

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

217806Orig1s000

STATISTICAL REVIEW(S)



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #: NDA 217806

Supplement #: Original

Drug Name: Tirzepatide

Indication(s): As an adjunct to a reduced-calorie diet and increased physical activity for chronic weight management in adults with an initial body mass index of

- 30 kg/m² or greater (obesity) or
- 27 kg/m² or greater (overweight) in the presence of at least one weight-related comorbid condition

Applicant: Eli Lilly and Company

Date(s): Receipt Date (Rolling submission)

- Sequence 1: 11/17/2022
- Sequence 2: 5/8/2023

Filing Meeting: 6/2/2023
Review Due: 10/6/2023
PDUFA: 11/8/2023

Review Priority: Priority

Biometrics Division: DBII

Statistical Reviewer: Kyunghee K. Song, PhD

Concurring Reviewers: Feng Li, PhD, Team Leader

Medical Division: Division of Diabetes, Lipid Disorders, and Obesity Products

Clinical Team: Raymond Soccio, MD, Reviewer
Laura Higginbotham, MD, Team Leader

Project Manager: Arati Kamath

TABLE OF CONTENTS

1	EXECUTIVE SUMMARY	5
2	INTRODUCTION	6
2.1	OVERVIEW	6
2.2	DATA SOURCES	7
3	STATISTICAL EVALUATION	7
3.1	DATA AND ANALYSIS QUALITY.....	7
3.2	EVALUATION OF EFFICACY.....	8
3.2.1	Study Design and Endpoints	8
3.2.2	Statistical Methodologies.....	13
3.2.3	Patient Disposition, Demographic and Baseline Characteristics	17
3.2.4	Results and Conclusions	25
3.3	EVALUATION OF SAFETY	34
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	35
4.1	SEX, RACE, AGE, AND GEOGRAPHIC REGION	36
4.2	OTHER SPECIAL/SUBGROUP POPULATIONS	40
5	SUMMARY AND CONCLUSIONS	44
5.1	STATISTICAL ISSUES	44
5.2	COLLECTIVE EVIDENCE	45
5.3	CONCLUSIONS AND RECOMMENDATIONS	45
5.4	LABELING RECOMMENDATIONS (AS APPLICABLE)	45

LIST OF TABLES

Table 1: %Change in body weight and 5% Body weight loss (from Baseline to Week 72)	6
Table 2: Patient Disposition in SURMOUNT-1	18
Table 3: Patient Disposition in SURMOUNT-2	19
Table 4: Summary of COVID-19 Related Events	20
Table 5: Baseline Demographics in SURMOUNT-1	21
Table 6: Baseline Demographics in SURMOUNT-2	22
Table 7: Baseline Characteristics in SURMOUNT-1	23
Table 8: Baseline Characteristics in SURMOUNT-2	24
Table 9: Summary of Observed/Missing Data	25
Table 10: Primary analysis: %Change in body weight and 5% body weight loss in SURMOUNT-1	26
Table 11: %Change in body weight and 5% body weight loss in SURMOUNT-2	27
Table 12: Key Secondary Endpoints in SURMOUNT-1 (at Week 72)	28
Table 13: Key Secondary Endpoint in SURMOUNT-1 (at Week 20)	28
Table 14: Key Secondary Endpoints in SURMOUNT-1 (at Week 72)	29
Table 15: Key Secondary Endpoints in SURMOUNT-1 (at Week 72)	30
Table 16: Key Secondary Endpoint in SURMOUNT-1 (at Week 72)	31
Table 17: Key Secondary Endpoints in SURMOUNT-2 (at Week 72)	32
Table 18: Key Secondary Endpoints in SURMOUNT-2 (at Week 72)	33
Table 19: Key Secondary Endpoints in SURMOUNT-2 (at Week 72)	34
Table 20: Overview of Adverse Events	35

LIST OF FIGURES

Figure 1: Study Schema	9
Figure 2: Study Schema	12
Figure 3: Type 1 error control strategy for primary end key secondary endpoints (SURMOUNT-1).....	16
Figure 4: Type 1 error control strategy for primary end key secondary endpoints (SURMOUNT-2).....	17
Figure 5: Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-1	37
Figure 6: Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-1	38
Figure 7: Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-2	39
Figure 8: Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-2	40
Figure 9: Other Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-1	41
Figure 10: Other Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-1	42
Figure 11: Other Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-2	43
Figure 12: Other Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-2	44

1 EXECUTIVE SUMMARY

Eli Lilly and Company (referred to as “the applicant” in this review) submitted an original NDA for tirzepatide as an adjunct to a reduced-calorie diet and increased physical activity for chronic weight management in adult patients with an initial body mass index (BMI) of 30 kg/m² or greater (obesity), or 27 kg/m² or greater (overweight) in the presence of at least one weight-related comorbid condition (e.g., hypertension, dyslipidemia, obstructive sleep apnea, cardiovascular disease, or type 2 diabetes).

Tirzepatide was previously approved under the tradename MOUNJARO™ as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus in 2022 (NDA 215866).

The submission consists of two phase 3 trials, I8F-MC-GPHK (SURMOUNT-1) and I8F-MC-GPHL (SURMOUNT-2).

SURMOUNT-1 was a phase 3, multicenter, randomized, placebo-controlled, double-blind, 72-week study of the safety and efficacy of 5, 10, and 15 mg tirzepatide once weekly (QW) compared with placebo for weight management when used in conjunction with a reduced-calorie diet and increased physical activity, in subjects with obesity or overweight with weight-related comorbidities (excluding Type 2 diabetes (T2D)). Comorbid condition included obstructive sleep apnea (OSA), hypertension, dyslipidemia or cardiovascular (CV) disease.

SURMOUNT-2 was a phase 3, multicenter, randomized, parallel-arm, placebo-controlled, double-blind, 72-week study that investigated the effects of treatment with tirzepatide 10 and 15 mg QW compared with placebo, in conjunction with a reduced-calorie diet and increased physical activity, on weight management in subjects with T2D who have obesity (BMI ≥30 kg/m²) or are overweight (BMI ≥27 kg/m²).

In both studies, the primary endpoints were the percent change from baseline to Week 72 in body weight and the proportion of subjects who achieved at least 5% body weight reduction from baseline to Week 72 for the tirzepatide 10 mg and 15 mg groups.

The primary efficacy results demonstrated the efficacy for body weight reduction at Week 72 for both the 10 mg and the 15 mg doses, and the results are shown in **Error! Reference source not found.** Missing values were handled using missing at random and retrieved-dropout multiple imputation approaches based on an Applicant defined hybrid estimand framework, for the primary analyses.

Table 1:%Change in body weight and 5% Body weight loss (from Baseline to Week 72)

	Tirzepatide 10 mg N=636	Tirzepatide 15 mg N=630	Placebo N=643
SURMOUNT-1			
%Change in body weight			
Baseline body weight (kg), mean (SD)	105.8 (23.3)	105.6 (22.9)	104.8 (21.4)
Estimate, LSMean (SE) ¹	-19.5 (0.5)	-20.9 (0.5)	-3.1 (0.6)
Difference from placebo, (95% CI)	-16.4 (-17.9, -14.8)	-17.8 (-19.3, -16.3)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)
5% Body weight loss			
Proportion estimate ²	89.6	91.4	34.0
Difference from placebo (95% CI)	55.6 (50.0, 61.2)	57.4 (51.9, 62.9)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)
SURMOUNT-2			
%Change in body weight			
Baseline body weight (kg), mean (SD)	100.9 (20.9)	99.6 (20.1)	101.7 (22.3)
Estimate, LSMean (SE) ¹	-12.8 (0.6)	-14.7 (0.5)	-3.2 (0.5)
Difference from placebo, (95% CI)	-9.6 (-11.1, -8.1)	-11.6 (-13.0, -10.1)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)
5% Body weight loss			
Proportion estimate ²	79.6	83.2	32.0
Difference from placebo (95% CI)	47.6 (40.2, 55.0)	51.2 (43.9, 58.4)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)

Abbreviations: N = number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline body weight (kg) as a covariate; ²Estimates based on a logistic regression using treatment and stratification group as factors and baseline body weight (kg) as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt]

There were no major statistical issues found during the review of this submission. Efficacy in comparison to placebo was further supported by the confirmatory secondary endpoints. Based on information from the clinical reviewer, it seems there were no major safety concerns identified during the review.

Collectively, the studies provided substantial evidence of effectiveness for the proposed indication.

2 INTRODUCTION

2.1 Overview

Tirzepatide is a 39-amino acid synthetic peptide, engineered from the native GIP peptide sequence, modified to bind to both glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide-1 (GLP-1) receptors. Tirzepatide has a mean half-life of approximately 5

days, which enables once-weekly dosing. The increased glucose lowering observed with tirzepatide, attributed to dual GIP and GLP-1 receptor agonist (RA), may reflect an added benefit of GIP receptor-mediated insulinotropic action and an increase of insulin sensitivity independent of body weight changes. The increased weight reduction observed with tirzepatide may be attributed to a combination of the GIP receptor and GLP-1 receptor-mediated mechanisms regulating satiety and food intake, and potentially energy expenditure.

The trials were designed to investigate that tirzepatide subcutaneous injection (s.c.) once-weekly as an adjunct to a reduced-calorie diet and increased physical activity, was effective and safe for treatment of overweight or obesity with and without type 2 diabetes (T2D). This program consisted of two phase 3 trials:

- Study Number: I8F-MC-GPHK, titled “Efficacy and Safety of Tirzepatide Once Weekly in Subjects Without Type 2 Diabetes Who Have Obesity or Are Overweight with Weight-Related Comorbidities: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-1)”
- Study Number: I8F-MC-GPHL, titled “Efficacy and Safety of Tirzepatide Once Weekly in Subjects with Type 2 Diabetes Who Have Obesity or Are Overweight: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-2)”

The clinical program was discussed with the Agency under IND 139721. The applicant addressed the statistical comments conveyed during the IND stage.

2.2 Data Sources

Materials for this statistical review, including the data and a clinical study report (CSR), were submitted electronically under the network path location:

<\\CDSESUB1\evsprod\NDA217806\0001> for SURMOUNT-1 and

<\\CDSESUB1\evsprod\NDA217806\0009> for SURMOUNT-2.

The information necessary for the statistical review was contained in Module 1 (cover letter and labeling) and Module 5 (clinical study report, study protocol, statistical analysis plan, datasets, and programs).

The applicant’s responses to the statistics information requests for a list of analysis reports and programs were submitted electronically and located under the network path:

<\\CDSESUB1\evsprod\NDA217806\0014> (for subgroup analysis) and

<\\CDSESUB1\evsprod\NDA217806\0018> (additional information related to the labeling).

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submitted efficacy data and analyses are generally acceptable in quality and documentation. The statistical reviewer was able to reproduce the results of primary and important secondary analyses and performed additional analyses as needed.

Blinding procedures were described in the study reports and acceptable. Until the end of the study, treatment assignments remained blinded for the sponsor, investigators, site staff, clinical monitors and study subjects.

For SURMOUNT-1, external data monitoring committee (DMC) reviewed unblinded safety data. Three DMC meetings were held and the committee recommended to continue study without modification. For SURMOUNT-2, there was no data monitoring committee.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

SURMOUNT-1

The trial was a multinational, multicenter, double-blind, parallel-group, placebo-controlled study that randomly assigned subjects to receive once-weekly (QW), injectable placebo or tirzepatide 5, 10, or 15 mg, in conjunction with a reduced-calorie diet and increased physical activity, for weight management in subjects without type 2 diabetes mellitus (T2DM) who have obesity ($BMI \geq 30 \text{ kg/m}^2$) or are overweight ($BMI \geq 27 \text{ kg/m}^2$).

The study was conducted at 118 sites in 9 countries including the United States.

A total of 2539 subjects were randomized 1:1:1:1 to tirzepatide 5 mg (N=630), 10 mg (N=636), 15 mg (N=630) or placebo (N=643) and the randomization was stratified by prediabetes status, country and sex.

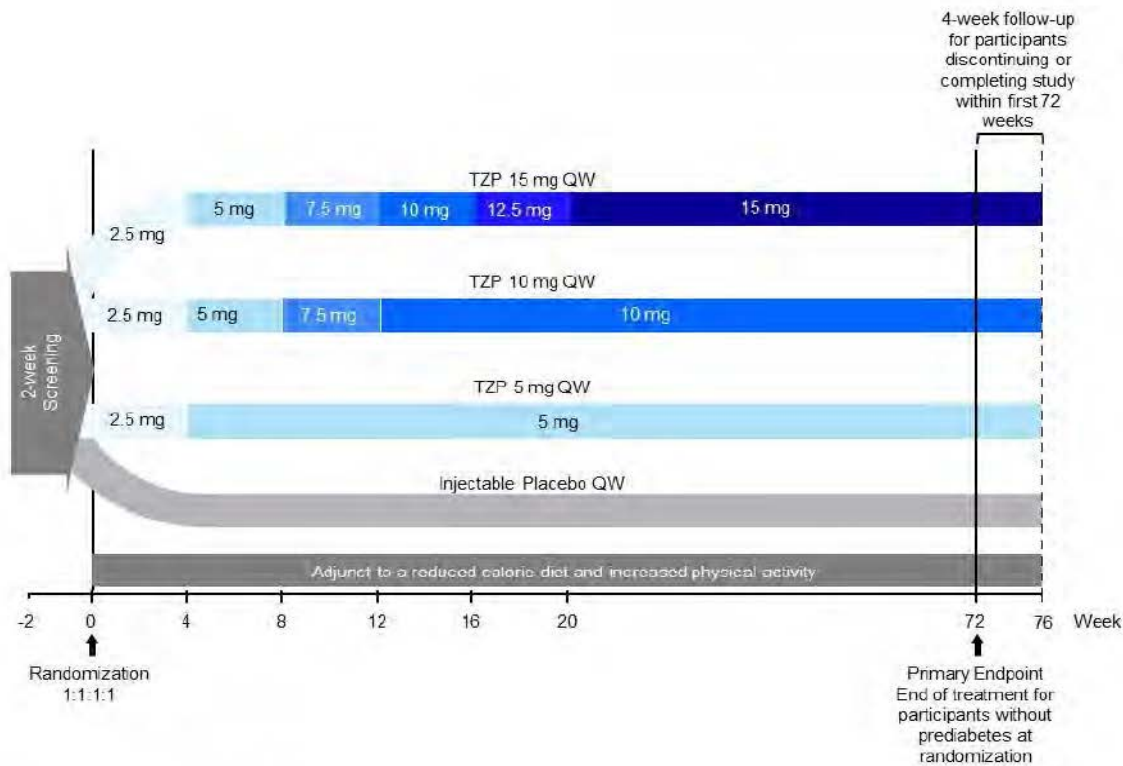
To be eligible, study subjects

- were 18 years or older
- either had
 - obesity, defined as having a BMI of 30 kg/m² or more, or
 - overweight, defined as having a BMI of 27 kg/m² or more, with at least 1 weight-related comorbid condition, including
 - obstructive sleep apnea (OSA)
 - hypertension
 - dyslipidemia, or
 - cardiovascular (CV) disease
- had a history of at least 1 self-reported unsuccessful dietary effort to lose weight

Subjects who had T2DM were not eligible for this study.

The trial design for this study is shown in Figure 1.

Figure 1: Study Schema



Abbreviations: QW = once weekly; TZP = tirzepatide.

[Source: Page 29 of Clinical Study Report (CSR)]

The primary study period included a

- 2-week screening period
- 72-week primary treatment period, and
- 4-week safety follow-up period for all subjects except for those with prediabetes at randomization continuing into the additional 2-year treatment period

Note that the 2-year treatment period for subjects with prediabetes at randomization is not in scope of this submission. Only data from the 72-week treatment period (the primary study period) in all study subjects were reported.

Dose escalation was double-blinded. For subjects randomly assigned to tirzepatide, the starting dose was 2.5 mg. The dose was increased by 2.5 mg every 4 weeks until reaching the maintenance dose. The dose escalation period lasted 20 weeks. Subjects in the placebo group received matching once-weekly placebo.

Subjects who developed intolerable gastrointestinal (GI) adverse events (AEs) during the treatment period could temporarily interrupt the study drug, omitting 1 dose, if the subjects had taken the last 3 weekly doses. After the temporary interruption, the participant was supposed to restart at the same dose either alone or with the participant taking symptomatic medication, per

Principal Investigator (PI) discretion, to alleviate their GI symptoms. If intolerable GI symptoms persisted, the tirzepatide-treated subjects could de-escalate study drug to a lower tolerated maintenance dose in a blinded fashion:

- Subjects at 5 mg or 2.5 mg could decrease to placebo
- Subjects at 7.5 mg or 10 mg could decrease to the dose to 5 mg, or
- Subjects at 12.5 mg or 15 mg could decrease to the dose to 10 mg

If intolerable GI symptoms continued to persist despite the above measures, the subjects were discontinued from the study drug.

The primary objective was

- to demonstrate that tirzepatide 10 mg QW was superior to placebo for percent change in body weight and percentage of subjects with $\geq 5\%$ body weight reduction at Week 72, and/or
- to demonstrate that tirzepatide 15 mg QW was superior to placebo for percent change in body weight and percentage of subjects with $\geq 5\%$ body weight reduction at Week 72

Primary endpoint (co-primary)

- Percent change in body weight from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve $\geq 5\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg) -> referred to as 5% responder in this review

Key secondary endpoints

- Percentage of study subjects who achieve $\geq 10\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg) -> referred to as 10% responder in this review
- Percentage of study subjects who achieve $\geq 15\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve $\geq 20\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg)
- Change in waist circumference (cm) from baseline at Week 72 (10 mg, 15 mg)
- Change in body weight (kg) from baseline at Week 20 (pooled 10 mg and 15 mg)
- Percent change in body weight from baseline at Week 72 and 5% responder in body weight from baseline at Week 72 (5 mg)
- Change in systolic blood pressure (mmHg) from baseline at Week 72 (pooled 5 mg, 10 mg, and 15 mg)
- Change in triglycerides (mg/dL) from baseline at Week 72 (pooled 5 mg, 10 mg, and 15 mg)
- Change in non-HDL-C (mg/dL) from baseline at Week 72 (pooled 5 mg, 10 mg, and 15 mg)
- Change in HDL-C (mg/dL) from baseline at Week 72 (pooled 5 mg, 10 mg, and 15 mg)
- Change in fasting insulin (pmol/L) from baseline at Week 72 (pooled 5 mg, 10 mg, and 15 mg)

- Change in SF-36v2 acute form Physical Functioning domain score from baseline to Week 72 (pooled 10 mg and 15 mg)

SURMOUNT-2

The trial was a multinational, multicenter, randomized, parallel-arm, placebo-controlled, double-blind, 72-week study that investigated the safety and efficacy of treatment with tirzepatide 10 mg and 15 mg QW compared to placebo QW, in conjunction with a reduced-calorie diet and increased physical activity, on weight management in subjects with T2DM who have obesity ($\text{BMI} \geq 30 \text{ kg/m}^2$) or are overweight ($\text{BMI} \geq 27 \text{ kg/m}^2$).

The study was conducted at 77 sites in 7 countries including the United States.

A total of 938 subjects were randomized 1:1:1 to tirzepatide 10 mg (N=312), 15 mg (N=311), or placebo (N=315) and the randomization was stratified by type of baseline antihyperglycemic medications (AHMs) (classified according to its potential effect on body weight), country and sex.

To be eligible, study subjects

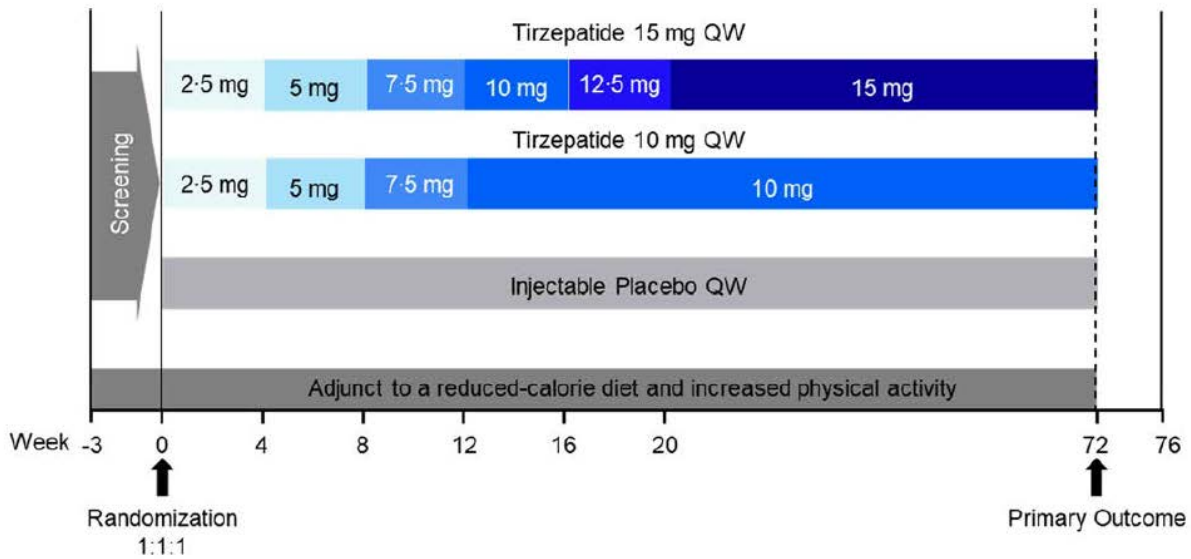
- were 18 years or older
- had a diagnosis of T2DM with glycosylated hemoglobin A1c (HbA1c) $\geq 7\%$ (53 mmol/mol) to $\leq 10\%$ (86 mmol/mol) at screening, on stable therapy for the last 3 months prior to screening (diet or exercise alone or any oral antihyperglycemic medications except dipeptidyl peptidase-4 (DPP-4) inhibitors and GLP-1 RA)
- $\text{BMI} \geq 27 \text{ kg/m}^2$
- had a history of at least 1 self-reported unsuccessful dietary effort to lose weight

The primary study period included a

- 3-week screening period
- 72-week primary treatment period, and
- 4-week safety follow-up

The trial design for this study is shown in Figure 2.

Figure 2: Study Schema



Abbreviation: QW = once weekly.

[Source: Page 25 of Clinical Study Report (CSR)]

Dose escalation was double-blinded. For subjects randomly assigned to tirzepatide, the starting dose was 2.5 mg. The dose was increased by 2.5 mg every 4 weeks until reaching the maintenance dose. The dose escalation period lasted 20 weeks. Subjects in the placebo group received matching once-weekly placebo.

Subjects who developed intolerable gastrointestinal (GI) adverse events (AEs) during the treatment period were advised to eat smaller meals, could be prescribed symptomatic medication, and could temporarily interrupt the study drug, omitting 1 dose, if the subjects had taken the last 3 weekly doses. After the temporary interruption, the participant was supposed to restart at the same dose either alone or with the participant taking symptomatic medication, per PI discretion, to alleviate their GI symptoms. If intolerable GI symptoms persisted, the tirzepatide-treated subjects could de-escalate the study drug to a lower tolerated maintenance dose in a blinded fashion:

- Subjects at ≤ 10 mg could decrease to placebo
- Subjects at 12.5 mg or 15 mg could decrease to the dose to 10 mg

If intolerable GI symptoms continued to persist despite the above measures, the subjects discontinued the study drug.

The primary objective was

- to demonstrate that tirzepatide 10 mg QW was superior to placebo for percent change in body weight and percentage of subjects with $\geq 5\%$ body weight reduction at Week 72, and/or
- to demonstrate that tirzepatide 15 mg QW was superior to placebo for percent change in body weight and percentage of subjects with $\geq 5\%$ body weight reduction at Week 72

Primary endpoint (co-primary)

- Percent change in body weight from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve $\geq 5\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg)

Key secondary endpoints

- Percentage of study subjects who achieve $\geq 10\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve $\geq 15\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg)
- Change in waist circumference (cm) from baseline at Week 72 (10 mg, 15 mg)
- Change in HbA1c (%) from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve HbA1c $< 7\%$ from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve HbA1c $\leq 6.5\%$ from baseline to Week 72 (10 mg, 15 mg)
- Change in fasting glucose (mg/dL) from baseline to Week 72 (10 mg, 15 mg)
- Change in non-HDL-C (mg/dL) from baseline at Week 72 (pooled 10 mg and 15 mg)
- Change in systolic blood pressure (mmHg) from baseline at Week 72 (pooled 10 mg and 15 mg)
- Change in triglycerides (mg/dL) from baseline at Week 72 (pooled 10 mg and 15 mg)
- Change in HDL-C (mg/dL) from baseline at Week 72 (pooled 10 mg and 15 mg)
- Percentage of study subjects who achieve HbA1c $\leq 5.7\%$ from baseline to Week 72 (10 mg, 15 mg)
- Percentage of study subjects who achieve $\geq 20\%$ body weight reduction from baseline to Week 72 (10 mg, 15 mg)

3.2.2 Statistical Methodologies

Primary estimand

All primary and key secondary efficacy assessments were guided by treatment-regimen estimand, which represented the efficacy irrespective of adherence to study drug. Treatment-regimen estimand was also referred to as “hybrid” estimand in this study, and under this estimand the intercurrent events were categorized into Category 1 and Category 2, where Category 1 was for missing data due to exceptional circumstances and Category 2 was for missing data due to all other intercurrent events.

Analysis population

Modified intent-to-treat (mITT): All randomly assigned subjects who were exposed to at least 1 dose of study drug.

Full Analysis Set (FAS): Data obtained during treatment period from mITT population, regardless of adherence to study drug. This was the primary analysis set for efficacy.

Safety Analysis Set (SAS): Data obtained during the treatment period plus safety follow-up period from mITT population, regardless of adherence to study drug. The SAS was used for safety analysis. The adverse events and other safety evaluations were reported by the treatment group to which subjects were randomly assigned.

In both studies, all subjects were treated as randomized.

Primary endpoint

Primary analysis

The analysis model for percent change in body weight was an analysis of covariance (ANCOVA) model with treatment and stratification groups as factors and baseline body weight (kg) as a covariate. In SURMOUNT-1, the stratification groups were country, sex and prediabetes status at randomization. In SURMOUNT-2, the stratification groups were country, sex and types of antihyperglycemic medication (AHM) at randomization. The estimated treatment differences between tirzepatide 10 mg and placebo, as well as tirzepatide 15 mg and placebo were reported with the associated 2-sided 95% confidence interval (CI) and the corresponding p-value.

The analysis model for 5% responder endpoint was a logistic regression model with treatment and stratification groups as factors and baseline body weight (kg) as a covariate. The estimated odds ratio between tirzepatide 10 mg and placebo, as well as tirzepatide 15 mg and placebo were reported with the associated 2-sided 95% confidence interval (CI) and the corresponding p-value.

Handling of missing Week 72 values

For efficacy analysis, missing values at Week 72 were imputed using hybrid imputation. Under the treatment-regimen estimand (also referred to as hybrid estimand), the intercurrent events and the resulting missing events were handled as below:

Category 1: Missing data due to exceptional circumstances

For missing data due to exceptional circumstances such as pandemic or natural disasters (after other reasons for missing data were ruled out), the analysis considered the missing data as missing at random. The missing data were imputed using all non-missing data of the primary outcome measurements from the same treatment arm.

Category 2: Missing data due to all other intercurrent events

For missing data due to all other intercurrent events, missing data were imputed based on the retrieved dropouts in the same treatment arm. Retrieved dropouts were defined as observed primary outcome measurements from subjects in the same treatment group who had their primary outcome measurements collected after early discontinuation of study drug.

The analyses of continuous data at the primary endpoint visit used an analyses of covariance model with missing endpoint imputed. Analysis of percentage of subjects achieving target thresholds used a logistic regression, with missing continuous values at primary endpoint imputed by multiple imputation followed by dichotomization.

For multiple imputation, a total of 100 datasets were generated and the final results were integrated using Rubin's rule.

Sensitivity analyses

- Retrieved dropouts multiple imputation (RDMI): Missing values of change in body weights at the 72-week were imputed based on observed body weight change from baseline values at the visit from subjects in the same treatment group who had their efficacy assessed after early discontinuation of study drug
- Placebo multiple imputation: Missing values of change in body weight at the 72-week visit were imputed based on observed body weight change from baseline values at the visit from subjects in the placebo group (SURMOUNT-2)
- Return to baseline imputation: Missing values of body weight at the 72-week visit were imputed using the return-to-baseline multiple imputation method to account for within subject variability (SURMOUNT-2)
- Tipping point analysis: Missing data were imputed according to the primary multiple imputation approach. Then a penalty was added to the imputed values at Week 72. The 2-dimensional tipping point analysis was performed to explore the values with which the conclusions would change.

Key secondary endpoint

All key secondary endpoints were analyzed using the same imputation approaches as those used for the primary endpoints and to address the treatment-regimen estimand. The imputation model was the same as the one used for the primary endpoints with body weight replaced by assessments of the endpoints to be analyzed. The statistical model for continuous endpoints was ANCOVA with factors and a covariate replaced by the baseline assessment of the endpoint to be analyzed. The statistical model for binary endpoints (such as 10% responder) was logistic regression with factors and a covariate replaced by the baseline assessment of the endpoint to be analyzed.

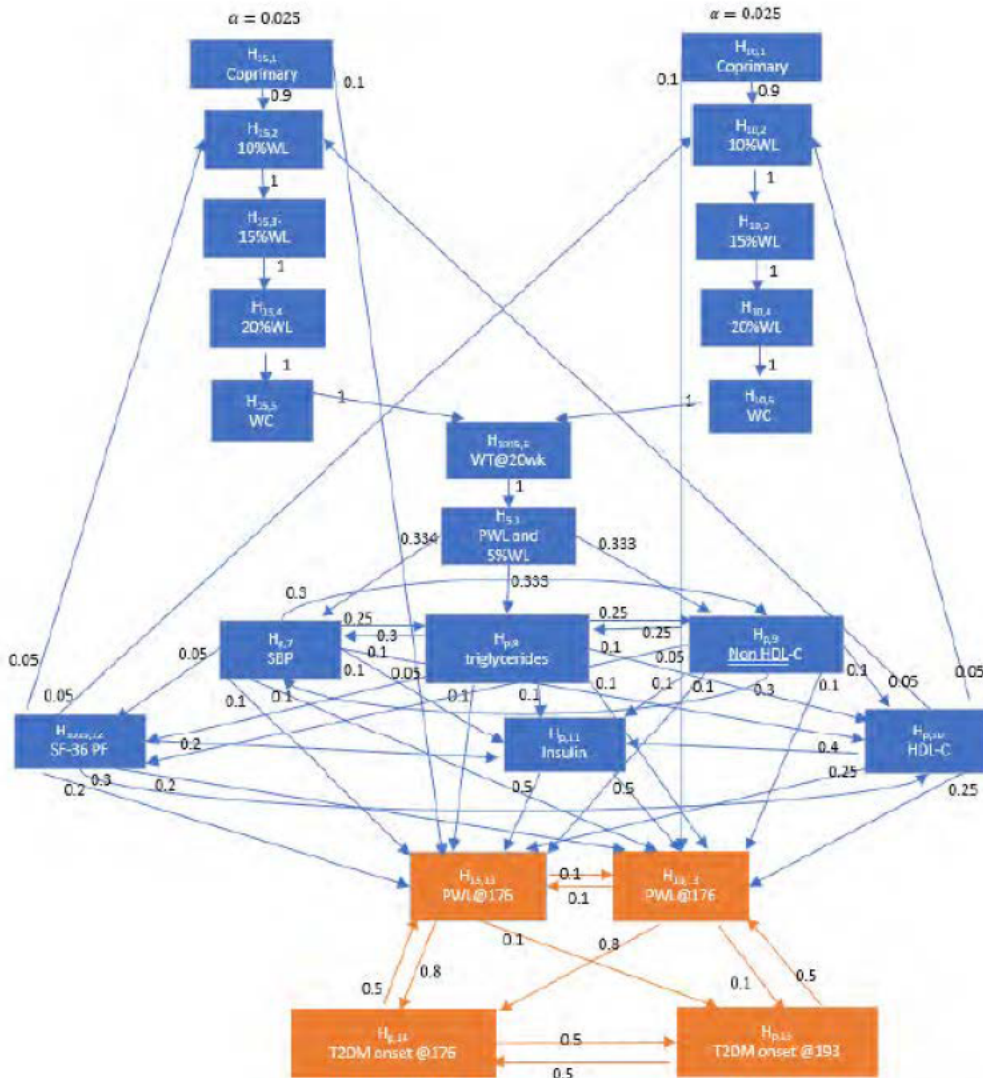
Subgroup analyses of the primary endpoints

Subgroup analyses were performed for age, sex, race, ethnicity, region, BMI group, and glycemic status at randomization in SURMOUNT-1. Similarly, subgroup analyses were performed for age, sex, race, ethnicity, region, BMI group, baseline HbA1c group in SURMOUNT-2. Subgroup analyses were carried out addressing the primary estimand using the same imputation approach and statistical models used in the analysis of the primary endpoints. The results of the subgroup analyses were presented with treatment contrasts and treatment-by-subgroup interaction p-values.

Multiplicity considerations

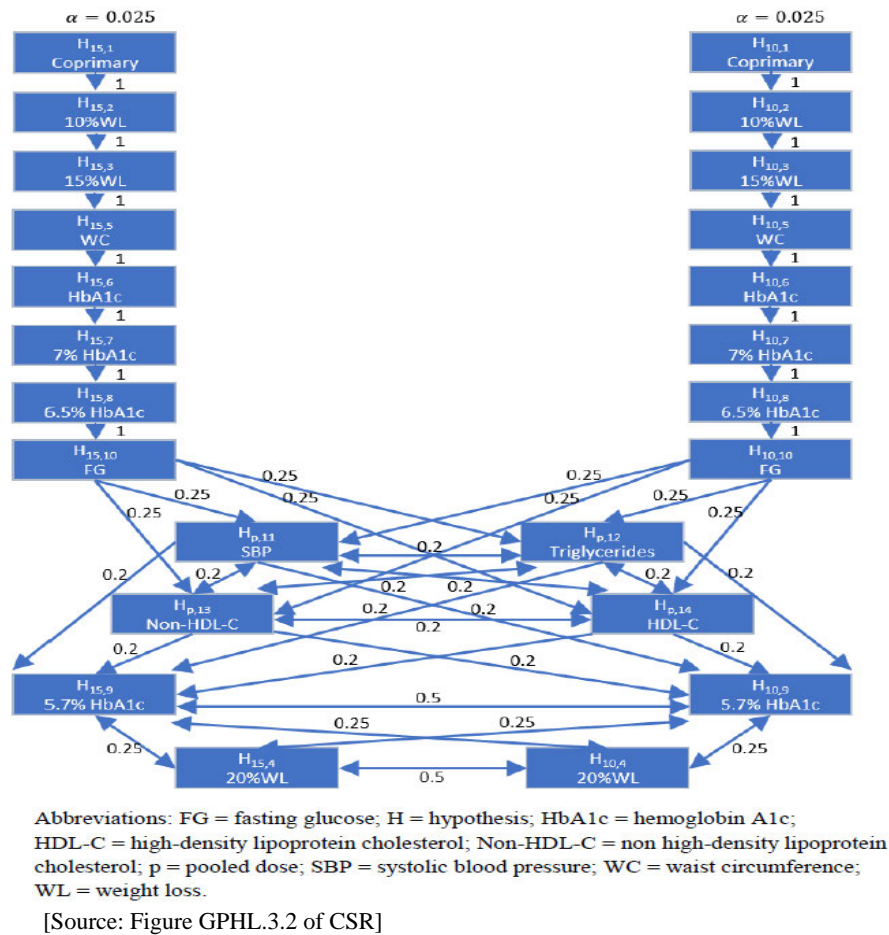
To control the overall Type 1 error rate at 0.05 across all primary and key secondary endpoints, a series of testing sequence was prespecified. Figure 3 and Figure 4 illustrate the testing sequence for each study. This testing strategy was implemented in both studies.

Figure 3: Type 1 error control strategy for primary end key secondary endpoints (SURMOUNT-1)



Abbreviations: HDL-C = high-density lipoprotein cholesterol; SBP = systolic blood pressure; SF-36v2 PF = Short Form survey-36 version 2 physical functioning; T2DM = type 2 diabetes mellitus; p = tirzepatide all doses combined; PwL = percent weight loss WC = waist circumference; wk = week, WT = weight loss in kg
 [Source: Figure GPHK.6.1 of CSR Appendix Statistical Methods]

Figure 4: Type 1 error control strategy for primary end key secondary endpoints (SURMOUNT-2)



3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Patient disposition

The summary of the subject disposition is given in Table 2 and Table 3. In SURMOUNT-1, the proportion of subjects who completed treatment was from 73.6 % to 85.7% across groups. The lowest completion rate (73.6%) was from subjects in the placebo group. In SURMOUNT 2, the proportion of subjects who completed treatment was from 85.1% to 90.7% across groups. The lowest completion rate (85.1%) was from subjects in the placebo group. In both trials, main reason for discontinuing treatment was withdrawal by subject and adverse event followed by various reasons listed in the tables. In both trials, the proportion of subjects who withdrew from the trial was highest in the placebo group.

Table 2: Patient Disposition in SURMOUNT-1

	TZP 5 mg	TZP 10 mg	TZP 15 mg	Placebo	Total
Randomized	630	636	630	643	2539
Full Analysis Set	630	636	630	643	2539
Completed treatment	540 (85.7%)	531 (83.5%)	535 (84.9%)	473 (73.6%)	2079 (81.9%)
Discontinued treatment	90 (14.3%)	104* (16.4%)	95 (15.1%)	170 (26.4%)	459 (18.1%)
adverse event	27 (4.3%)	45 (7.1%)	39 (6.2%)	17 (2.6%)	128 (5.0%)
Withdrawal by subject	34 (5.4%)	29 (4.6%)	28 (4.4%)	87 (13.5%)	178 (7.0%)
Lost to follow up	15 (2.4%)	14 (2.2%)	11 (1.7%)	38 (5.9%)	78 (3.1%)
Pregnancy	4 (0.6%)	6 (0.9%)	5 (0.8%)	4 (0.6%)	19 (0.7%)
Physician decision	2 (0.3%)	3 (0.5%)	2 (0.3%)	2 (0.3%)	9 (0.4%)
Protocol deviation	0	0	0	2 (0.3%)	2 (0.1%)
Other	5 (0.8%)	6 (0.9%)	9 (1.4%)	17 (2.6%)	37 (1.5%)
Death	3 (0.5%)	1 (0.2%)	1 (0.2%)	3 (0.5%)	8 (0.3%)
Completed trial	561 (89.0%)	562 (89.0%)	566 (89.8%)	495 (77.0%)	2184 (86.0%)
Discontinued study	69 (11.0%)	74 (11.6%)	64 (10.2%)	148 (23.0%)	355 (14.0%)
Adverse event	7 (1.1%)	11 (1.7%)	8 (1.3%)	7 (1.1%)	33 (1.3%)
Withdrawal by subject	28 (4.4%)	29 (4.6%)	22 (3.5%)	68 (10.6%)	147 (5.8%)
Lost to follow up	19 (3.0%)	20 (3.1%)	15 (2.4%)	41 (6.4%)	95 (3.7%)
Pregnancy	4 (0.6%)	3 (0.5%)	5 (0.8%)	4 (0.6%)	16 (0.6%)
Physician decision	0	3 (0.5%)	1 (0.2%)	0	4 (0.2%)
Protocol deviation	0	0	0	1 (0.2%)	1 (0.04%)
Missing weight at Week 72	0	1 (0.2%)	0	1 (0.2%)	2 (0.08%)
Other	7 (1.1%)	5 (0.8%)	12 (1.9%)	22 (3.4%)	46 (1.8%)
Death	4 (0.6%)	2 (0.3%)	1 (0.2%)	4 (0.6%)	11 (0.4%)

Abbreviation: TZP=tirzepatide; [Source: *N=104 instead of 105 because one subject in TZP 10 mg discontinued safety follow-up visit due to change residence city and did not complete treatment disposition form. Therefore, this subject was not included in the count of treatment discontinuation: Excerpted from Figure GPHK.4.1. of CSR, Statistical Reviewer, adsl.xpt]

Table 3: Patient Disposition in SURMOUNT-2

	TZP 10 mg	TZP 15 mg	Placebo	Total
Randomized	312	311	315	938
Full Analysis Set	312	311	315	938
Completed treatment	283 (90.7%)	268 (86.2%)	268 (85.1%)	819 (87.3%)
Discontinued treatment	29 (9.3%)	43 (13.8%)	47 (14.9%)	119 (12.7%)
adverse event	12 (3.8%)	23 (7.4%)	12 (3.8%)	47 (5.0%)
Withdrawal by subject	10 (3.2%)	8 (2.6%)	21 (6.7%)	39 (4.2%)
Lost to follow up	3 (1%)	9 (2.9%)	7 (2.2%)	19 (2.0%)
Pregnancy	0	2 (0.6%)	1 (0.3%)	3 (0.3%)
Physician decision	1 (0.3%)	0	0	1 (0.1%)
Protocol deviation	0	0	2 (0.6%)	2 (0.2%)
Other	1 (0.3%)	1 (0.3%)	4 (1.3%)	6 (0.6%)
Death	2 (0.6%)	0	0	2 (0.2%)
Completed trial	296 (94.9%)	282 (90.7%)	281 (89.2%)	859 (91.6%)
Discontinued study	16 (5.1%)	29 (9.3%)	34 (10.8%)	79 (8.4%)
Adverse event	1 (0.3%)	5 (1.6%)	5 (1.6%)	11 (1.2%)
Withdrawal by subject	7 (2.2%)	10 (3.2%)	16 (5.1%)	33 (3.5%)
Lost to follow up	4 (1.3%)	10 (3.2%)	8 (2.5%)	22 (2.3%)
Pregnancy	0	2 (0.6%)	2 (0.6%)	4 (0.4%)
Other	2 (0.6%)	2 (0.6%)	3 (1%)	7 (0.7%)
Death	2 (0.6%)	0	0	2 (0.2%)

Abbreviation: TZP=tirzepatide; [Source: Excerpted from Figure GPHL.4.1. of CSR, Statistical Reviewer, adsl.xpt]

Summary of COVID-19 Impact

Both trials were conducted during the COVID-19 pandemic,

- SURMOUNT-1: conducted between 04 December 2019 to 01 April 2022
- SURMOUNT-2: conducted between 29 March 2021 and 10 April 2023

To mitigate the impact of COVID-19, sites implemented changes to the conduct of the study. These changes included remote visits (phone calls or video conference), mobile home health visits, extended visit windows, and local laboratory testing. Missing data due to COVID-19 were considered missing at random for the statistical analysis and imputed as described in Section 3.2.2 of this review. Table 4 presents a summary of subjects who experienced listed events due to the COVID-19. The COVID-19 related study discontinuation or treatment discontinuation was low across all treatment groups in SURMOUNT-1 and none discontinued due to the COVID-19 in SURMOUNT-2. There were 3 deaths reported due to COVID in SURMOUNT-1 and no death reported in SURMOUNT-2 due to the COVID-19.

Table 4: Summary of COVID-19 Related Events

SURMOUNT-1					
	TZP 5 mg (N=630)	TZP 10 mg (N=636)	TZP 15 mg (N=630)	Placebo (N=643)	Total (N=2539)
Discontinued study	3 (0.5%)	2 (0.3%)	6 (1.0%)	4 (0.6%)	15 (0.6%)
Discontinued treatment	3 (0.5%)	2 (0.3%)	5 (0.8%)	8 (1.2%)	18 (0.7%)
≥1 Covid related adverse event	114 (18.1%)	109 (17.1%)	93 (14.8%)	109 (17.0%)	425 (16.7%)
≥1 Covid related serious adverse event	9 (1.4%)	101 (1.6%)	5 (0.8%)	10 (1.6%)	34 (1.3%)
Death	2 (0.3%)	0	1 (0.2%)	0	3 (0.1%)
SURMOUNT-2					
		TZP 10 mg (N=312)	TZP 15 mg (N=311)	Placebo (N=315)	Total (N=938)
Discontinued study		0	0	0	0
Discontinued treatment		0	0	0	0
≥1 Covid related adverse event		56 (17.9%)	38 (12.2%)	63 (20.0%)	157 (16.7%)
≥1 Covid related serious adverse event		1 (0.3%)	1 (0.3%)	2 (0.6%)	4 (0.4%)
Death		0	0	0	0

Abbreviation: TZP=tirzepatide; [Source: SURMOUNT-1=Table GPHK.5.3 and Table GPHK.6.1 of I8F-MC-GPHK CSR Appendix COVID-19; SURMOUNT-2=Table GPHL.5.3 and Table GPHL.5.10 of I8F-MC-GPHL CSR Appendix COVID-19]

Demographic and other baseline characteristics

Baseline demographics and other characteristics are shown in Table 5 to Table 8. In both trials, most baseline demographics and other characteristics were similarly distributed across all treatment groups.

In SURMOUNT-1, most of the subjects were females (67.5%) and white (70.6%). The mean age was 44.9 years of age, with 94.0% in the <65 years old stratum. For the ≥65 years of age, the tirzepatide 5.0 mg group had more subjects (8.3%) compared to the other groups (ranged from 4.9% to 5.6%). The proportion of subjects from the United States was 44.9%. The proportion of subjects who were prediabetes at baseline was 59.4%. The mean BMI was 38.0 which was similar across all treatment groups. The majority of subjects (59.9%) had BMI between 30 and 40 at baseline.

In SURMOUNT-2, males and females were similarly distributed across groups. Most of the subjects were white (75.7%). The mean age was 54.2 years of age, with 82.4% in the <65 years old stratum. Most of subjects were from Argentina (36.9%) and the United States (36.5%). The mean duration of type 2 diabetes was 8.5 years for the study population and the shortest duration was 8.0 years in the tirzepatide 15 mg group. The mean BMI was 36.1 at baseline which was

similar across all treatment groups. The majority of subjects (62.9%) had BMI between 30 and 40 at baseline.

Table 5: Baseline Demographics in SURMOUNT-1

	TZP 5 mg N=630	TZP 10 mg N=636	TZP 15 mg N=630	Placebo N=643	Total N=2539
Sex, n (%)					
Female	426 (67.6)	427 (67.1)	425 (67.5)	436 (67.8)	1714 (67.5)
Male	204 (32.4)	209 (32.9)	205 (32.5)	207 (32.2)	825 (32.5)
Age categories, n (%)					
<65	578 (91.7)	605 (95.1)	595 (94.4)	609 (94.7)	2387 (94.0)
≥65	52 (8.3)	31 (4.9)	35 (5.6)	34 (5.3)	152 (6.0)
Age, years					
Mean (SD)	45.6 (12.7)	44.7 (12.4)	44.9 (12.3)	44.4 (12.5)	44.9 (12.5)
Median	46	45	45	44	45
Q1, Q3	36, 55	35, 54	36, 54	35, 54	36, 54
Min, Max	18, 79	18, 75	18, 84	18, 78	18, 84
Race, n (%)					
American Indian or Alaskan Native	56 (8.9)	58 (9.1)	59 (9.4)	58 (9.0)	231 (9.1)
Asian	68 (10.8)	71 (11.2)	66 (10.5)	71 (11.0)	276 (10.9)
Black or African American	48 (7.6)	47 (7.4)	51 (8.1)	55 (8.6)	201 (7.9)
Native Hawaiian or Other Pacific Islander	2 (0.3)	2 (0.3)	3 (0.5)	2 (0.3)	9 (0.4)
White	447 (71.0)	452 (71.1)	443 (70.3)	450 (70.0)	1792 (70.6)
Multiple	9 (1.4)	6 (0.9)	8 (1.3)	7 (1.1)	30 (1.2)
Country/Region, n (%)					
Argentina	90 (14.3)	90 (14.2)	91 (14.4)	93 (14.5)	364 (14.3)
Brazil	59 (9.4)	61 (9.6)	60 (9.5)	59 (9.2)	239 (9.4)
China	9 (1.4)	7 (1.1)	7 (1.1)	7 (1.1)	30 (1.2)
India	9 (1.4)	9 (1.4)	6 (1.0)	8 (1.2)	32 (1.3)
Japan	30 (4.8)	30 (4.7)	31 (4.9)	33 (5.1)	124 (4.9)
Mexico	110 (17.5)	107 (16.8)	108 (17.1)	108 (16.8)	433 (17.1)
Russian Federation	29 (4.6)	30 (4.7)	27 (4.3)	32 (5.0)	118 (4.6)
Taiwan	12 (1.9)	15 (2.4)	16 (2.5)	15 (2.3)	58 (2.3)
United States	282 (44.8)	287 (45.1)	284 (45.1)	288 (44.8)	1141 (44.9)
Ethnicity, n (%)					
Hispanic or Latino	308 (48.9)	297 (46.7)	299 (47.5)	310 (48.2)	1214 (47.8)
Not Hispanic or Latino	276 (43.8)	286 (45.0)	280 (44.4)	281 (43.7)	1123 (44.2)
Missing	46 (7.3)	53 (8.3)	51 (8.1)	52 (8.1)	202 (8.0)

Abbreviations: N = number of patients randomized; Q1 = 25th percentile; Q3=75th percentile; SD = standard deviation; TZP=tirzepatide; [Source: Statistical Reviewer Analysis; adsl.xpt]

Table 6: Baseline Demographics in SURMOUNT-2

	TZP 10 mg N=312	TZP 15 mg N=311	Placebo N=315	Total N=938
Sex, n (%)				
Female	158 (50.6)	159 (51.1)	159 (50.5)	476 (50.7)
Male	154 (49.4)	152 (48.9)	156 (49.5)	462 (49.3)
Age categories, n (%)				
<65	258 (82.7)	257 (82.6)	258 (81.9)	773 (82.4)
≥65	54 (17.3)	54 (17.4)	57 (18.1)	165 (17.6)
Age, years				
Mean (SD)	54.3 (10.7)	53.6 (10.6)	54.7 (10.5)	54.2 (10.6)
Median	55	55	55	55
Q1, Q3	47, 61	46, 61	48, 62	47, 62
Min, Max	22, 85	22, 79	18, 80	18, 85
Race, n (%)				
Asian	44 (14.1)	42 (13.5)	39 (12.4)	125 (13.3)
Black or African American	33 (10.6)	22 (7.1)	22 (7.0)	77 (8.2)
Native Hawaiian or Other Pacific Islander	1 (0.3)	1 (0.3)	1 (0.3)	3 (0.3)
White	228 (73.1)	234 (75.2)	248 (78.7)	710 (75.7)
Multiple	6 (1.9)	12 (3.9)	5 (1.6)	23 (2.5)
Country/Region, n (%)				
Argentina	115 (36.9)	115 (37.0)	116 (36.8)	346 (36.9)
Brazil	29 (9.3)	30 (9.6)	30 (9.5)	89 (9.5)
India	8 (2.6)	9 (2.9)	9 (2.9)	26 (2.8)
Japan	24 (7.7)	22 (7.1)	21 (6.7)	67 (7.1)
Russian Federation	13 (4.2)	13 (4.2)	17 (5.4)	43 (4.6)
Taiwan	9 (2.9)	8 (2.6)	8 (2.5)	25 (2.7)
United States	114 (36.5)	114 (36.7)	114 (36.2)	342 (36.5)
Ethnicity, n (%)				
Hispanic or Latino	184 (59.0)	189 (60.8)	188 (59.7)	561 (59.8)
Not Hispanic or Latino	124 (39.7)	112 (36.0)	122 (38.7)	358 (38.2)
Not reported	4 (1.3)	10 (3.2)	5 (1.6)	19 (2.0)

Abbreviations: N = number of patients randomized; Q1=25th percentile; Q3=75th percentile; SD = standard deviation; TZP=tirzepatide; [Source: Statistical Reviewer Analysis; adsl.xpt]

Table 7: Baseline Characteristics in SURMOUNT-1

	TZP 5 mg N=630	TZP 10 mg N=636	TZP 15 mg N=630	Placebo N=643	Total N=2539
Prediabetes, n (%)					
No	383 (60.8)	374 (58.8)	377 (59.8)	373 (58.0)	1507 (59.4)
Yes	247 (39.2)	262 (41.2)	253 (40.2)	270 (42.0)	1032 (40.6)
BMI (kg/m ²) categories, n (%)					
<30	38 (6.0)	38 (6.0)	40 (6.3)	24 (3.7)	140 (5.5)
30-<35	241 (38.3)	209 (32.9)	199 (31.6)	227 (35.3)	876 (34.5)
35-<40	174 (27.6)	187 (29.4)	179 (28.4)	180 (28.0)	720 (28.4)
>=40	177 (28.1)	202 (31.8)	212 (33.7)	212 (33.0)	803 (31.6)
Body weight, kg					
Mean (SD)	102.9 (20.7)	105.8 (23.3)	105.6 (22.9)	104.8 (21.4)	104.8 (22.1)
Median	98.8	101.4	102.0	101.0	100.7
Q1, Q3	88.8, 114.1	88.3, 119.3	88.9, 118.5	89.7, 116.5	88.8, 116.8
Min, Max	64.0, 199.0	60.1, 206.2	65.7, 214.3	67.6, 202.3	60.1, 214.3
BMI, kg/m ²					
Mean (SD)	37.4 (6.6)	38.2 (7.0)	38.1 (6.7)	38.2 (6.9)	38.0 (6.8)
Median	36.0	36.5	36.9	36.4	36.5
Q1, Q3	32.5, 41	33.1, 41.9	33.1, 42	32.9, 42	32.9, 41.7
Min, Max	27.0, 68.9	27.2, 67.7	27.1, 67.5	26.9, 69.9	26.9, 69.9
Height, cm					
Mean (SD)	165.7 (9.0)	166.1 (9.3)	166.1 (9.7)	165.6 (9.3)	165.9 (9.3)
Median	165.0	165.0	165.1	165.0	165.0
Q1, Q3	159.4, 172	160, 172.1	159, 172.7	158.1, 171.4	159, 172
Min, Max	142.2, 193.0	144.0, 195.6	138.9, 192.4	131.5, 196.5	131.5, 196.5
Waist circumference, cm					
Mean (SD)	113.2* (14.3)	114.8 (15.8)	114.4 (15.6)	114.0 (14.9)	114.1 (15.2)
Median	111.6	112.9	112.6	112.0	112.1
Q1, Q3	103, 121	103, 124	102.5, 124	104, 123	103, 123
Min, Max	82.0, 177.3	80.0, 191.0	76.2, 169.0	80.0, 177.5	76.2, 191.0
Systolic blood pressure, mmHg					
Mean (SD)	123.6 (12.5)	123.8 (12.8)	123.0 (12.9)	122.9 (12.8)	123.3 (12.7)
Median	123.0	123.0	122.8	122.7	123.0
Q1, Q3	114.3, 132.3	115, 132.7	113.6, 131.7	113, 130.3	114, 131.7
Min, Max	91.3, 179.3	92.3, 184.3	88.0, 162.0	91.0, 159.3	88.0, 184.3
Diastolic blood pressure, mmHg					
Mean (SD)	79.3 (8.1)	79.9 (8.3)	79.3 (8.2)	79.6 (8.0)	79.5 (8.2)
Median	79.3	80.3	79.7	80.0	80.0
Q1, Q3	74, 84.7	74, 85.3	73.7, 85	74.3, 85.3	74, 85
Min, Max	52.7, 98.3	57.7, 103.3	53.0, 112.0	58.0, 101.3	52.7, 112.0

Abbreviations: N = number of patients randomized; Q1=25th percentile; Q3=75th percentile; SD = standard deviation;
TZP=tirzepatide; *N=629 for waist circumference; [Source: Statistical Reviewer Analysis; adsl.xpt]

Table 8: Baseline Characteristics in SURMOUNT-2

	TZP 10 mg N=312	TZP 15 mg N=311	Placebo N=315	Total N=938
Duration of type 2 diabetes, years				
Mean (SD)	8.8 (6.9)	8.0 (6.4)	8.8 (6.2)	8.5 (6.5)
Median	7.2	6.4	7.6	7.1
Q1, Q3	3.4, 12.5	3.3, 11.5	4.3, 11.7	3.5, 11.7
Min, Max	0.3, 49.6	0.2, 50.7	0.3, 33.7	0.2, 50.7
BMI (kg/m²) categories, n (%)				
<30	60 (19.2)	51 (16.4)	52 (16.5)	163 (17.4)
30-<35	92 (29.5)	114 (36.7)	105 (33.3)	311 (33.2)
35-<40	94 (30.1)	85 (27.3)	71 (22.5)	250 (26.7)
>=40	66 (21.2)	61 (19.6)	87 (27.6)	214 (22.8)
Body weight, kg				
Mean (SD)	100.9 (20.9)	99.6 (20.1)	101.7 (22.3)	100.7 (21.1)
Median	98.0	96.3	96.7	97.2
Q1, Q3	87.8, 111.2	86.3, 111.2	86, 115.2	86.6, 112.8
Min, Max	62.4, 209.5	58.5, 219.8	55.1, 195.9	55.1, 219.8
BMI, kg/m²				
Mean (SD)	36.0 (6.4)	35.7 (6.1)	36.6 (7.3)	36.1 (6.6)
Median	35.1	34.7	35.0	34.9
Q1, Q3	30.9, 39.2	31.4, 38.4	31.0, 40.8	31.1, 39.5
Min, Max	26.1, 66.1	26.1, 62.3	26.3, 74.4	26.1, 74.4
Height, cm				
Mean (SD)	167.3 (9.1)	166.9 (10.4)	166.5 (9.9)	166.9 (9.8)
Median	167.4	166.0	167.0	167.0
Q1, Q3	161, 173	160, 174	160, 173.4	160, 173.5
Min, Max	144.0, 195.6	141.5, 209.0	140.2, 192.0	140.2, 209.0
Waist circumference, cm				
Mean (SD)	114.2 (14.1)	114.6 (13.1)	116.0 (15.7)	114.9 (14.4)
Median	113.0	113.5	113.5	113.3
Q1, Q3	104.3, 122.1	105.5, 122	105, 125	105, 122.9
Min, Max	87.0, 173.5	83.0, 167.0	83.0, 172.0	83.0, 173.5
HbA1c, %				
Mean (SD)	8.0 (0.8)	8.1 (1.0)	8.0 (0.8)	8.0 (0.9)
Median	7.9	7.9	7.8	7.9
Q1, Q3	7.4, 8.5	7.4, 8.7	7.4, 8.6	7.4, 8.6
Min, Max	5.9, 10.6	5.8, 11.7	6.0, 10.8	5.8, 11.7
Fasting serum glucose, mg/dL				
Mean (SD)	158.3 (44.0)	161.2 (49.3)	158.5 (46.5)	159.3 (46.6)
Median	149.5	151.3	151.3	149.5
Q1, Q3	126.1, 185.6	129.7, 183.8	126.1, 183.8	126.1, 183.8
Min, Max	57.6, 383.7	75.7, 387.3	66.7, 396.3	57.6, 396.3
Systolic blood pressure, mmHg				
Mean (SD)	130.6 (12.2)	130.0 (12.3)	131.0 (11.9)	130.5 (12.1)
Median	130.5	131.0	132.0	131.0
Q1, Q3	122, 139	121.5, 138.5	123, 139.5	122.5, 139
Min, Max	99.0, 180.0	90.0, 162.0	98.5, 163.0	90.0, 180.0
Diastolic blood pressure, mmHg				
Mean (SD)	80.2 (8.1)	79.7 (8.7)	79.4 (8.4)	79.8 (8.4)
Median	80.5	80.5	79.0	80.0
Q1, Q3	74.5, 85.5	74, 85.5	73.5, 85	74, 85.5
Min, Max	59.0, 97.5	50.0, 100.0	58.0, 107.0	50.0, 107.0

Abbreviations: N = number of patients randomized; Q1=25th percentile; Q3=75th percentile; SD = standard deviation; TZP=tirzepatide; [Source: Statistical Reviewer Analysis; adsl.xpt]

3.2.4 Results and Conclusions

Missing Data

The amount of missing data at Week 72 are shown in Table 9 for each trial. Across trials, the proportion of missing data ranged from 4.8% to 10.5% for the tirzepatide groups. The proportion of missing data for the placebo group was 21.6% in SURMOUNT-1 and 11.1% in SURMOUNT-2. Most missing values were from Category 2 missing in SURMOUNT-1. In SURMOUNT-2, all missing values were from Category 2 missing. For the primary efficacy analyses, missing body weight measurements at Week 72 were imputed using hybrid imputation approach as described in Section 3.2.2 of this review.

Table 9: Summary of Observed/Missing Data

SURMOUNT-1				
	Tirzepatide 5 mg (N=630)	Tirzepatide 10 mg (N=636)	Tirzepatide 15 mg (N=630)	Placebo (N=643)
Observed	566 (89.8%)	569 (89.5%)	571 (90.6%)	504 (78.4%)
Completed treatment and observed	538 (85.4%)	530 (83.3%)	535 (84.9%)	471 (73.3%)
Retrieved dropouts	28 (4.4%)	39 (6.1%)	36 (5.7%)	33 (5.1%)
Missing	64 (10.2%)	67 (10.5%)	59 (9.4%)	139 (21.6%)
Category 1 missing	3 (0.5%)	3 (0.5%)	5 (0.8%)	6 (0.9%)
Category 2 missing	61 (9.7%)	64 (10.1%)	54 (8.6%)	133 (20.7%)
SURMOUNT-2				
		Tirzepatide 10 mg (N=312)	Tirzepatide 15 mg (N=311)	Placebo (N=315)
Observed		297 (95.2%)	285 (91.6%)	280 (88.9%)
Completed treatment and observed		282 (90.4%)	265 (85.2%)	264 (83.8%)
Retrieved dropouts		15 (4.8%)	20 (6.4%)	16 (5.1%)
Missing		15 (4.8%)	26 (8.4%)	35 (11.1%)
Category 1 missing		0	0	0
Category 2 missing		15 (4.8%)	26 (8.4%)	35 (11.1%)

Abbreviations: N=number of subjects randomized; cell content shows frequency and percentage relative to N in the parentheses; [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt]

Primary endpoints results

In SURMOUNT-1, both tirzepatide dose groups resulted in statistically significant greater percent reductions in body weight compared to placebo. The treatment effect on % change in body weight was -16.4% in the tirzepatide 10 mg group and -17.8% in the tirzepatide 15 mg group. The proportion of subjects who had at least 5% body weight loss was greater with statistical significance in the tirzepatide groups compared to the placebo group (Table 10).

In SURMOUNT-2, both tirzepatide dose groups resulted in statistically significant greater percent reductions in body weight compared to placebo. The treatment effect on % change in body weight was -9.6% in the tirzepatide 10 mg group and -11.6% in the tirzepatide 15 mg group. The proportion of subjects who had at least 5% body weight loss was statistically significant and greater in the tirzepatide groups compared to the placebo group (Table 11).

Table 10: Primary analysis: %Change in body weight and 5% body weight loss in SURMOUNT-1

	Tirzepatide 10 mg N=636	Tirzepatide 15 mg N=630	Placebo N=643
SURMOUNT-1			
%Change in body weight			
Baseline body weight (kg), mean (SD)	105.8 (23.3)	105.6 (22.9)	104.8 (21.4)
Estimate, LSMean (SE) ¹	-19.5 (0.5)	-20.9 (0.5)	-3.1 (0.6)
Difference from placebo, (95% CI)	-16.4 (-17.9, -14.8)	-17.8 (-19.3, -16.3)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)
5% Body weight loss			
Proportion estimate ²	89.6	91.4	34.0
Difference from placebo (95% CI)	55.6 (50.0, 61.2)	57.4 (51.9, 62.9)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)

Abbreviations: N=number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline body weight (kg) as a covariate; ²Estimates based on a logistic regression using treatment and stratification group as factors and baseline body weight (kg) as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt]

Table 11: %Change in body weight and 5% body weight loss in SURMOUNT-2

	Tirzepatide 10 mg N=312	Tirzepatide 15 mg N=311	Placebo N=315
SURMOUNT-2			
%Change in body weight			
Baseline body weight (kg), mean (SD)	100.9 (20.9)	99.6 (20.1)	101.7 (22.3)
Estimate, LSMean (SE) ¹	-12.8 (0.6)	-14.7 (0.5)	-3.2 (0.5)
Difference from placebo, (95% CI)	-9.6 (-11.1, -8.1)	-11.6 (-13.0, -10.1)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)
5% Body weight loss			
Proportion estimate ²	79.6	83.2	32.0
Difference from placebo (95% CI)	47.6 (40.2, 55.0)	51.2 (43.9, 58.4)	
P-value	<0.001	<0.001	
Missing, n(%)	15(4.8)	26 (8.4)	35 (11.1)

Abbreviations: N=number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline body weight (kg) as a covariate; ²Estimates based on a logistic regression using treatment and stratification group as factors and baseline body weight (kg) as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt]

Pre-specified sensitivity analyses using different imputation approaches were conducted to evaluate the robustness of the conclusions based on the primary analyses. All sensitivity analyses yielded results that were consistent with the primary analyses results.

Two-dimensional tipping point analyses with incremental changes ranging from -30% to 30% applied to imputed values at Week 72 were performed for %change in body weight. The conclusion of the primary analysis was not overturned in the range of the incremental changes explored, supporting the robustness of the conclusion based on the primary analysis.

Key secondary endpoint results

SURMOUNT-1

Both tirzepatide 10 mg and 15 mg doses achieved superiority compared with placebo for the proportions of subjects who had at least 10%, 15% and 20% body weight reduction from baseline to Week 72. Both tirzepatide 10 mg and 15 mg doses achieved superiority compared with placebo for mean change in waist circumference at Week 72 (Table 12).

Table 12: Key Secondary Endpoints in SURMOUNT-1 (at Week 72)

	Tirzepatide 10 mg N=636	Tirzepatide 15 mg N=630	Placebo N=643
SURMOUNT-1			
10% Body weight loss			
Proportion estimate ¹	79.0	84.3	18.1
Difference from placebo (95% CI)	61.0 (55.7, 66.3)	66.3 (61.2, 71.4)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)
15% Body weight loss			
Proportion estimate ¹	67.5	71.4	8.3
Difference from placebo (95% CI)	59.2 (54.3, 64.1)	63.1 (58.3, 68.0)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)
20% Body weight loss			
Proportion estimate ¹	50.2	57.0	2.8
Difference from placebo (95% CI)	47.5 (42.8, 52.1)	54.2 (49.6, 58.8)	
P-value	<0.001	<0.001	
Missing, n(%)	67 (10.5)	59 (9.4)	139 (21.6)
Change in Waist Circumference			
Baseline (cm), mean (SD)	114.8 (15.8)	114.4 (15.6)	114.0 (14.9)
Estimate, LSMean (SE) ²	-17.7 (0.5)	-18.5 (0.4)	-4.0 (0.6)
Difference from placebo (95% CI)	-13.8 (-15.2, -12.3)	-14.5 (-15.9, -13.0)	
P-value	<0.001	<0.001	
Missing, n(%)	68 (10.7)	59 (9.4)	139 (21.6)

Abbreviations: N = number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Estimates based on a logistic regression using treatment and stratification group as factors and baseline value as a covariate; ²Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline value as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adccovid.xpt, adds.xpt]

Pooled tirzepatide 10 mg and 15 mg (10/15 mg) resulted in a statistically significant greater reduction in body weight at Week 20 compared to placebo (Table 13).

Table 13: Key Secondary Endpoint in SURMOUNT-1 (at Week 20)

	Tirzepatide 10/15 mg N=1266	Placebo N=643
SURMOUNT-1		
Change in body weight		
Baseline body weight (kg), mean (SD)	105.7 (23.1)	104.8 (21.4)
Estimate, LSMean (SE) ¹	-12.8 (0.2)	-2.7 (0.2)
Difference from placebo (95% CI)	-10.1 (-10.7, -9.6)	
P-value	<0.001	
Missing, n(%)	78 (6.2)	50 (7.8)

Abbreviations: N = number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline value as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adccovid.xpt, adds.xpt]

Tirzepatide 5 mg achieved superiority compared with placebo for %change in body weight and the proportions of subjects who had at least 5% body weight reduction from baseline to Week 72 (Table 14).

Table 14: Key Secondary Endpoints in SURMOUNT-1 (at Week 72)

SURMOUNT-1	Tirzepatide	Placebo
	5 mg N=630	N=643
%Change in body weight		
Baseline body weight (kg), mean (SD)	102.9 (20.7)	104.8 (21.4)
Estimate, LSMean (SE) ¹	-15.0 (0.4)	-3.1 (0.6)
Difference from placebo (95% CI)	-11.9 (-13.4, -10.4)	
P-value	<0.001	
Missing, n(%)	64 (10.2)	139 (21.6)
5% Body weight loss		
Proportion estimate ²	85.5	34.0
Difference from placebo (95% CI)	51.5 (45.7, 57.4)	
P-value	<0.001	
Missing, n(%)	64 (10.2)	139 (21.6)

Abbreviations: N = number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline value as a covariate; ²Estimates based on a logistic regression using treatment and stratification group as factors and baseline value as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adccovid.xpt, adds.xpt]

Pooled tirzepatide 5 mg, 10 mg, and 15 mg (tirzepatide 5/10/15 mg) was compared to placebo for selected cardiometabolic parameters included in the key secondary endpoints, and the results are shown in Table 15. Fasting insulin is shown in a conventional unit (mIU/L) because the conventional unit is a clinically preferred unit. Similarly, %change from baseline at Week 72 (shown with a subscript _{pctchange}) was listed along with change from baseline at Week 72 (shown with a subscript _{change}) for lipid parameters and fasting insulin because %change from baseline results are considered clinically more informative. For all parameters listed in Table 15, the difference between the pooled tirzepatide group and the placebo group was statistically significant, in favor of the pooled tirzepatide group.

Table 15: Key Secondary Endpoints in SURMOUNT-1 (at Week 72)

	Tirzepatide 5/10/15 mg N=1896	Placebo N=643
SURMOUNT-1		
Change in SBP		
Baseline (mmHg), mean (SD)	123.5 (12.7)	122.9 (12.8)
Estimate, LSMean (SE) ¹	-7.2 (0.3)	-1.0 (0.7)
Difference from placebo (95% CI)	-6.2 (-7.7, -4.8)	
P-value	<0.001	
Missing, n(%)	201 (10.6)	146 (22.7)
Triglyceride		
Baseline ² (mg/dL)	127.5	130.8
Estimate _{change} , LSMean (SE) ¹	-31.8 (1.0)	-7.2 (2.9)
Difference _{change} from placebo (95% CI)	-24.6 (-24.6, -24.5)	
Estimate _{pctchange} , LSMean (SE) ¹	-24.8 (0.8)	-5.6(2.2)
Difference _{pctchange} from placebo (95% CI)	-20.3 (-24.3, -16.1)	
P-value	<0.001	
Missing, n(%)	239 (12.6)	154 (24.0)
Non-HDL-C		
Baseline ² (mg/dL)	138.3	138.3
Estimate _{change} , LSMean (SE) ¹	-13.4 (0.7)	-3.2 (1.8)
Difference _{change} from placebo, (95% CI)	-10.2 (-10.2, -10.2)	
Estimate _{pctchange} , LSMean (SE) ¹	-9.7 (0.4)	-2.3 (1.1)
Difference _{pctchange} from placebo (95% CI)	-7.5 (-9.6, -5.4)	
P-value	<0.001	
Missing, n(%)	244 (12.9)	153 (23.8)
HDL-C		
Baseline ² (mg/dL)	47.6	46.6
Estimate _{change} , LSMean (SE) ¹	3.8 (0.3)	-0.3 (0.5)
Difference _{change} from placebo, (95% CI)	4.1 (4.1, 4.1)	
Estimate _{pctchange} , LSMean (SE) ¹	8.0 (0.4)	-0.7 (1.0)
Difference _{pctchange} from placebo (95% CI)	8.8 (6.5, 11.0)	
P-value	<0.001	
Missing, n(%)	242 (12.8)	153 (23.8)
Fasting insulin		
Baseline ² (mIU/L)	11.7	12.0
Estimate _{change} , LSMean (SE) ¹	-5.1 (0.1)	-0.8 (0.5)
Difference _{change} from placebo (95% CI)	-4.3 (-4.4, -4.2)	
Estimate _{pctchange} , LSMean (SE) ¹	-42.9 (1.0)	-6.6 (4.5)
Difference _{pctchange} from placebo (95% CI)	-38.9 (-44.8, -32.4)	
P-value	<0.001	
Missing, n(%)	297 (15.7)	185 (28.8)

Abbreviations: N = number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline body weight (kg) as a covariate; ²Geometric mean, and log transformation was applied to raw data; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt, adlbcn.xpt]

Treatment with the pooled tirzepatide 10/15 mg resulted in a statistically significant greater improvement in SF-36v2 physical functioning domain score compared to placebo (Table 16). The treatment effect was 1.9. Currently, the treatment effect along with the instrument (SF-26v2

physical functioning domain score) is being reviewed by Patient Focused Statistical Support Team within Division of Biometrics III.

Table 16: Key Secondary Endpoint in SURMOUNT-1 (at Week 72)

SURMOUNT-1	Tirzepatide	Placebo
	10/15 mg N=1266	N=643
Change in SF-36v2		
Baseline, mean (SD)	49.5 (0.2)	49.6 (0.3)
Estimate, LSMean (SE) ¹	3.6 (0.2)	1.7 (0.5)
Difference from placebo (95% CI)	1.9 (1.0, 2.9)	
P-value	<0.001	
Missing, n(%)	141 (11.1)	148 (23.0)

Abbreviations: N=number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline value as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt, adsf36.xpt]

SURMOUNT-2

Both tirzepatide 10 mg and 15 mg doses achieved superiority compared with placebo for all endpoints listed in Table 17. The proportions of subjects who had at least 10% and 15% body weight reduction from baseline to Week 72 were greater in the tirzepatide groups compared to the placebo group. Mean change in waist circumference at Week 72 was greater in the tirzepatide groups compared to the placebo group. Mean change in HbA1c at Week 72 was greater in the tirzepatide groups compared to the placebo group. The proportions of subjects who achieved HbA1c <7% and HbA1c ≤6.5% at Week 72 were greater in the tirzepatide groups compared to the placebo group. Mean change in fasting glucose at Week 72 was greater in the tirzepatide groups compared to the placebo group.

Table 17: Key Secondary Endpoints in SURMOUNT-2 (at Week 72)

	Tirzepatide 10 mg N=312	Tirzepatide 15 mg N=311	Placebo N=315
SURMOUNT-2			
10% Body weight loss			
Proportion estimate ¹	61.0	65.4	8.9
Difference from placebo (95% CI)	52.1 (45.4, 58.8)	56.5 (49.8, 63.1)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)
15% Body weight loss			
Proportion estimate ¹	39.0	47.8	2.5
Difference from placebo (95% CI)	36.6 (30.5, 42.6)	45.4 (39.1, 51.6)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)
Change in waist circumference			
Baseline (cm), mean (SD)	114.2 (14.1)	114.6 (13.1)	116.0 (15.7)
Estimate, LSMean (SE) ²	-10.8 (0.6)	-13.1 (0.5)	-3.3 (0.5)
Difference from placebo (95% CI)	-7.4 (-8.9, -5.9)	-9.8 (-11.2, -8.3)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)
Change in HbA1c (%)			
Baseline, mean (SD)	8.0 (0.8)	8.1 (1.0)	8.0 (0.8)
Estimate, LSMean (SE) ²	-2.1 (0.1)	-2.1 (0.1)	-0.5 (0.1)
Difference from placebo (95% CI)	-1.6 (-1.7, -1.4)	-1.6 (-1.8, -1.4)	
P-value	<0.001	<0.001	
Missing, n(%)	18 (5.8)	31 (10.0)	42 (13.3)
HbA1c <7%			
Proportion estimate ¹	87.3	85.4	35.4
Difference from placebo (95% CI)	51.9 (44.5, 59.4)	50.0 (42.1, 57.9)	
P-value	<0.001	<0.001	
Missing, n(%)	18 (5.8)	31 (10.0)	42 (13.3)
HbA1c ≤6.5%			
Proportion estimate ¹	80.4	81.0	18.8
Difference from placebo (95% CI)	61.7 (54.6, 68.8)	62.3 (55.0, 69.6)	
P-value	<0.001	<0.001	
Missing, n(%)	18 (5.8)	31 (10.0)	42 (13.3)
Change in fasting glucose			
Baseline (mg/dL), mean (SD)	158.3 (44.0)	161.2 (49.3)	158.5 (46.5)
Estimate, LSMean (SE) ²	-48.9 (2.1)	-48.8 (2.3)	-11.0 (2.3)
Difference from placebo (95% CI)	-37.9 (-44.1, -31.8)	-37.9 (-44.4, -31.4)	
P-value	<0.001	<0.001	
Missing, n(%)	19 (6.1)	30 (9.6)	41 (13.0)

Abbreviations: N=number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Estimates based on a logistic regression using treatment and stratification group as factors and baseline value as a covariate; ²Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline value as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt, ada1c.xpt, adlbcn.xpt]

Pooled tirzepatide 10 mg and 15 mg (tirzepatide 10/15 mg) was compared to placebo for selected cardiometabolic parameters included in the key secondary endpoints, and the results are shown in Table 18. For its practical utility, %change from baseline at Week 72 (shown with a subscript pctchange) was listed along with change from baseline at Week 72 (shown with a subscript change)

for lipid parameters. For all parameters listed in Table 18, the difference between the pooled tirzepatide group and the placebo group was statistically significant, in favor of the pooled tirzepatide group.

Table 18: Key Secondary Endpoints in SURMOUNT-2 (at Week 72)

	Tirzepatide 10/15 mg N=623	Placebo N=315
SURMOUNT-2		
Triglyceride		
Baseline ¹ (mg/dL)	158.6	165.0
Estimate _{change} , LSMean (SE) ²	-43.7 (2.1)	-5.4 (4.5)
Difference _{change} from placebo, (95% CI)	-38.4 (-38.4, -38.3)	
Estimate _{pctchange} , LSMean (SE) ¹	-27.2 (1.3)	-3.3 (2.8)
Difference _{pctchange} from placebo (95% CI)	-24.7 (-29.5, -19.5)	
P-value	<0.001	
Missing, n(%)	47 (7.5)	39 (12.4)
HDL-C		
Baseline ¹ (mg/dL)	43.0	42.7
Estimate _{change} , LSMean (SE) ²	3.9 (0.4)	0.1 (0.6)
Difference _{change} from placebo (95% CI)	3.7 (3.7, 3.8)	
Estimate _{pctchange} , LSMean (SE) ¹	9.0 (0.9)	0.2 (1.4)
Difference _{pctchange} from placebo (95% CI)	8.7 (5.3, 12.2)	
P-value	<0.001	
Missing, n(%)	47 (7.5)	39 (12.4)
Non-HDL-C		
Baseline ¹ (mg/dL)	124.6	129.6
Estimate, LSMean (SE) ²	-7.5 (1.4)	4.6 (2.3)
Difference from placebo (95% CI)	-12.1 (-12.1, -12.0)	
Estimate, LSMean (SE) ¹	-5.9 (1.1)	3.7 (1.8)
Difference from placebo (95% CI)	-9.2 (-13.0, -5.3)	
P-value	<0.001	
Missing, n(%)	47 (7.5)	39 (12.4)
Change in SBP		
Baseline (mmHg), mean (SD)	130.3 (12.2)	131.0 (11.9)
Estimate, LSMean (SE) ²	-6.3 (0.5)	-1.2 (0.9)
Difference from placebo (95% CI)	-5.2 (-7.2, -3.1)	
P-value	<0.001	
Missing, n(%)	44 (7.1)	38 (12.1)

Abbreviations: N=number of subjects randomized; SD=standard deviation; LSMean=least squares mean; SE=standard error; CI=confidence interval; ¹Geometric mean, and log transformation was applied to raw data; ²Model based estimates and standard error, the ANCOVA model included treatment and stratification group as factors and baseline body weight (kg) as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt, ada1c.xpt, adlbcn.xpt]

Both tirzepatide 10 mg and 15 mg doses achieved superiority compared with placebo for endpoints listed in Table 19. The proportion of subjects who achieved HbA1c <5.7% at Week 72 was greater in the tirzepatide groups compared to the placebo group. The proportion of subjects who had at least 20% body weight reduction from baseline to Week 72 was greater in the tirzepatide groups compared to the placebo group.

Table 19: Key Secondary Endpoints in SURMOUNT-2 (at Week 72)

SURMOUNT-2	Tirzepatide	Tirzepatide	Placebo
	10 mg N=312	15 mg N=311	N=315
HbA1c <5.7%			
Proportion estimate ¹	45.5	48.8	3.5
Difference from placebo (95% CI)	42.1 (35.7, 48.5)	45.3 (38.8, 51.9)	
P-value	<0.001	<0.001	
Missing, n(%)	18 (5.8)	31 (10.0)	42 (13.3)
20% Body weight loss			
Proportion estimate ¹	20.2	29.5	1.0
Difference from placebo (95% CI)	19.2 (14.4, 24.1)	28.5 (23.0, 34.0)	
P-value	<0.001	<0.001	
Missing, n(%)	15 (4.8)	26 (8.4)	35 (11.1)

Abbreviations: N = number of subjects randomized; CI=confidence interval; ¹Estimates based on a logistic regression using treatment and stratification group as factors and baseline value as a covariate; Missing observations were imputed using hybrid imputation (refer to Section 3.2.2 of this review for details); [Source: Statistical Reviewer Analysis; adsl.xpt, advs.xpt, adcovid.xpt, adds.xpt, ada1c.xpt]

3.3 Evaluation of Safety

All safety analyses were conducted on the safety analysis set, which was defined as all randomized subjects who were treated with at least one dose of randomized treatment. The results are summarized in Table 20.

In SURMOUNT-1, the proportion of subjects who discontinued treatment due to adverse events (AEs) was highest in the tirzepatide 10 mg (7.2%). In SURMOUNT-2, the proportion of subjects who discontinued treatment due to AEs was highest in the tirzepatide 15 mg (7.4%).

In both trials, the proportion of subjects who had treatment emergent adverse events (TEAEs) was ranged from 72.0% to 81.8% across all groups including placebo. Most common TEAEs were gastrointestinal events and most events were mild to moderate in severity.

There were 11 deaths in SURMOUNT-1, including 4 deaths in the placebo group, 4 deaths in the tirzepatide 5 mg group, 2 deaths in the tirzepatide 10 mg group, and 1 death in the tirzepatide 15 mg group. All deaths were considered not related to the study drugs by the applicant, except for one death due to hepatic failure in the tirzepatide 5 mg group. There were 2 deaths in SURMOUNT-2 and both deaths were in the tirzepatide 10 mg group. These two deaths were considered not related the study drug by the applicant.

For more details regarding the safety findings including reported deaths, refer to the review from the Medical Reviewer, Dr. Raymond Soccio.

Table 20: Overview of Adverse Events

SURMOUNT-1				
	TZP 5 mg (N=630)	TZP 10 mg (N=636)	TZP 15 mg (N=630)	Placebo (N=643)
Discontinued treatment due to adverse event	30 (4.8%)	46 (7.2%)	40 (6.3%)	21 (3.3%)
Serious adverse event	40 (6.3%)	44 (6.9%)	32 (5.1%)	44 (6.8%)
Treatment emergent adverse event	510 (81.0%)	520 (81.8%)	497 (78.9%)	463 (72.0%)
Death	4 (0.6%)	2 (0.3%)	1 (0.2%)	4 (0.6%)
SURMOUNT-2				
		TZP 10 mg (N=312)	TZP 15 mg (N=311)	Placebo (N=315)
Discontinued treatment due to adverse event		14 (4.5%)	23 (7.4%)	12 (3.8%)
Serious adverse event		18 (5.8%)	27 (8.7%)	23 (7.3%)
Treatment emergent adverse event		242 (77.6%)	222 (71.4%)	239 (75.9%)
Death		2 (0.6%)	0	0

Abbreviations: N=number of subjects randomized and received at least one dose of treatment; cell content shows the number of subjects experiencing at least one event, subjects may be counted more than one category; TZP=tirzepatide; [Source: excerpted from Section 5.2 Table GPHK.5.33. of CSR Section 12 for SURMOUNT-1 and Section 5.2 Table GPHL.5.39 of CSR for SURMOUNT-2]

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The subgroup analysis using an ANCOVA model compared %change from baseline at Week 72 in body weight across treatment groups within subgroups. The LS mean differences and the corresponding 95% CIs are shown in Figure 5 to Figure 12.

There were some random highs and random lows in sample estimates of subgroup treatment effect due to small sample size and large variability for some subgroups. Therefore, we also calculated shrinkage estimates of subgroup treatment effects using a Bayesian hierarchical model based on summary sample estimates. The total variability in the sample estimates is the sum of the within subgroup variability of the sample estimator and the across subgroups variability in underlying/true parameter values. A shrinkage estimate of the subgroup treatment effect, which borrows information from the other subgroups while estimating the treatment effect for a specific subgroup, is a “weighted” average of the sample estimate and overall estimate. We used the same flat prior to derive shrinkage estimates for all subgroups. The Bayesian hierarchical model assumptions are:

For $i=1, 2, \dots$, Y_i represents the observed sample estimate of treatment effect in a subgroup level i , assume $Y_i \sim N(\mu_i, \sigma_i^2)$ where

- σ_i^2 are the observed variance for sample estimates
- $\mu_i \sim N(\mu, \tau^2)$

- $\mu \sim N(0, 40^2)$, $\tau^2 \sim N(0, 10^2, \text{lower}=1e-12)$

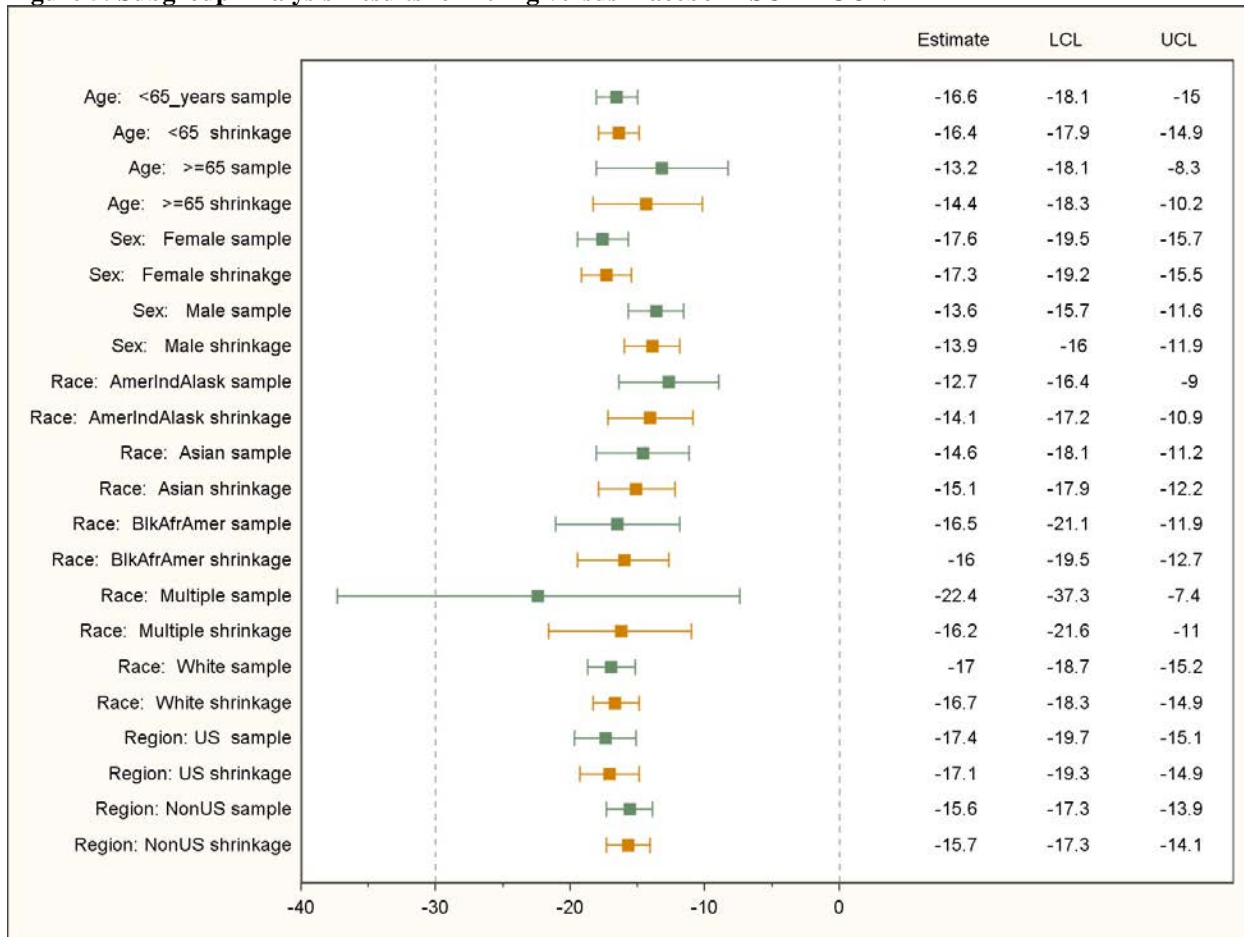
A standard deviation of 40% was chosen for both studies so that the standard deviation was approximately 4 times subject-level standard deviation. Results from both the sample and shrinkage estimates of the treatment effects for the subgroups are presented in Figure 5 to Figure 12.

4.1 Sex, Race, Age, and Geographic Region

For the %change from baseline at Week 72 in body weight, subgroup analyses were performed for age, sex, race and region (Figure 5 and Figure 8). All subgroups reported the upper limit of the 95% confidence interval less than zero, in favor of both tirzepatide 10 mg and 15 mg groups in SURMOUNT-1 and SURMOUNT-2. With shrinkage estimates, the upper limits of the 95% credible intervals were also less than zero, in favor of tirzepatide.

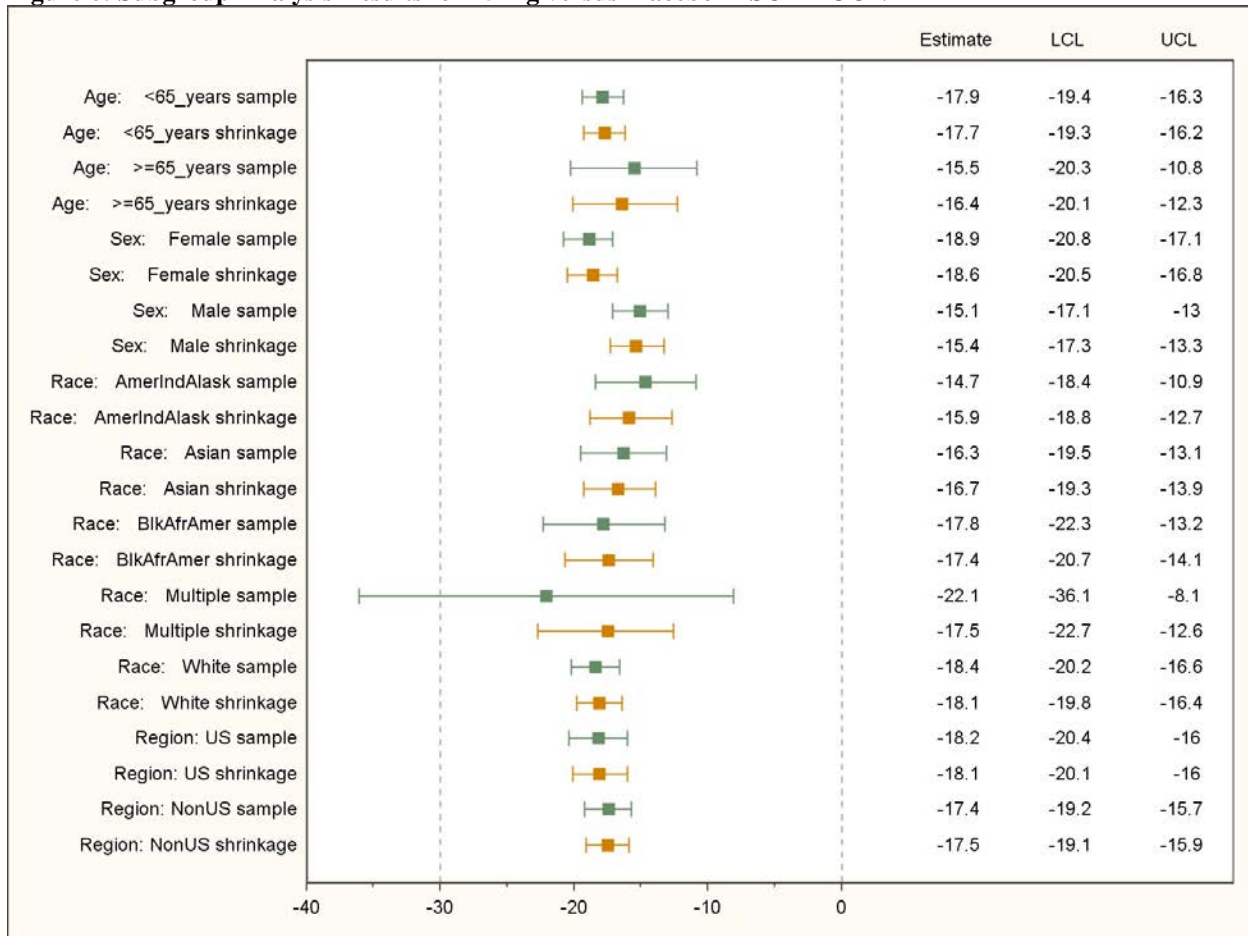
Subgroups with a few subjects were either not shown or the results based on combined categories were shown below. For SURMOUNT-1, “Multiple” category in race consisted of several race categories combined because the number of subjects was too small to obtain reliable estimates.

Figure 5: Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-1



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; AmerIndAlask=American Indian or Alaskan Native; BlkAfrAmer=Black or African American; US=United States; Dotted vertical lines are drawn at zero and at -30; [Source: Statistical Reviewer]

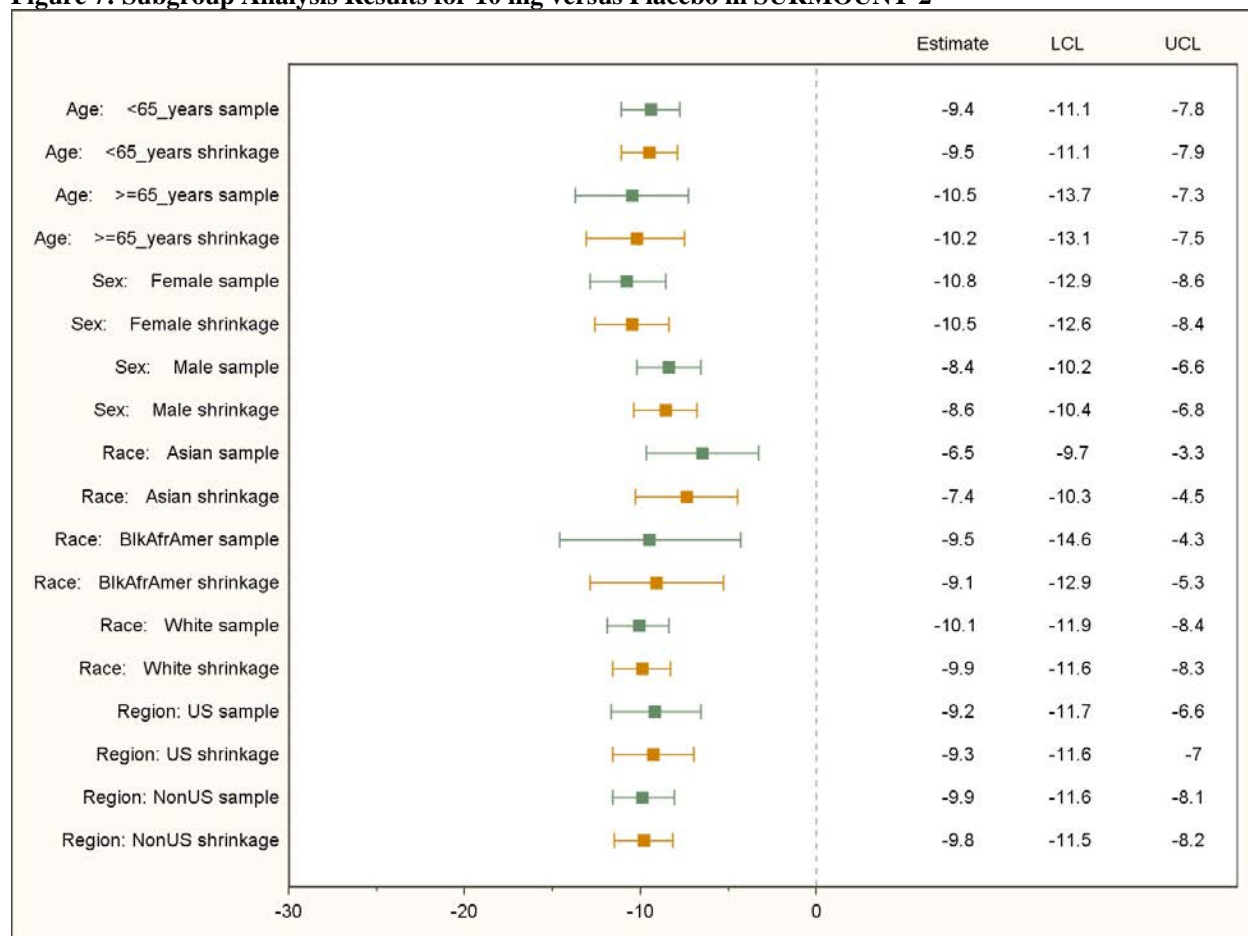
Figure 6: Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-1



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; AmerIndAlask=American Indian or Alaskan Native; BlkAfrAmer=Black or African American; US=United States; Dotted vertical lines are drawn at zero and at -30; [Source: Statistical Reviewer]

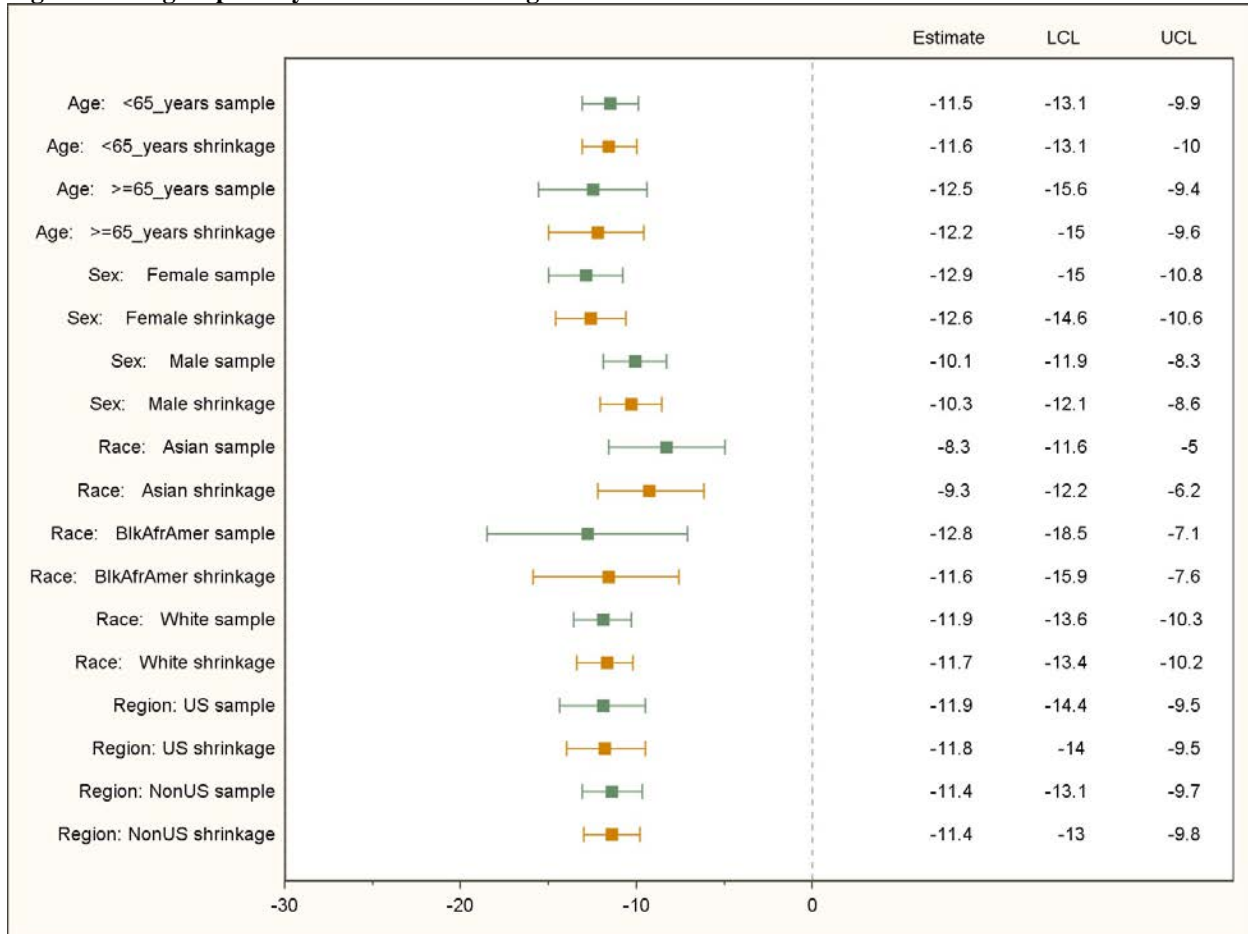
There was a significant interaction effect between sex and treatment in SURMOUNT-1. It appears that weight reduction was more favorable for females than for males in tirzepatide groups. The interaction was quantitative and not qualitative as the study drug was superior to placebo for both males and female. However, it will need further investigation to better understand the observed difference in treatment effects between female and male.

Figure 7: Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-2



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; BlkAfrAmer=Black or African American; US=United States; Dotted vertical line indicates zero; [Source: Statistical Reviewer]

Figure 8: Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-2

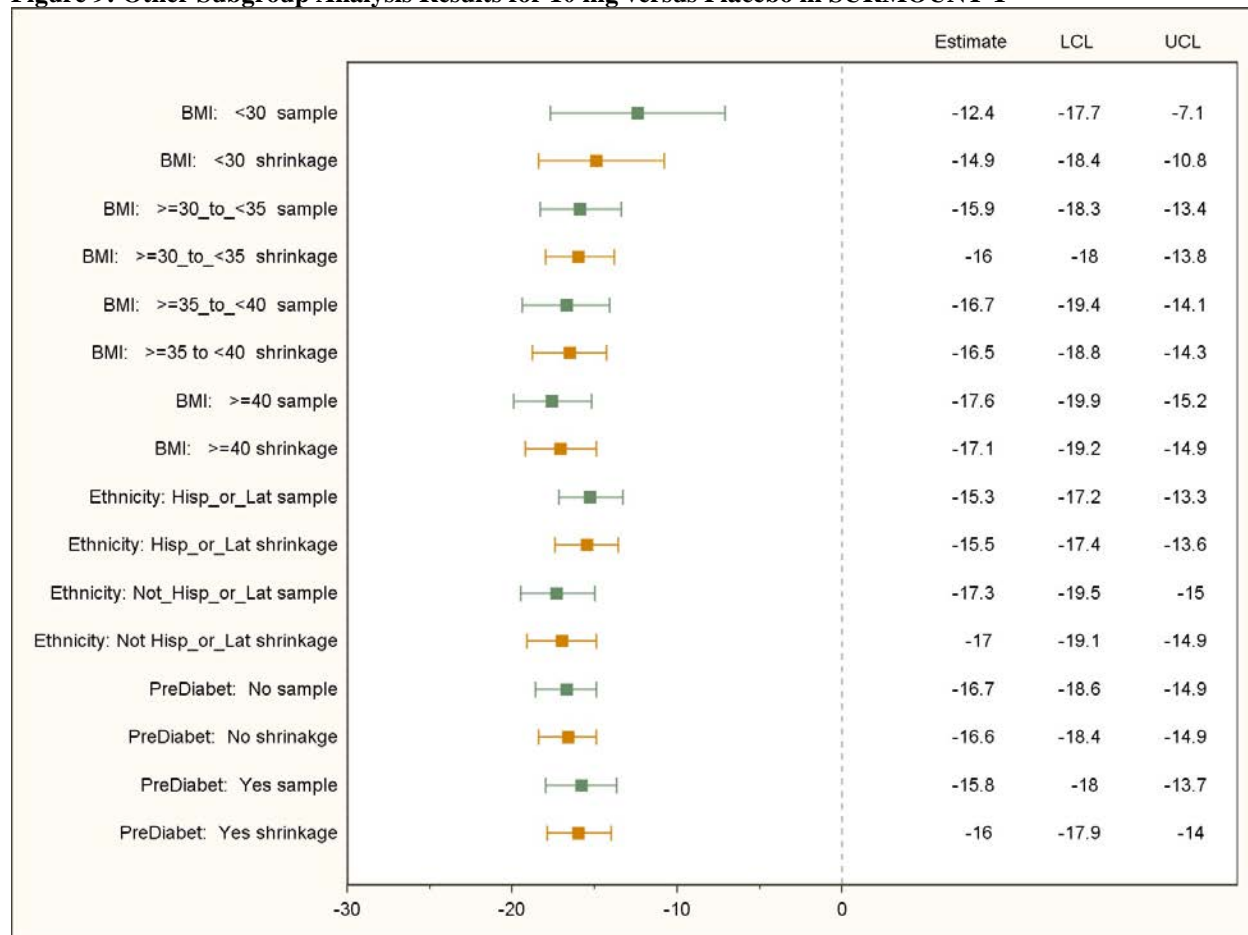


Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; BlkAfrAmer=Black or African American; US=United States Dotted vertical line indicates zero; [Source: Statistical Reviewer]

4.2 Other Special/Subgroup Populations

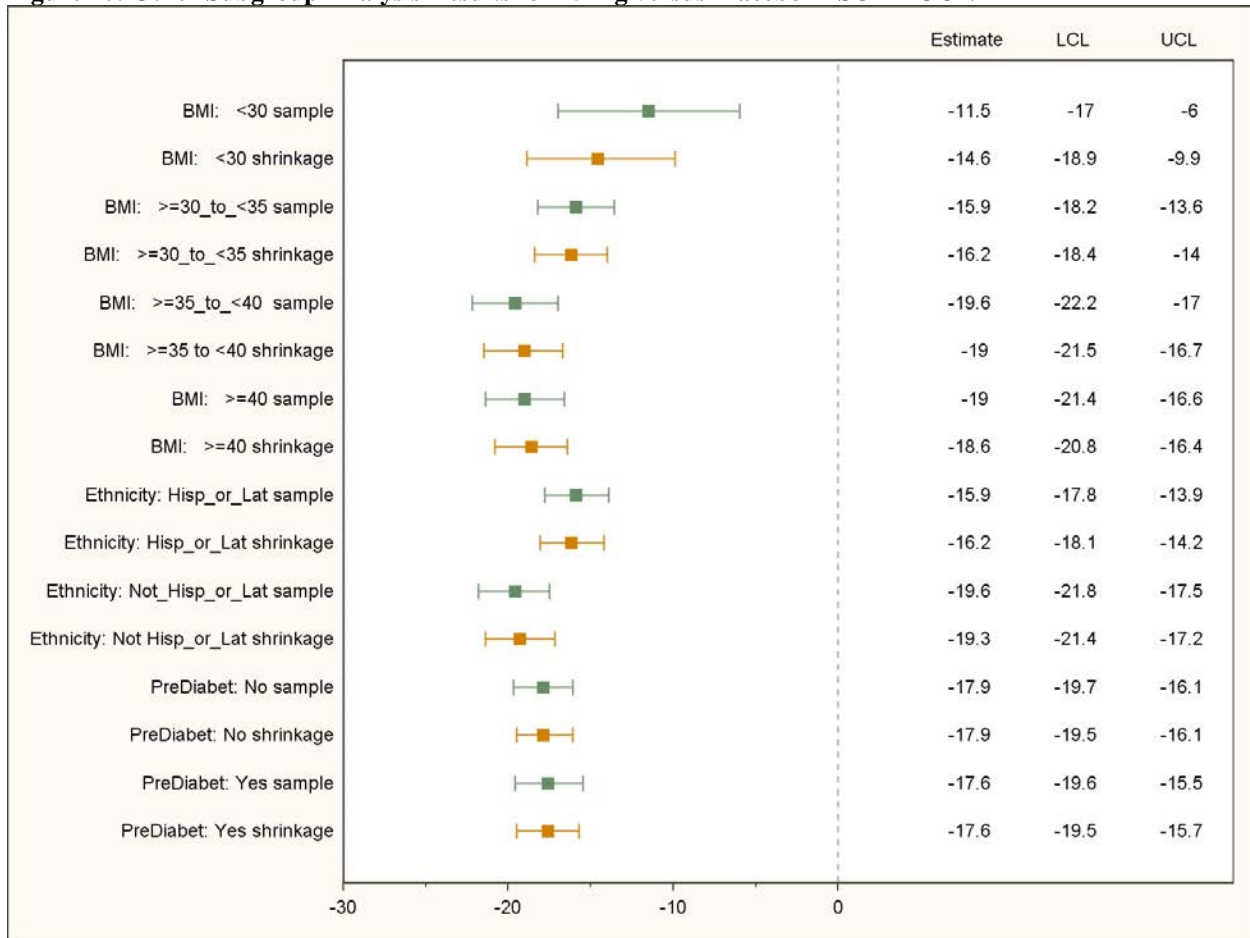
Additional subgroup analyses were performed for baseline BMI (<30, >=30 to <35, >=35 to <40 and >=40), ethnicity and prediabetes status at baseline (SURMOUNT-1) or baseline HbA1c level (SURMOUNT-2), and the results are shown below (Figure 9 to Figure 12). All subgroups reported the upper limits of intervals less than zero, in favor of both tirzepatide 10 mg and 15 mg groups. With shrinkage estimates, the upper limit of the 95% credible interval was also less than zero, in favor of tirzepatide groups.

Figure 9: Other Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-1



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; BMI=body mass index (in kg/m²); Hisp_or_Lat=Hispanic or Latino; PreDiabet=prediabetes status at baseline; Dotted vertical line indicates zero; [Source: Statistical Reviewer]

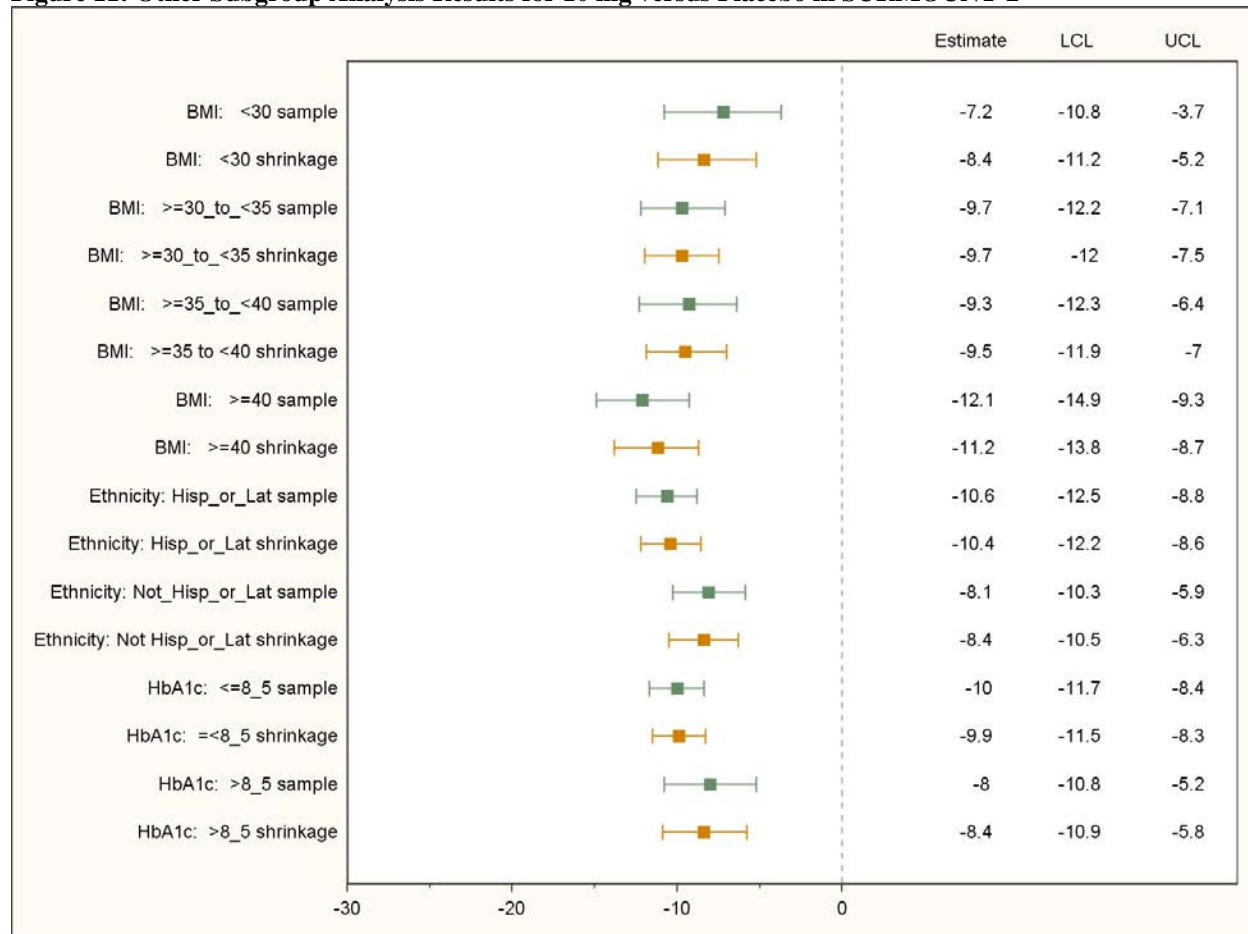
Figure 10: Other Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-1



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; BMI=body mass index (in kg/m²); Hisp_or_Lat=Hispanic or Latino; PreDiabet=prediabetes status at baseline; Dotted vertical line indicates zero; [Source: Statistical Reviewer]

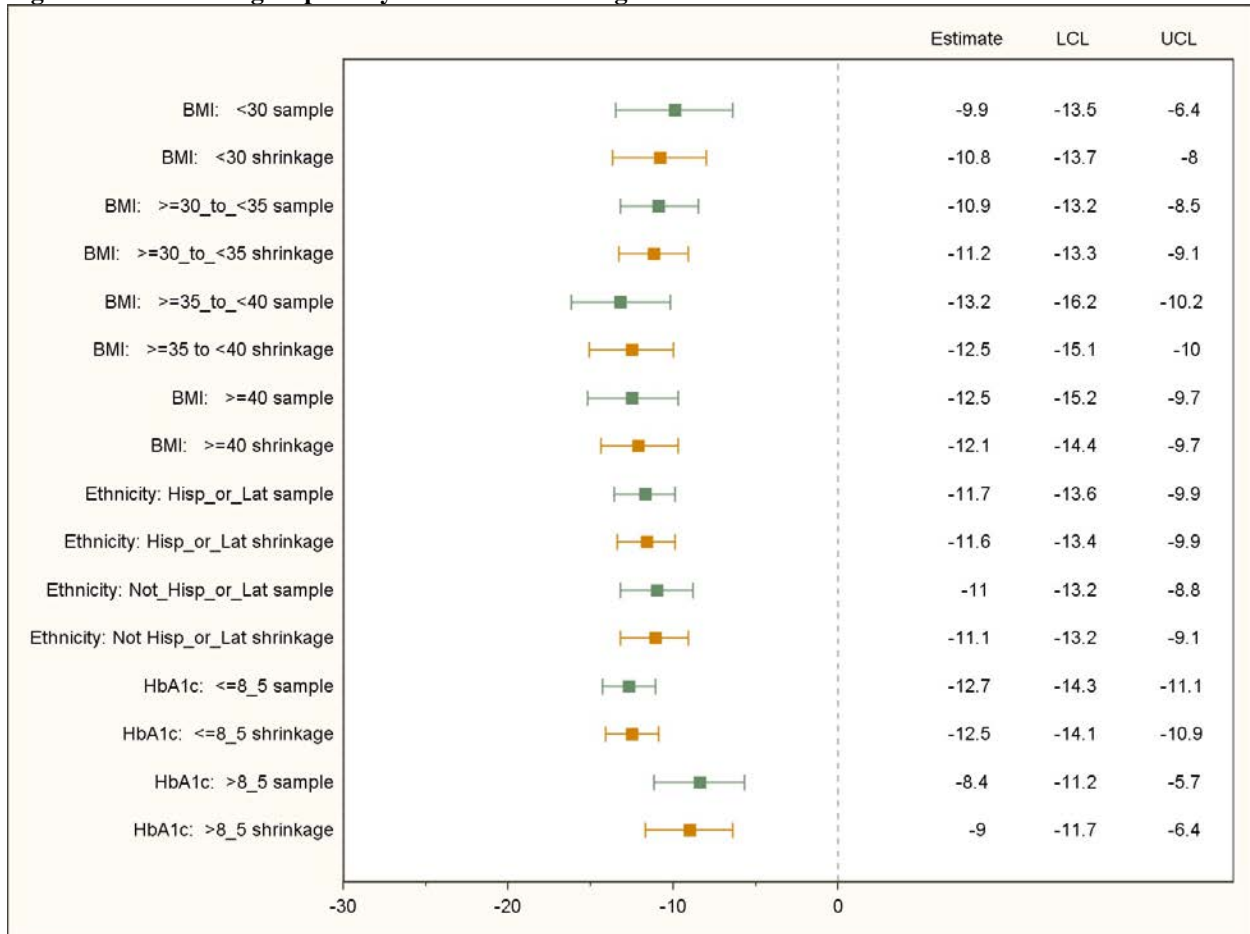
There were significant quantitative interaction effects between tirzepatide groups and subgroups such as BMI groups and ethnicity in SURMOUNT-1. It appears that weight reduction was more favorable for high baseline BMI than for low baseline BMI in tirzepatide groups. In addition, it appears that weight reduction was more favorable for Not Hispanic or Latino group than for Hispanic or Latino group in tirzepatide groups. However, it will need further investigation to better understand the treatment effect on different BMI groups and ethnicity.

Figure 11: Other Subgroup Analysis Results for 10 mg versus Placebo in SURMOUNT-2



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; BMI=body mass index (in kg/m²); Hisp_or_Lat=Hispanic or Latino; HbA1c:<=8_5=baseline HbA1c<=8.5%; HbA1c:>8_5=baseline HbA1c>8.5%; Dotted vertical line indicates zero; [Source: Statistical Reviewer]

Figure 12: Other Subgroup Analysis Results for 15 mg versus Placebo in SURMOUNT-2



Note: Subgroup analysis results are shown for the %change from baseline at Week 72 in body weight; Sample estimates are shown with the corresponding 95% confidence interval (in green) and shrinkage estimates are shown with the corresponding 95% credible interval (in orange); LCL: lower confidence (or credible) limit; UCL: upper confidence (or credible) limit; BMI=body mass index (in kg/m²); Hisp_or_Lat=Hispanic or Latino; HbA1c:<=8_5=baseline HbA1c ≤8.5%; HbA1c: >8_5=baseline HbA1c>8.5%; Dotted vertical line indicates zero; [Source: Statistical Reviewer]

There was a significant quantitative interaction effect between baseline HbA1c group and tirzepatide groups in SURMOUNT-2. It appears that weight reduction was more favorable for subjects with baseline HbA1c ≤8.5% than for subject with baseline HbA1c >8.5% in tirzepatide groups.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

There were no major statistical issues that would impact or change the overall conclusion. Across trials, the proportion of missing data ranged from 4.8% to 10.5% for the tirzepatide

groups. The proportion of missing data for the placebo group was 21.6% in SURMOUNT-1 and 11.1% in SURMOUNT-2. The sensitivity analyses using the pre-specified approaches supported the robustness of the primary efficacy results. Both primary and confirmatory secondary efficacy endpoints were statistically significant, in favor of the tirzepatide groups.

In SURMOUNT-1, there was a significant interaction effect between sex and treatment. It appears that weight reduction was more favorable for females than for males in both tirzepatide groups. In addition, there were significant interaction effects between tirzepatide groups and subgroups such as BMI groups and ethnicity. It appears that weight reduction was more favorable for high baseline BMI than for low baseline BMI in tirzepatide groups, and it appears that weight reduction was more favorable for Not Hispanic or Latino group than for Hispanic or Latino group in tirzepatide groups. In SURMOUNT-2, there was a significant interaction effect between baseline HbA1c group and tirzepatide groups. It appears that weight reduction was more favorable for subjects with baseline HbA1c $\leq 8.5\%$ than for subject with baseline HbA1c $> 8.5\%$ in tirzepatide groups. All the interactions are quantitative instead of qualitative. However, it will need further investigation to better understand what lead to the interaction effects.

5.2 Collective Evidence

The primary analysis showed statistically significant treatment effect in body weight reduction (%) at Week 72 in both tirzepatide dose groups (10 mg and 15 mg). The proportion of subjects achieved at least 5% body weight reduction was also statistically significant, in favor of both dose groups. Sensitivity analyses supported the robustness of the primary efficacy results. Confirmatory secondary endpoints were consistently in favor of tirzepatide.

5.3 Conclusions and Recommendations

The collective evidence from the submitted data demonstrated efficacy of tirzepatide in the study population. We recommend approval for the proposed indication based on findings from the submitted results.

5.4 Labeling Recommendations (as applicable)

Reviewing of labeling is still ongoing while this statistical review is finalized.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

KYUNGHEE K SONG
10/06/2023 10:53:02 AM

FENG LI
10/10/2023 10:55:33 AM



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

NDA NO.:	217806
DATE RECEIVED BY OB:	06/02/2023
DRUG NAME:	Zepbound (tirzepatide)
SPONSOR:	Eli Lilly and Company
REVIEW FINISHED DATE:	09/27/2023
CMC STATISTICAL REVIEWER:	Tengfei Li, Ph.D.
NAME OF REQUESTOR:	Arati Kamath, Ph.D., RPM CDER/OND/ORO/DROCHEN

Secondary Reviewer:

Meiyu Shen, Ph.D., Team Leader, Mathematical Statistician, CDER/OTS/OB/DB VI

Concur:

Yi Tsong, Ph.D., Division Director, DBVI, CDER/OTS/OB/DB VI

Distribution:

NDA 217806
CDER/OTS/OB/DBVI/ Yi Tsong
CDER/OTS/OB/DBVI/ Meiyu Shen
CDER/OTS/OB/DBVI/ Atiar Mohammad Rahman
CDER/OND/ORO/DROCHEN/ Arati Kamath
CDER/OPQ/OBP/DBRRII/ Faruk Sheikh
CDER/OPQ/OBP/DBRRII/ Harold Dickensheets

TABLE OF CONTENTS

1. Executive Summary	4
2. Introduction and Background	4
3. FDA CMC Statistical Reviewers' Comments on the Applicant's Approach	5
4. FDA CMC Statistical Reviewers' Independent Data Analysis.....	6
4.1 Data preparation	6
4.2 Model Setup for CP estimation.....	6
4.3 Data analysis results.....	7
5. Conclusions and Recommendations.....	8
6. References.....	9

LIST OF TABLES

Table 1. Comparison of Disease State Cut Point (DSCP) provided by applicant and DSCP calculated by statistical reviewers..... 4

Table 2 Summary of the Applicant’s Disease State Cut Points (DSCP) 5

Table 3. Comparison of sponsor’s disease specific cut point (DSCP) and statistical reviewer’s DSCP..... 7

1. Executive Summary

Eli Lilly and Company (“The applicant”) submitted an NDA 217806 for tirzepatide for chronic weight management (CWM), in which the applicant estimated the ADA assay disease state cut points (DSCP).

A general comment on the DSCP for CWM is that the applicant did not differentiate the variability between samples and the variability between analysts. In addition, the applicant ignored the correlation of two observed values of the same sample measured by a different analyst.

Based on CMC statistical reviewer’s independent CP analysis, the DSCP are factor 1.35 for Tier 1, 22.7% inhibition for Tier 2a, 21.6% inhibition for Tier 2b, 19.2% inhibition for Tier 2c, 9.5% neutralization for Tier 4a and 6.0% neutralization for Tier 4b. Table 1 compares Disease State Cut Point (DSCP) provided by the applicant and DSCP calculated by statistical reviewers.

Table 1. Comparison of Disease State Cut Point (DSCP) provided by applicant and DSCP calculated by statistical reviewers

Assay Parameter	Attribute	Values for (p, q)	DSCP calculated by the p% lower confidence limit for the q-th quantile		
			Applicant’s DSCP from T2DM	Applicant’s DSCP from CWM	Reviewer’s DSCP from CWM
Tier 1*	CP factor	(90, 95)	1.22	1.30	1.35
Tier 2a	% Inhibition	(80, 99)	30.4%	28.8%	22.7%
Tier 2b	% Inhibition	(80, 99)	14.5%	24.0%	21.6%
Tier 2c	% Inhibition	(80, 99)	18.1%	19.7%	19.2%
Tier 4a	% Neutralization	(80, 99)	10.8%	10.5%	9.5%
Tier 4b	% Neutralization	(80, 99)	6.6%	4.7%	6.0%

*Data in Tier 1 including the average per sample in Tier 2a were used in the reviewer’s statistical analysis. The sponsor added 2 observations per sample in Tier 2a as independent and identical observations to their Tier 1 data for DSCP calculation.

Considering obvious differences in DSCP for all tiers between CWM and T2DM, statistical reviewers would not recommend DSCP from T2DM for use in CWM population given all data are available for determining DSCP specifically for CWM.

Introduction and background are given in **Section 2**. More detailed comments are in **Section 3**. Details of the CMC statistical reviewer’s independent statistical analysis are in **Section 4**. The applicant’s proposal is summarized in **Table 2**. The CMC reviewer’s results are summarized in **Table 3**.

2. Introduction and Background

The applicant submitted an NDA 217806 for tirzepatide for chronic weight management (CWM) on November 17, 2022 for the first sequence and for the 2nd and final sequence on May 8, 2023. Office of Biotechnology Products (OBP) was consulted to review the immunogenicity results from the completed phase 3 studies. We sent an immunogenicity information request (IR) on April 4,

2023 requesting the data used for determining the disease cut point for each of the anti-drug antibody (ADA) assays. The applicant provided the data for each of the ADA assays for Tiers 1, 2a, 2b, 2c, 4a and 4b on April 21, 2023.

In the study for CWM, the applicant removed outliers using the mean and a 1.5*IQR (interquartile range), and test for normality for the raw data and transformed data. The applicant estimated the cut points using a one-sided parametric tolerance limit if the raw data or converted data is normally distributed and a one-sided non-parametric tolerance limit otherwise. Specifically, they took the lower 90% confidence limit of the 95th percentile ($p=95%$, $1 - \alpha=90%$) as the estimator for Tier 1, and the lower 80% confidence limit of the 99th percentiles ($p=99%$, $1 - \alpha=80%$) for Tier 2a, 2b, 2c, 4a and 4b. **Table 2** shows the estimated disease state cut points for CWM.

Table 2 Summary of the Applicant’s Disease State Cut Points (DSCP)

Assay Parameter	DSCP from T2DM (The applicant’s proposal)	DSCP from CWM
Tier 1	Factor: 1.22	Factor: 1.30
Tier 2a	30.4%	28.8%
Tier 2b	14.5%	24.0%
Tier 2c	18.1%	19.7%
Tier 4a	10.8%	10.5%
Tier 4b	6.6%	4.7%

The CMC statistical team in Office of Biostatistics (OB) was asked to evaluate the applicant’s statistical analysis approach for the cut points for the binding assays (Tiers 1, 2a, 2b and 2c) and Nab assays (Tiers 4a and 4b) for CWM, and also to comment on the proposal using cut points from T2DM for analyzing serum samples from CWM subjects.

In **Section 3**, CMC statistical reviewers provide comment on the applicant’s approach. In **Section 4**, CMC statistical reviewers present an independent analysis.

3. FDA CMC Statistical Reviewers’ Comments on the Applicant’s Approach

We sent an Information Request (IR) on August 10, 2023 asking for clarifications of statistical methods, data management and related issues ([link](#)). The applicant responded on August 17, 2023 ([link](#)).

The applicant confirmed the Tier 1 samples were from CWM study. They also confirmed the use of a one-sided lower tolerance bound to provide 90% confidence of at least a 5% false positive rate for Tier 1 and 80% confidence for at least a 1% false positive rate for Tiers 2a, 2b and 2c.

We had the following comments:

Firstly, it follows the FDA Guidance for Immunogenicity Testing of Therapeutic Protein Products – Developing and Validating Assays for Anti-Drug Antibody Detection, that the screening assay cut-point is to use an 90% one-sided lower confidence interval for the 95th percentile and

confirmatory assay cut-point is to use an 80% to 90% one-sided lower confidence interval for the 99th percentile.

Secondly, the applicant did not differentiate the variability between samples and the variability between analysts. Furthermore, the applicant ignored the correlation of two observed values of the same sample measured by a different analyst.

Thirdly, the parametric approach may be inappropriate for Tier 1 as the data was not normally distributed even after log transformation.

Fourthly, regarding the proposal using cut points from T2DM for analyzing serum samples from CWM subjects, comparing only the DSCPs from the two studies is not sufficient to conclude the appropriateness of replacing DSCP in CWM with DSCP in T2DM. On top of that, we need to confirm the samples are comparable from the two studies which means the samples from the two studies are balanced. However, we do not have such data for the characterization of the sample data from the two studies.

In **Section 4**, CMC statistical reviewers evaluated the DSCP, using tolerance interval-based method considering random effects on samples and analysts.

4. FDA CMC Statistical Reviewers' Independent Data Analysis

4.1 Data preparation

We excluded outliers detected by the applicant, and then conducted normality check. If the data were normally distributed, we would use parametric approach. The nonparametric approach would be used otherwise.

In the CP estimation for Tier 1, we combined data from Tier 1 and average of two observations from each sample from Tier 2a. In the CP estimation for Tiers 2a, 2b, 2c, 4a and 4b, we used negative-control-normalized data and differentiated the variability between samples and variability between analysts.

The FDA Guidance for Immunogenicity Testing of Therapeutic Protein Products – Developing and Validating Assays for Anti-Drug Antibody Detection, recommends that the screening cut-point is determined by the 90% lower confidence limit for upper 95th percentile and the confirmatory cut-point is determined by the 80% to 90% lower confidence limit for the upper 99th percentile. Here the CP is defined as the 90% one-sided lower confidence limit for the 95th percentile for Tier 1, and the 80% one-sided lower confidence limit for the 99th percentile for Tiers 2a, 2b, 2c, 4a and 4b.

4.2 Model Setup for CP estimation

For Tier 1 with one observation per sample, we used non-parametric tolerance intervals (PCTLDEF = 4 for SAS PROC UNIVARIATE).

For other tiers, we differentiated the variability between samples and variability of the same sample repeatedly measured by different analysts. Therefore, Y_{ijk} can be modeled by the following two-way REM model for each Tier that

$$Y_{ijk} = \mu + s_i + r_j + \varepsilon_{ijk}, \quad i = 1, \dots, n, j = 1, \dots, m, k = 1, 2.$$

Here, Y_{ijk} is the observed value for the k th replicate of sample i measured by analyst j . n is the number of samples per Tier excluding outliers. m is the number of analysts per Tier. μ is the fixed overall mean, s_i is a random effect for sample i , r_j is a random effect induced by analyst j , and ε_{ijk} is the error term of the model. s_i , r_j , ε_{ijk} are assumed to be mutually independent and $s_i \sim N(0, \sigma_s^2)$, $r_j \sim N(0, \sigma_r^2)$, and $\varepsilon_{ijk} \sim N(0, \sigma_e^2)$.

The detailed DSCP estimation is given in Shen, Dong and Tsong (2015) and Shen and Dai (2021).

4.3 Data analysis results

Tier 1 has 2500 samples, one observation per sample. Tier 2a, 2b, 2c, 4a and 4b have 200 samples, each measured by different analysts, respectively. There were five analysts for Tier 2a, 2b, 2c, 4a and 4b in total. The results are shown in **Table 3**.

Table 3. Comparison of sponsor's disease specific cut point (DSCP) and statistical reviewer's DSCP

Assay Parameter	Attribute	Values for (p, q)	DSCP calculated by the p% lower confidence limit for the q-th quantile		
			Applicant's DSCP from T2DM	Applicant's DSCP from CWM	Reviewer's DSCP from CWM
Tier 1*	CP factor	(90, 95)	1.22	1.30	1.35
Tier 2a	% Inhibition	(80, 99)	30.4%	28.8%	22.7%
Tier 2b	% Inhibition	(80, 99)	14.5%	24.0%	21.6%
Tier 2c	% Inhibition	(80, 99)	18.1%	19.7%	19.2%
Tier 4a	% Neutralization	(80, 99)	10.8%	10.5%	9.5%
Tier 4b	% Neutralization	(80, 99)	6.6%	4.7%	6.0%

*Data in Tier 1 including the average per sample in Tier 2a were used in the reviewer's statistical analysis. The sponsor added 2 observations per sample in Tier 2a as independent and identical observations to their Tier 1 data for DSCP calculation.

In CP estimation for Tier 1, we used non-parametric tolerance interval since most of the samples only had one observation, and the data were not normally distributed even after log transformation. The DSCP is 1.35, higher than 1.30 and 1.22 (that equals to the Applicant's DSCP for CWM and T2DM, respectively).

In terms of Tier 2a, the data follows a normal distribution, so no transformation was considered. We used two-way REM method taking into consideration the variability between samples and between analysts. We had a much smaller DSCP estimation of 22.7% than 28.8% and 30.4% (that equals to the Applicant's DSCP for CWM and T2DM, respectively).

For Tier 2b, the data does not follow a normal distribution. So, we took a log transformation. The same two-way REM method was employed for CP estimation. We had a DSCP estimation of 21.6%, smaller than 24% (the Applicant's DSCP for CWM), but much larger than 14.5% (the Applicant's DSCP for T2DM).

For Tier 2c, the data does not follow a normal distribution. So, we took a log transformation. The same two-way REM method was employed for CP estimation. The DSCP is 19.2%, similar to the applicant's DSCP of 19.7% for CWM.

For Tier 4a, the data does not follow a normal distribution. We did not take a log transformation as the data had negative values. The same two-way REM method was employed for CP estimation. We had a smaller DSCP estimation of 9.5% than 10.5% and 10.8% (that equals to the Applicant's DSCP for CWM and T2DM, respectively).

For Tier 4b, the data does not follow a normal distribution. We did not take a log transformation as the data had negative values. The same two-way REM method was employed for CP estimation. The DSCP is 6.0%, larger than the applicant's DSCP of 4.7% for CWM.

Considering obvious differences in DSCP for all tiers between CWM and T2DM, statistical reviewers would not recommend DSCP from T2DM for use in CWM population given all data are available for determining DSCP specifically for CWM.

5. Conclusions and Recommendations

Based on CMC statistical reviewer's independent analyses, considering the variability between samples and analysts, the DSCP is bigger for Tier 1 and Tier 4b, and smaller for Tiers 2a, 2b, 2c and 4a, compared to the applicant's DSCP for CWM. The DSCP are factor 1.35 for Tier 1, 22.7% inhibition for Tier 2a, 21.6% inhibition for Tier 2b, 19.2% inhibition for Tier 2c, 9.5% neutralization for Tier 4a and 6.0% neutralization for Tier 4b.

Considering obvious differences in DSCP for all tiers between CWM and T2DM, statistical reviewers would not recommend DSCP from T2DM for use in CWM population given all data are available for determining DSCP specifically for CWM.

6. References

- [1] FDA Guidance for Industry, Immunogenicity Testing of Therapeutic Protein Products – Developing and Validating Assays for Anti-Drug Antibody Detection, January 2019, <https://www.fda.gov/media/119788/download>
- [2] Shen, M., & Dai, T. (2021). Statistical methods of screening cut point determination in immunogenicity studies. *Bioanalysis*, 13(7), 551-563.
- [3] Shen, M., Dong, X., & Tsong, Y. (2015). Statistical evaluation of several methods for cut-point determination of immunogenicity screening assay. *Journal of biopharmaceutical statistics*, 25(2), 269-279.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

TENGFEI LI
09/27/2023 11:10:05 AM

MEIYU SHEN
09/27/2023 12:14:26 PM

YI TSONG
09/27/2023 12:17:01 PM