to Tirzepatide*

The number of patients and reasons for prematurely discontinuing study treatment prior to the 52-week visit will be provided by study treatment. Time-to-event analysis of premature study treatment discontinuation will be conducted.

A summary of patients who re-escalate Tirzepatide due to missing ≥3 doses will be provided.

Compliance to tirzepatide will be defined as taking at least 75% of the scheduled tirzepatide injections. Compliance will be calculated by taking the number of injections administered (regardless of the actual dose administered) divided by the total number of injections expected to be administered × 100. Treatment compliance will be summarized descriptively for tirzepatide arms.

## *6.10.2. Exposure and Compliance to Insulin*

### 6.10.2.1. Exposure to Insulin

Summary information on total daily dose of insulin glargine (U100) will be reported by visit and by treatment (tirzepatide 5 mg, 10 mg, 15 mg and insulin lispro) for the mITT population during Study Period I and II. A mixed-effect model for repeated measures (MMRM) will be used to analyze the ratio of post-baseline daily insulin glargine doses to baseline daily insulin glargine dose from randomization to Week 52 using the EAS data set. A log transformation will be performed on insulin glargine doses. The model terms will include baseline HbA1c category (≤8.5%, >8.5% [≤69, > 69 mmol/mol]), baseline metformin use (Yes/No), country/pooled country, treatment, visit, and treatment-by-visit interaction as fixed effects, and log-transformed baseline insulin glargine dose as a covariate. Summary information on total daily dose of insulin lispro (U100) will be reported by visit for the insulin lispro (U100) arm (patients who have received at least one dose of insulin lispro [U100]) during Study Period II.

If applicable, the ratio of total daily dose of insulin glargine (U100) to total daily dose of insulin lispro (U100) will be summarized by visit during Study Period II for the insulin lispro (U100) arm.

Summary information on total daily dose of insulin will be reported during Study Period II.

Unless specified otherwise, all insulin related data will be presented for both units: U and U/kg.

### 6.10.2.2. Compliance to Insulin Glargine (U100)

Compliance to insulin glargine (U100) will be defined as taking at least 75% of the scheduled insulin glargine (U100) injections. Compliance will be calculated by the formulas below:

Overall Insulin compliance (%) = $\frac{(total\ \#\ of\ injections\ expected - total\ \#\ of\ missed\ injections)}{total\ \#\ of\ injections\ expected} * 100.$

Treatment compliance will be summarized descriptively for mITT population during Study Period 2.

Information related to compliance to treat-to-target algorithm including number of assessments performed correctly, number of assessments that required a dose change, number of assessments for which the outcomes were correctly followed by the patient, and reasons for non-compliance will be summarized for Study Period 2.

#### 6.10.2.3. Compliance to Insulin Lispro (U100)

Compliance to insulin lispro (U100) will be defined as taking at least 75% of the scheduled insulin lispro (U100) injections. Compliance will be calculated by the formula below:

Overall Insulin compliance (%) = $\frac{(total\ \#\ of\ injections\ expected - total\ \#\ of\ missed\ injections)}{total\ \#\ of\ injections\ expected} * 100.$

Treatment compliance will be summarized descriptively for the insulin lispro (U100) arm during Study Period 2.

Information related to compliance to treat-to-target algorithm including number of assessments performed correctly, number of assessments that required a dose change, number of assessments for which the outcomes were correctly followed by the patient, and reasons for non-compliance will be summarized for Study Period 2.

## 6.11. Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan (TIMP). A listing and a summary of important protocol deviations will be provided for the randomized population by treatment. If applicable, a separate listing of important protocol deviations will be provided for the enrolled population.

## 6.12. Efficacy Analyses

Potentially for regulatory agencies, primary and key secondary efficacy assessments will be guided by the "treatment-regimen" estimand conducted using the FAS. Assessment of the primary and secondary efficacy objectives subject to type 1 error rate control (key secondary) will be conducted with multiple imputation of missing data (see Section 6.12.1.3) at 52 weeks. For publications and other purposes, the assessment of efficacy objectives will be guided by the "efficacy" estimand using the EAS without imputation of missing data.

Unless specified otherwise, the treatment term will include tirzepatide 5 mg, 10 mg, and 15 mg and insulin lispro (U100). The efficacy of the pooled tirzepatide arms compared with insulin lispro (U100) will be estimated via a linear contrast which averages the effect sizes of individual tirzepatide doses compared with insulin lispro (U100).

### *6.12.1. Primary Efficacy Analysis*

The primary efficacy measure will be change in HbA1c from baseline (postbaseline – baseline). Both HbA1c values as well as change from baseline in HbA1c will be summarized by treatment (pooled cohort of tirzepatide doses and insulin lispro) and nominal visit (week).

#### 6.12.1.1. 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 52-week visit with the aid of an 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 the primary measure and model terms of interest will include treatment (5 mg, 10 mg, and 15 mg tirzepatide and insulin lispro [U100]), visit, treatment-by-visit interaction, country/pooled country, and baseline metformin use (Yes/No) as fixed effects, and baseline HbA1c as a covariate. A linear contrast, averaging estimates from the individual tirzepatide doses at 52 week visit, will be used to estimate the treatment effect of the combined tirzepatide arms compared with insulin lispro (U100). 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 order:

- Heterogeneous Toeplitz
- Heterogeneous First Order Autoregressive
- Heterogeneous Compound Symmetry
- Toeplitz
- First Order Autoregressive, and
- 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 (ie, the pooled cohort of tirzepatide doses and insulin lispro [U100]).

With the aid of the MMRM analysis, 2-sided 95% confidence interval (CI) for mean change in HbA1c from baseline to the 52-week visit will be derived and summarized for the combined tirzepatide doses compared to insulin lispro (tirzepatide – insulin lispro). If the upper limit of the CI is ≤0.3%, then the pooled cohort of tirzepatide doses will be declared noninferior to insulin lispro (U100) relative to change in HbA1c from baseline.

#### 6.12.1.2. 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 52-week visit with the aid of an analysis of covariance (ANCOVA) model. The response variable will be the primary measure and model terms will include treatment (5 mg, 10 mg, and 15 mg tirzepatide and insulin lispro), country/pooled country, and baseline metformin use (Yes/No) as fixed effects and baseline

HbA1c as a covariate. A linear contrast, averaging estimates from the individual doses, will be used to estimate the treatment effect of the combined tirzepatide arms compared with insulin lispro (U100). The ANCOVA analysis will be conducted using multiple imputation of missing primary measures (see Section 6.12.1.3 for details) and statistical inference over multiple imputation of missing data guided by Rubin (1987).

With the aid of the ANCOVA analysis, 2-sided 95% CI for mean change in HbA1c from baseline to the 52-week visit will be derived and summarized for the combined tirzepatide doses compared to insulin lispro (tirzepatide – insulin lispro). If the upper limit of the CI is ≤0.3%, then the pooled cohort of tirzepatide doses will be declared noninferior to insulin lispro (U100) relative to change in HbA1c from baseline.

#### 6.12.1.3. Methods for Multiple Imputations

For efficacy analyses relative to the "treatment-regimen" estimand, missing HbA1c data at the 52-week visit will be imputed based on "retrieved dropouts," defined as patients who had their HbA1c value measured at the 52-week visit in the same treatment arm who prematurely discontinued study drug. In cases where there are not enough retrieved dropouts to provide a reliable imputation model (ie, model implemented by the SAS program does not converge), an alternative multiple imputation method with reference to the baseline data ("return to baseline" approach) will be used as the primary analysis relative to the treatment-regimen estimand. If value of 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. Analyses will be conducted with multiple imputations, and statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987).

Imputation of missing data using retrieved dropouts may bias towards the null which favors the noninferiority outcome. A sensitivity analysis using imputation under the non-inferiority null hypothesis will also be conducted to address this limitation (Koch 2018). In this analysis, missing data will first be imputed using retrieved dropouts from each arm. Since larger negative values represent better response for HbA1c reduction, the noninferiority margin of 0.3% will be added to the imputed values in the tirzepatide arms. This approach helps to avoid bias in the direction of making tirzepatide and insulin lispro arms more similar, therefore it reduces the chance of incorrectly demonstrating noninferiority.

### *6.12.2. Secondary Efficacy Analyses Subject to Type 1 Error Rate Control*

#### 6.12.2.1. Mean Change in HbA1c from Baseline at the 52 Week Visit

Noninferiority of individual doses of tirzepatide 5 mg, 10 mg, and 15 mg to insulin lispro (U100) will be conducted in the same manner as the primary analyses in Section 6.12.1, with individual doses of tirzepatide used in the treatment term.

Assessment of superiority of tirzepatide doses compared to insulin lispro (U100) will be conducted using the same statistical models as those used for evaluating the primary objective in Section 6.12.1. The treatment term will include tirzepatide 5 mg, 10 mg, 15 mg and insulin lispro

(U100). A linear contrast, averaging estimates from individual doses, will be used to assess the superiority of combined tirzepatide doses compared to insulin lispro (U100). Decisions will be guided by the 2-sided p-values for mean comparisons between tirzepatide doses and insulin lispro (U100) (see details in Section 6.12.3).

#### 6.12.2.2. Mean Change in Body Weight from Baseline at the 52 Week Visit

The analysis for change in body weight from baseline (postbaseline – baseline) will be conducted in a manner similar to the primary analyses in Section 6.12.1. Baseline HbA1c category (≤8.5%, >8.5% [≤69, > 69 mmol/mol]) will be used as a fixed factor in place of baseline HbA1c as a covariate and baseline of body weight will be used as an additional covariate in the statistical model. The treatment term will include tirzepatide 5 mg, 10 mg and 15 mg and insulin lispro (U100). Least squares mean estimates of mean change in body weight from baseline will be summarized by nominal visit and by study treatment. A linear contrast, averaging estimates from individual doses, will be used to estimate the mean change in body weight from the baseline for the pooled cohort of tirzepatide doses. For the multiple imputation of missing values, if value of the imputed weight change from baseline is <-50kg or >50kg, that value will be set to -50kg or 50kg, respectively, to avoid unrealistic values.

#### 6.12.2.3. Analyses of Binary Outcomes

The proportion of patients achieving HbA1c values <7.0% at the 52-week visit relative to the "efficacy" estimand will be conducted utilizing data from the EAS from baseline through the 52-week visit with the aid of a logistic regression. Missing HbA1c data were imputed using all the available data in the same treatment using fully conditional specification regression method. The missing data is imputed as a continuous outcome. Statistical inference over multiple imputations will be guided by Rubin (1987)

The analysis relative to the "treatment-regimen" estimand will be conducted utilizing HbA1c data in the FAS at baseline and at the 52-week visit with the aid of a logistic regression using multiple imputation of missing HbA1c data at the 52-week visit (see Section 6.12.1.3 for details) using retrieved dropouts.

After the different missing imputation methods for the "efficacy estimand" and the "treatment-regimen" estimand, the imputed values will then be dichotomized into binary outcomes based on cutoff values. The logistic regression with include model terms of treatment, country/pooled country, and baseline metformin use (Yes/No) as fixed effects, and baseline HbA1c as a covariate. The treatment term will include tirzepatide 5 mg, 10 mg, and 15 mg and insulin lispro (U100). A linear contrast, averaging estimates from individual tirzepatide doses, will be used to estimate the proportion of patients achieving HbA1c <7.0% for the combined tirzepatide doses. Statistical inference over multiple imputations will be guided by Rubin (1987).

### *6.12.3. Type 1 Error Rate Control Strategy for Primary and Key Secondary Efficacy Analyses*

Since they are intended for different purposes, no type 1 error rate adjustments will be made for conducting analyses relative to "efficacy" and "treatment-regimen" estimands. For analyses

within each estimand, the type 1 error control strategy for evaluation of the primary and key secondary objectives is illustrated in Figure GPHD.6.1.

The primary and key secondary objective hypotheses are as follows:

- $H_{p,1}$: The pooled cohort of tirzepatide (5 mg, 10 mg, and 15 mg) QW is noninferior to insulin lispro (U100) TID, relative to the mean change in HbA1c from baseline to 52 weeks.

- $H_{p,2}$: The pooled cohort of tirzepatide (5 mg, 10 mg, and 15 mg) QW is superior to insulin lispro (U100) TID, relative to mean change in HbA1c from baseline to 52 weeks.

- $H_{p,3}$: The pooled cohort of tirzepatide (5 mg, 10 mg, and 15 mg) QW is superior to insulin lispro (U100) TID, relative to mean change in body weight from baseline to 52 weeks.

- $H_{p,4}$: The pooled cohort of tirzepatide (5 mg, 10 mg, and 15 mg) QW is superior to insulin lispro (U100) TID, relative to proportion of participants achieving target value of HbA1c <7% at 52 weeks.

- $H_{5,1}$, $H_{10,1}$, $H_{15,1}$: Each of tirzepatide arms (5 mg, 10 mg, and 15 mg) QW is noninferior to insulin lispro (U100) TID, relative to mean change in HbA1c from baseline to 52 weeks.

- $H_{5,2}$, $H_{10,2}$, $H_{15,2}$: Each of tirzepatide arms (5 mg, 10 mg, and 15 mg) QW is superior to insulin lispro (U100) TID, relative to mean change in HbA1c from baseline to 52 weeks.

- $H_{5,3}$, $H_{10,3}$, $H_{15,3}$: Each of tirzepatide arms (5 mg, 10 mg, and 15 mg) QW is superior to insulin lispro (U100) TID, relative to mean change in body weight from baseline to 52 weeks.

A graphical testing scheme (Bretz 2009) presented in Figure GPHD.6.1 will be used to strongly control for type 1 error. $H_{p,1}$ will be initially tested at a two-sided 0.05 (one-sided 0.025) significance level:

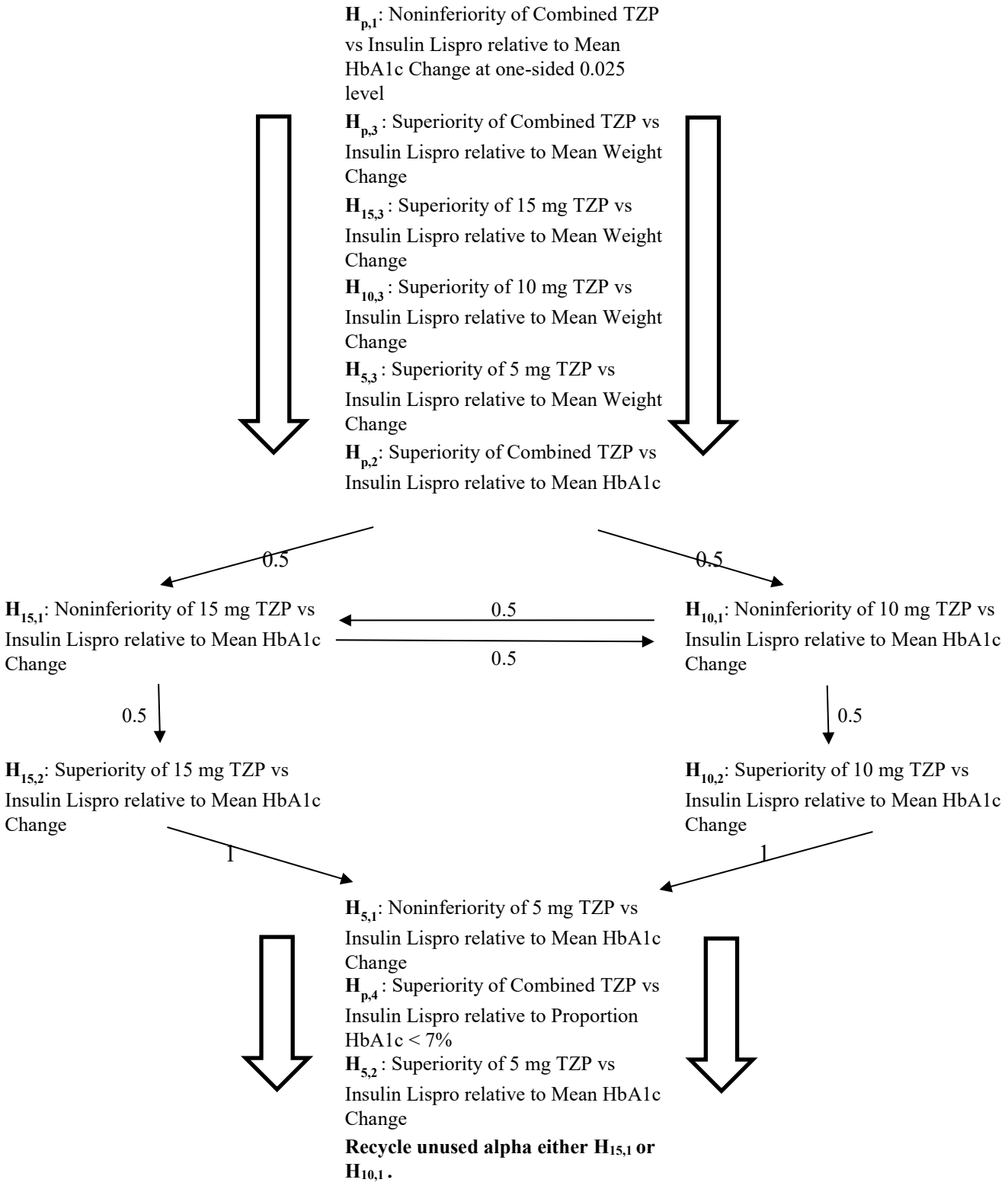



**Figure GPHD.6.1. Type 1 error control strategy for primary and key secondary efficacy endpoints.**

### *6.12.4. Other Secondary and Exploratory Efficacy Analyses*

Other secondary and exploratory efficacy measures will use the efficacy estimand and will be summarized by treatment and nominal visit. 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 GPHD.6.2. Secondary and 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 |
|---|---|---|---|
| **Secondary Analyses** | | | |
| To evaluate superiority of tirzepatide (5 mg, 10 mg, and/or 15 mg) QW to insulin lispro (U100) at 52 weeks in: | Proportion of participants achieving an HbA1c target value of <7% (53 mmol/mol) | 6.12.2.3 | None |
| To evaluate superiority of tirzepatide (pooled cohort of 5 mg, 10 mg, and 15 mg) QW to insulin lispro (U100), And to evaluate superiority of tirzepatide 5 mg, 10 mg, and/or 15 mg QW to insulin lispro (U100) at 52 weeks in: | Proportion of participants achieving an HbA1c target value of ≤6.5% (48 mmol/mol) | 6.12.2.3 | None |
| | Mean change in fasting serum glucose (central laboratory) from baseline | 6.12.1.1 | Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) in place of baseline HbA1c. Use baseline FSG as a covariate. |
| | Change from baseline in 7-point self-monitored blood glucose (SMBG) profiles | 6.12.1.1 | Use baseline SMBG as a covariate. LSM estimates at 52-weeks will be summarized by treatment and 7-points. |

| Objective | Relative to the efficacy measure: | | Analysis Conducted in a manner similar to section: | Additional Information |
|---|---|---|---|---|
| | | | | In addition to the analyses on each of the 7-points, similar analyses will be done for the 2-hour morning, midday, and evening meal excursions, the mean of all meals 2-hour excursion, the mean of all 7-point measurements, the mean of all pre-meal measurements, and the mean of all 2-hour postprandial measurements. |
| | Proportion of participants who achieved HbA1c target value of <7.0% (53 mmol/mol) without hypoglycemia (confirmed glucose <54 mg/dL [3.0 mmol/L] or report of severe hypoglycemia) | | 6.12.2.3 | Severe hypoglycemia follows ADA level 3 definition. |
| | Proportion of patients who achieved weight loss of ≥5%, from baseline | | 6.12.2.3 | Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) as in place of baseline HbA1c. Use baseline weight as a covariate. |
| | Mean change from randomization in SF-36v2 acute form | Physical Component summary score | 6.12.1.2 | Use ANCOVA. Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) as in place of baseline HbA1c. Use corresponding baseline score as a covariate. |
| | | Mental Component summary score | | |
| | | Physical Functioning domain score | | |
| | | General Health domain score | | |
| | | Vitality domain score | | |
| | | Role-Physical domain score | | |
| | | Bodily Pain domain score | | |

| Objective | Relative to the efficacy measure: | | Analysis Conducted in a manner similar to section: | Additional Information |
|---|---|---|---|---|
| | | Social Functioning domain score | | |
| | | Role-Emotional domain score | | |
| | | Mental Health domain score | | |
| **Exploratory Objectives** | | | | |
| To evaluate superiority of tirzepatide (pooled cohort of 5 mg, 10 mg, and 15 mg) QW to insulin lispro (U100) at 52 weeks in: | Change from baseline in lipid parameters (Total-Cholesterol, HDL-C, LDL-C, VLDL-C, TG) | | 6.12.1.1 | Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) in place of baseline HbA1c. Use corresponding baseline lipid parameter as a covariate |
| | Change from baseline in waist circumference | | 6.12.1.1 | Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) in place of baseline HbA1c. Use baseline waist circumference as a covariate |
| | Change from baseline in BMI | | 6.12.1.1 | Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) in place of baseline HbA1c. Use baseline BMI as a covariate. |

| Objective | Relative to the efficacy measure: | Analysis Conducted in a manner similar to section: | Additional Information |
|---|---|---|---|
| | Change from baseline in patient reported outcomes: APPADL, IW-SP, EQ-5D-5L | 6.12.1.2 | Use ANCOVA. Use baseline HbA1c category (≤8.5%, >8.5% [≤69, >69 mmol/mol]) as a fixed effect. Use corresponding baseline patient outcome score as a covariate. |

Abbreviations: ADA = antidrug antibodies; ANCOVA = analysis of covariance; 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; LSM = least squares mean; QW = once-weekly; SF-36v2 = Short-Form-36 Health Survey Version 2; TG = triglycerides; VLDL-C = very low-density lipoprotein.

## 6.13. Safety Analyses

Unless specified otherwise, safety assessments will be based on the Safety analysis Set (see Table GPHD.6.1). All events that occur between the date of first dose of study drug (ie, tirzepatide or insulin lispro [U100]) to the date of the patient's safety follow-up visit or the patient's end of study participation will be included, regardless of the adherence to study drug or initiation of rescue therapy. For assessing the benefit- risk profile through 52 weeks, selected safety analyses will be conducted by utilizing safety data from first dose through the date of the 52-week visit. Some safety analyses may be conducted after excluding data after initiation of new antihyperglycemic therapy. Additional safety analyses may be conducted using the LAS data set.

Unless specified otherwise, statistical assessment of homogeneity of the distribution of categorical safety responses among treatment arms (pooled cohort of tirzepatide doses and insulin lispro) will be conducted using Fisher exact test.

Unless specified otherwise, differences among treatment mean changes from baseline in continuous safety parameters at all scheduled visits will be assessed via an MMRM using REML. The model will include treatment, visit, treatment-by-visit interaction, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline metformin use (Yes/No) as fixed effects, and baseline value of the safety parameter as a covariate. To model the covariance structure within patients, the unstructured covariance matrix will be used. If this model fails to converge, the covariance structures specified in Section 6.12.1.1 will be tested in order.

Unless specified otherwise, statistical summaries and analyses of safety assessments will be displayed in the following treatment order: individual tirzepatide doses (5 mg, 10 mg, 15 mg), pooled cohort of tirzepatide doses and insulin lispro (U100).

For selected safety parameters, time-to-first-event analysis via the Cox proportional hazards model may be conducted. For patients without an event, the "time-to-event" will be time (in days) from first dose to end of study participation. For patients experiencing the event, "time-to-first-event" will be the time (in days) from first dose to first occurrence of the event.

Where necessary, the rate of events will be analyzed using a generalized linear mixed-effects model assuming the number of events follow a negative binomial distribution, and including treatment as a fixed effect. The logarithm of days during the active treatment period will be adjusted as an offset to account for possible unequal treatment duration of follow-up between patients.

### *6.13.1. Adverse Events*

An overview of the number and percentage of patients with TEAEs, SAEs, death due to an AE, and discontinuations from study due to an AE will be summarized for the enrolled population using the LAS data set.

An overview of the number and percentage of patients with TEAEs, SAEs, death due to an AE, and discontinuation from the study or from study treatment due to an AE, as well as the relationship to study treatment (tirzepatide, insulin lispro [U100]) will be summarized by treatment for the SS data set.

A TEAE is defined as an event that first occurred or worsened in severity after baseline. The baseline of a TEAE for LAS is defined as prior to the first insulin glargine (U100) after Visit 2. The baseline of a TEAE for the SS data set is defined as prior to first dose of randomized study drug (tirzepatide or insulin lispro [U100]). The Medical Dictionary for Regulatory Activities (MedDRA) Lowest Level Term (LLT) will be used in the treatment-emergent derivation. The maximum severity for each 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. For TEAEs after randomization, the summary will be provided by treatment.

The percentages of patients with TEAEs will be summarized using MedDRA PT nested within SOC. Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency within System Organ Class (SOC). For events that are sex-specific, the denominator and computation of percentages will include only patients from the given sex.

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

The percentages of patients with TEAEs by maximum severity will be summarized 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 be shown 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:

- death
- serious adverse event, or
- permanent discontinuation of study treatment due to an AE.

A listing of AEs occurring after the patient's last date of study participation will be provided using the SS data set. The listings will include patient identification including the treatment (for patients who are randomized), site number, and event information:AE group ID, event start date, MedDRA SOC, and PT, seriousness, severity, outcome, first dose date of insulin glargine (U100) and first and last dose date of the study drug (tirzepatide or insulin lispro [U100] for patients who are randomized).

#### 6.13.1.1. Deaths

A listing of deaths will be provided using the SS data set. The listing will include patient identification including the treatment, site number, date of death, age at the time of enrollment, sex, associated AE group ID, time from first dose of study drug (tirzepatide/insulin lispro [U100]) to death, time from last dose of study drug (tirzepatide/insulin lispro [U100]) to death (if patient had discontinued study drug), cause of death as reported by the investigator, and cause of death as adjudicated by the clinical endpoint committee (CEC). If applicable, a separate listing of deaths will be provided for the insulin glargine (U100) optimization period using the LAS data set. The listing will include patient identification information.

#### 6.13.1.2. Other Serious Adverse Events

The number and percentage of patients who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) during the treatment period and the follow-up period will be summarized by treatment using MedDRA PT nested within SOC. If applicable, the number and percentage of patients who experienced an SAE during the insulin glargine (U100) optimization period will be summarized using MedDRA PT for the LAS data set.

Listings of all SAEs will be provided. For the insulin glargine (U100) optimization period (Visit 2 to Visit 6), the listing will include insulin optimization group (Group 1A, Group 1B and Group 2), randomized (Yes/No), patient identification including the site number, date of the event, age at the time of enrollment, sex, AE group ID, MedDRA PT, severity, outcome, and time from first dose of insulin glargine (U100) to the event. For the treatment period including the follow-up period, the listing will include treatment, patient identification including the site

number, date of the event, age at the time of enrollment, sex, AE group ID, MedDRA SOC and PT, severity, outcome, relationship to study drug (tirzepatide or insulin lispro [U100]), time from first dose of study drug (tirzepatide or insulin lispro [U100]) to the event, and time from most recent dose to the event (if the patient discontinued study drug prior to the event).

#### 6.13.1.3. Discontinuation from the Study due to an 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 for the treatment and follow-up period. Events will be ordered by decreasing frequency within SOC.

If applicable, the number and percentage of patients who prematurely discontinue the study due to an AE will also be summarized using MedDRA PT for the insulin glargine (U100) optimization period.

#### 6.13.1.4. Discontinuation from Study Drug due to an 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 time-to-event analysis will be conducted by treatment on time to study drug (tirzepatide or insulin lispro [U100]) discontinuation.

#### 6.13.1.5. Treatment of Overdose

A listing of patients reporting over-dosing of tirzepatide will be provided.

### *6.13.2. Special Safety Topics*

#### 6.13.2.1. Hypoglycemic Events

Definitions of different categories of hypoglycemic events are included in Table GPHD.6.3.

**Table GPHD.6.3. Definitions of Hypoglycemic Event Categories**

| | Symptoms and/or Signs of Hypoglycemia | Blood Glucose Level |
|---|---|---|
| Glucose Alert Value (Level 1): | | |
| Documented symptomatic hypoglycemia | Yes | ≤70 mg/dL (3.9 mmol/L) |
| Documented asymptomatic hypoglycemia | No | ≤70 mg/dL (3.9 mmol/L) |
| Documented unspecified hypoglycemia | Unknown | ≤70 mg/dL (3.9 mmol/L) |
| Clinically Significant Hypoglycemia (Level 2): | | |
| Documented symptomatic hypoglycemia | Yes | <54 mg/dL (3.0 mmol/L) |
| Documented asymptomatic hypoglycemia | No | <54 mg/dL (3.0 mmol/L) |
| Documented unspecified hypoglycemia | Unknown | <54 mg/dL (3.0 mmol/L) |
| Severe Hypoglycemia (Level 3) | | |

**Severe hypoglycemia:** Defined as an episode with severe cognitive impairment requiring the assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions. Severe hypoglycemia will be reported as an SAE.

**Nocturnal hypoglycemia:** Defined as any hypoglycemic event that occurs between bedtime and waking.

**Severe hypoglycemia requiring hospitalization, documented medical help, or is life threatening:** Defined as any severe hypoglycemic event that has one of the following records in the electronic case report form (eCRF):

- The hypoglycemia event required emergency room visit or initial/prolonged hospitalization.
- The patient could not treat themselves and received assistance from medical personnel such as an emergency medical technician, nurse or physician.
- The hypoglycemia event was treated with glucagon or intravenous glucose.
- The hypoglycemia event caused disability or permanent damage to the patient.
- The patient experienced seizure or loss of consciousness.
- The hypoglycemia event caused death or was life threatening and fatal.

To avoid duplicate reporting, all consecutive hypoglycemic events in the same category, defined in Table GPHD.6.3, occurring within a 1-hour period may be considered to be a single hypoglycemic event. Severe hypoglycemia will be considered as adverse events of special interest (AESIs).

For selected categories of hypoglycemia (including severe hypoglycemia and level 2 hypoglycemia/severe hypoglycemia), statistical summaries and analyses will exclude hypoglycemic events occurring after initiation of a new antihyperglycemic therapy. Incidence as well as rate per year of exposure will be provided by treatment.

Statistical summaries and analyses for severe hypoglycemia requiring hospitalization, documented medical help or is life threatening will be conducted using the Safety Analysis Set.

The incidence of hypoglycemic event will be analyzed using logistic regression with treatment, baseline HbA1c category, country/pooled country, and baseline metformin use (Yes/No) as fixed effects. The baseline incidence of hypoglycemic event during the insulin glargine (U100) optimization period and insulin optimization group may be adjusted 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 country/pooled country, baseline metformin use (Yes/No), baseline HbA1c category, and treatment as fixed effects. The logarithm of days during the active treatment period will be adjusted as an offset to account for possible unequal treatment duration of follow-up between patients. The treatment term will include the combined tirzepatide arm and insulin lispro (U100). The baseline incidence/rate of hypoglycemic event during the

lead-in period and insulin optimization group may be adjusted as fixed effects. A nonparametric empirical model may be used to analyze the rate of hypoglycemic episodes if the event rate is low.

#### 6.13.2.2. Severe Persistent Hyperglycemia

A summary of initiation of rescue therapy in response to severe, persistent hyperglycemia will be provided by treatment. If there are a sufficient number of episodes, a time-to-first-event analysis for the initiation of rescue therapy will be conducted by treatment using a cox proportional regression model. For patients without an event the "time-to-event" will be time (in days) from first dose to end of study participation (study discontinuation or safety follow-up). A listing of patients initiating rescue therapy will be provided.

#### 6.13.2.3. Pancreatitis

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 1.Treatment-emergent adjudicated-confirmed pancreatitis will be considered as AESIs.

##### *6.13.2.3.1. Pancreatic Enzyme Assessment*

Observed pancreatic enzyme data (p-amylase and lipase) will be summarized by treatment and nominal visit. The number and proportion of patients with maximum postbaseline pancreatic enzyme values exceeding the following thresholds will be provided by baseline pancreatic enzyme value (≤ upper limit of normal [ULN], >ULN), and treatment: (>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 country/pooled country, baseline metformin use (Yes/No), baseline HbA1c category (≤8.5%, >8.5% [≤69, > 69 mmol/mol]), treatment, visit, and treatment-by-visit interaction as fixed effects.

#### 6.13.2.4. Thyroid Malignancies and C-Cell Hyperplasia

Treatment-emergent thyroid disease, C-cell hyperplasia, and neoplasms will be identified using pre-defined MedDRA high level terms (HLTs) of thyroid neoplasms, and PT of C-cell hyperplasia. Detailed searching criteria can be found in Appendix 1. A summary by treatment and PT/PT within HLT and a listing will be provided. Thyroid malignancies and C-cell hyperplasia will be considered as AESIs.

##### *6.13.2.4.1. Calcitonin*

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

#### 6.13.2.5. 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 1. A summary by treatment and PT within SMQ and a listing of TEAEs will be provided. Malignancy will be considered as an AESI.

#### 6.13.2.6. Major Adverse Cardiovascular Events

Major adverse cardiovascular events (MACE) reported by investigators are adjudicated by an independent CEC in a blinded fashion. The MACE of special interest are: deaths due to cardiovascular cause, myocardial infarction, hospitalization for unstable angina, hospitalization for heart failure, coronary interventions (such as coronary artery bypass graft [CABG] or percutaneous coronary intervention [PCI]); and cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack (TIA).

A listing of patients reporting MACE, either reported by investigator or identified by the CEC, will be provided. The listing will include treatment, patient identification including the site number, date of event, type of event as reported by the investigator, type of event as adjudicated by the CEC, time from first dose of study drug to the event, and time from the last dose to the event (if patient has discontinued study drug prior to the event). Only adjudication-confirmed MACE will be considered as AESIs.

#### 6.13.2.7. Supraventricular Arrhythmias and Cardiac Conduction Disorders

The AE database will be searched using pre-defined SMQ or MedDRA HLT to identify events consistent with supraventricular arrhythmias and cardiac conduction disorders. Detailed searching criteria can be found in Appendix 1. Incidence of the resulting TEAEs will be summarized by treatment and PT. Treatment-emergent severe/serious supraventricular arrhythmias and cardiac conduction disorders will be considered as AESIs.

#### 6.13.2.8. Hypersensitivity Events

Hypersensitivity reactions and related information reported via the "Hypersensitivity and Anaphylactic Reactions" eCRF 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.
- **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.

All potential hypersensitivity reactions will be reported by PT within SMQ. Preferred Term will be used for the summary within each category in decreasing order of incidence. Detailed searching criteria for hypersensitivity events can be found in Appendix 1. The serious/severe cases of hypersensitivity will be considered as AESIs.

#### 6.13.2.9. Injection Site Reactions

Injection site reactions, incidence and rates, 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, pruritus, and edema.

Additionally, potential injection site reactions will be reported by PT within MedDRA HLTs of injection site reactions and administration site reactions. The PT will be used for the summary within each category in decreasing order of incidence. Only the severe/serious injection site reactions will be considered as AESIs.

#### 6.13.2.10. Anti-Drug Antibodies

#### 6.13.2.11. Diabetic Retinopathy Complications

Results of the baseline dilated fundoscopic exam will be summarized by treatment. 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 change from baseline in dilated fundoscopic exam will be summarized by treatment and PT.

The cases with repeat 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 1 will be considered as AESI and summarized.

#### 6.13.2.12. Hepatic Safety

##### *6.13.2.12.1. Hepatobiliary Disorders*

The AE database will be searched using SMQs to identify events consistent with hepatobiliary disorders. Detailed searching criteria can be found in Appendix 1. A summary by treatment and PT within SMQ will be provided. Severe/serious hepatobiliary disorders will be considered as AESI.

##### *6.13.2.12.2. Acute Gallbladder Disease*

The AE database will be searched using pre-defined SMQs to identify events consistent with acute gallbladder diseases. Severe/serious acute gallbladder disease will be considered as AESIs. Detailed searching criteria for these AEs can be found in Appendix 1.

##### *6.13.2.12.3. Liver Enzymes*

Analyses for laboratory analyte measurements are described in Section 6.13.5. This section describes additional analyses of liver enzymes. The following will be provided by treatment group:

- A shift table of maximum to maximum alanine aminotransferase (ALT) measurement from baseline to postbaseline. Baseline categories include ≤ULN, >1 × ULN and missing.

Postbaseline categories include ≤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 to postbaseline. Baseline categories include ≤ULN, >1 × ULN and missing. Postbaseline categories include ≤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: ≤ULN, >1 to <2 × ULN, ≥2 × ULN.
- Shift tables of serum alkaline phosphatase (ALP) from baseline to postbaseline with the following categories: ≤ULN, >1 to <2 × ULN, ≥2 × ULN.

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

#### 6.13.2.13. Severe Gastrointestinal Adverse Events

Summaries of all events of treatment-emergent gastrointestinal (GI) AEs, including nausea, vomiting, and diarrhea will be provided by treatment and PT with decreasing frequency.

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

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

Preferred Terms with serious/severe cases in the gastrointestinal SOC will be considered as AESIs.

#### 6.13.2.14. Acute Renal Events

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 6.13.5.

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/1.73 $m^2$, using categories (<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73 $m^2$). 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).

To examine AEs for any suggestion of decrease in renal function, summary of acute renal failure searched by MedDRA SMQ will be presented. Detailed searching criteria can be found in Appendix 1. Severe/serious acute renal events will be considered as AESIs.

#### 6.13.2.15. 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 1. Severe/serious dehydration events will be considered as AESIs.

#### 6.13.2.16. Metabolic Acidosis, Including Diabetic Ketoacidosis

The AE database will be searched using MedDRA PT to identify events consistent with metabolic acidosis. Detailed searching criteria can be found in Appendix 1. The incidence of the resulting TEAEs will be summarized by treatment and PT. Severe/serious metabolic acidosis, including diabetic ketoacidosis will be captured as AESIs.

#### 6.13.2.17. 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. Amputation/peripheral revascularization will be considered as AESIs.

#### 6.13.2.18. 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 1. The incidence of the resulting TEAEs will be summarized by treatment and PT. Severe/serious major depressive disorder/suicidal ideation or behavior will be considered as AESIs.

### *6.13.3. Vital Signs*

Descriptive summaries at the screening visit will be summarized for the enrolled population. Descriptive summaries by treatment and by nominal visit will be provided for baseline (with respect to the treatment period) 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 will be used to fit the changes from baseline in vital signs at all scheduled postbaseline visits. The model will include treatment, visit, treatment-by-visit interaction, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline metformin use (Yes/No) 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 GPHD.6.4.

**Table GPHD.6.4. 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 |

Abbreviation: BP = blood pressure.

### *6.13.4. Electrocardiograms*

No statistical analysis will be conducted for ECG parameters in this study.

### *6.13.5. Clinical Laboratory Evaluation*

All laboratory data will be reported in the International System of Units. Selected laboratory measures will also be reported using 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 (with respect to the treatment period) and postbaseline values as well as the change from baseline values. If applicable, descriptive summaries for the screening visit will be provided for the enrolled population.

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.

For qualitative laboratory analytes, the number and percentage of patients with normal and abnormal values will be summarized by treatment.

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

## 6.14. Health Outcomes

The patient-reported outcome questionnaires will be completed by the patients at baseline and at 52 weeks (or early termination visit prior to 52 weeks). These include use of the mITT

population on the FAS, and use of a 2-sided alpha level of 0.05 and a 2-sided 95% CI for pairwise comparisons.

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.

### *6.14.1. Short-Form-36 Health Survey Version 2, Acute Form*

The SF-36v2 acute, 1-week recall version is a 36-item, generic, patient-completed measure designed to assess the following 8 domains:

- Physical Functioning
- Role-Physical
- Bodily Pain
- General Health
- Vitality
- Social Functioning
- Role-Emotional
- Mental Health

The Physical Functioning domain assesses limitations due to health "now" while the remaining domains assess functioning "in the past week." Each domain is scored individually and information from these 8 domains are further aggregated into 2 health-component summary scores: Physical Component Summary and Mental Component Summary. Items are answered on Likert scales of varying lengths (3-, 5-, or 6- point scales). Scoring of each domain and both summary scores are norm-based and presented in the form of T-scores, with a mean of 50 and SD of 10; higher scores indicate better levels of function and/or better health (Maruish 2011).

Per copyright owner, the QualityMetric Health Outcomes™ Scoring Software 4.5 will be used to derive the domains and component summary scores from the raw scores. Each domain and component summary scores will be summarized descriptively by treatment at each scheduled visit at which the SF-36v2 is administered. The changes from baseline to week 52 (last observation after baseline carried forward [LOCF]) in each domain and component summary scores will be analyzed using ANCOVA models with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) and baseline metformin use (Yes/No) as fixed effects, and the corresponding baseline score as a covariate.

### *6.14.2. EQ-5D-5L*

Each item will be summarized descriptively by treatment at each scheduled visit at which the EQ-5D-5L is administered. The changes from baseline to Week 52 (LOCF) in the index values and visual analog scale (VAS) scores will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69,

>69 mmol/mol]) and baseline metformin use (Yes/No) as fixed effects, and baseline EQ-5D-5L score as a covariate.

### *6.14.3. Impact of Weight on Self-Perceptions Questionnaire*

Descriptive summaries by treatment at each scheduled visit at which the IW-SP is administered will be presented for each item. Treatment comparisons of the raw and transformed overall IW-SP score changes from baseline to Week 52 (LOCF) will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) and baseline metformin use (Yes/No) as fixed effects, and baseline IW-SP score as a covariate.

### *6.14.4. Ability to Perform Physical Activities of Daily Living*

Descriptive summaries by treatment at each scheduled visit at which the APPADL is administered will be presented for each item. Treatment comparisons of the raw and transformed overall APPADL score changes from baseline to week 52 (LOCF) will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) and baseline metformin use (Yes/No) as fixed effects, and baseline APPADL score as a covariate.

## 6.15. Subgroup Analyses

Efficacy subgroup analyses will be guided by the efficacy estimand. Safety subgroup analyses will use mITT population with SS set.

Subgroup analyses may be done by country to support local regulatory registrations.

### *6.15.1. Subgroup Analysis of HbA1c Change at 52-weeks*

Subgroup analyses by the following baseline characteristics will be provided: age group (<65, ≥65 years), age group (<75, ≥75 years), race, gender, region of enrollment (US, Latin America, Europe including Russia and Turkey), duration of diabetes (<median, ≥median), HbA1c (≤8.5%, >8.5%), baseline metformin use (Yes/No), renal impairment (eGFR <60, ≥60 mL/min/1.73m$^2$), BMI group (<27, ≥27 kg/m$^2$), BMI group (<30, ≥30 to <35, ≥35 kg/m$^2$) and requirement of insulin optimization (Group 1A, combined Group 1B and 2) during lead-in period.

### *6.15.2. Subgroup Analysis of Weight Change at 52 Weeks*

Subgroup analyses by the following baseline characteristics will be provided: age group (<65, ≥65 years), age group (<75, ≥75 years), race, gender, region of enrollment (US, Latin America, Europe including Russia and Turkey), duration of diabetes (<median, ≥median), HbA1c (≤8.5%, >8.5%), baseline metformin use (Yes/No), renal impairment (eGFR <60, ≥60 mL/min/1.73m$^2$), BMI group (<27, ≥27 kg/m$^2$), BMI group (<30, ≥30 to <35, ≥35 kg/m$^2$), and requirement of insulin optimization (Group 1A, combined Group1B and 2) during lead-in period.

## 6.16. Interim Analyses and Data Monitoring Committee

No interim analyses are planned for this study.

## 6.17. COVID-19 Impact Assessment

This section lists the potential statistical analyses that may be performed to assess the impact of COVID-19 pandemic when appropriate.

### *6.17.1. Patients Impacted by COVID-19*

Listings of patients with protocol deviation or mitigation due to COVID-19/ related restrictions, patients with COVID-19 adverse events or death, and patient's dispositions with reasons related to COVID-19 will be provided. Listing will include patient identification, treatment, and description of impact.

### *6.17.2. Adverse Events*

A summary table for patients with adverse events related to COVID-19, including death due to COVID-19, serious COVID-19 adverse events, and COVID-19 adverse events, will be provided by study treatment.

### *6.17.3. Patient Disposition*

Patient disposition with reasons related to COVID-19 (such as COVID-19 adverse event, etc.) will be summarized for study and study treatment discontinuation by treatment group.

### *6.17.4. Study Visits*

A summary of patients with study visit impacted by COVID-19/ related restrictions will be provided by treatment group. In this table, number and proportion of patients missing study visit including primary endpoint visit will be summarized.

# 7. Unblinding Plan

Even through GPHD is an open-label study, the study team will remain as blinded as possible to the tirzepatide doses. The unblinding plan will be located in a separate document.

# 8. References

Bretz F, Maurer W, Brannath B, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604

Koch GG. Comments on 'Current issues in non-inferiority trials' by Thomas R. Fleming, Statistics in Medicine. DOI: 10.1002/sim.2855. *Stat Med*. 2008;27(3):333-342. https://doi.org/10.1002/sim.2923

Maruish ME, editor. *User's Manual for the SF-36v2 Health Survey*. 3rd ed. QualityMetric Incorporated; 2011.

Protocol I8F-MC-GPHD A Randomized, Phase 3, Open-label Trial Comparing the Effect of the Addition of Tirzepatide Once Weekly versus Insulin Lispro (U100) Three Times Daily in Participants with Type 2 Diabetes Inadequately Controlled on Insulin Glargine (U100) with or without Metformin (SURPASS-6).

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

# 9. Appendices

# Appendix 1. Searching Criteria for Adverse Events of Special Interest

The adverse events of special interest (AESI) analyses are detailed in Section 6.13.2. The search criteria for each AESI are stored in CLUWE:

CCI [redacted]

Signature Page for VV-CLIN-068485 v1.0

| | |
|---|---|
| Approval | PPD Management<br>31-Oct-2022 04:27:19 GMT+0000 |

Signature Page for VV-CLIN-068485 v1.0